

Caroline Willkommen

Eine Logik von Handlungen und Überzeugungen



λογος

Die Open-Access-Stellung der Datei erfolgte mit finanzieller Unterstützung des Fachinformationsdiensts Philosophie (<https://philportal.de/>)



Dieses Werk ist lizenziert unter der Creative Commons Attribution 4.0 Lizenz CC BY-SA (<https://creativecommons.org/licenses/by-sa/4.0/>). Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z.B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.



DOI: <https://doi.org/10.30819/3484>

Caroline Willkommen

Eine Logik von Handlungen und Überzeugungen

Logos Verlag Berlin



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Dissertation der Ruhr-Universität Bochum, Fakultät für Philosophie und Erziehungswissenschaft.

© Copyright Logos Verlag Berlin GmbH 2013

Alle Rechte vorbehalten.

ISBN 978-3-8325-3484-4

Logos Verlag Berlin GmbH
Comeniushof, Gubener Str. 47,
10243 Berlin
Tel.: +49 (0)30 / 42 85 10 90
Fax: +49 (0)30 / 42 85 10 92
<http://www.logos-verlag.de>

Es ist möglich, sich Arten von Welten vorzustellen, die von der Welt, wie wir sie kennen, ganz verschieden sind.

Peter Strawson in *Die Grenzen des Sinns*

Here is what I consider one of the biggest mistakes of all in modal logic: concentration on a system with just one modal operator. The only way to have any philosophically significant results in deontic logic or epistemic logic is to combine these operators with: Tense operators (otherwise how can you formulate principles of change?); the logical operators (otherwise how can you compare the relative with the absolute?); the operators like historical or physical necessity (otherwise how can you relate the agent to his environment?); and so on and so on.

Dana Scott in *Advice on Modal logic*

Inhaltsverzeichnis

Einleitung	1
1 Eine Logik für Handlungssätze - Stit-Theorie	7
1.1 Branching Time	12
1.2 Stit-Operatoren: Achievement oder Deliberativ	14
1.3 Das Problem der konfligierenden Historie	23
2 Zustände, Sachverhalte und Ereignisse	27
2.1 Zustände und Sachverhalte	33
2.2 Ein Zustand und das Vorliegen des Zustandes	51
2.3 Zustände und Ereignisse	54
3 Handlungen oder die Frage „Actions as Events“	65
3.1 Handlung und das darunter liegende Ereignis	66
3.2 Klassifikation von Ereignissen	72
3.3 Mentale Zustände und mentale Ereignisse	83
4 Überzeugungen, Wünsche und Absichten	109
4.1 Überzeugungen und Wünsche	113
4.2 Wünsche und Absichten	158
4.3 Absichten und Handlungen	185
5 Von Stit und <i>BDI</i> zu <i>bdi-stit</i>	217
5.1 Von <i>BDI</i> zu <i>bdi-stit</i>	218
5.2 Die <i>bdi-stit</i> Logik	228
5.3 Eine Generalisierung von Stit-Theorie und $B^{KD45}D^{KD}I_{CTL}^{KD}$	238
6 Vollständigkeit und Erweiterungen von <i>bdi-stit</i>	243
6.1 Vollständige Axiomatisierung der <i>bdi-stit</i>	244
6.2 Ein Tableauekalkül für <i>bdi-stit</i>	258

6.3	Temporale Operatoren in der <i>bdi-stit</i> Logik	283
6.4	<i>bdi-stit</i> _{int} — Interaktionen zwischen Stit- und Intensionsoperatoren .	295
7	Zusammenfassung	309

Einleitung

Der Titel dieser Arbeit lautet „Eine Logik von Handlungen und Überzeugungen“. Zum einen impliziert dieser Titel, dass sich hinter dem Ausführen von Handlungen und dem Haben von Überzeugungen logische Prinzipien finden lassen. Zum anderen schließt sich die Frage an, wie diese zu beschreiben sind. Ausgangspunkt dieser Arbeit ist die erste Implikation zu bejahen und, um die aufgeworfene Frage zu beantworten, sich den folgenden Fragestellungen im Einzelnen zu widmen: Welcher Begriff von Handlung bzw. Handlungsbeschreibung liegt zugrunde? Wie werden Überzeugungen bzw. die Zuschreibung von Überzeugungen zu einem Akteur aufgefasst? Wie wird beides einheitlich in einem logischen System beschrieben? Der Beantwortung dieser Fragen komme ich in den folgenden Kapiteln nach.

Zu Beginn werde ich die Handlungslogik vorstellen, die als Grundlage der zu konstruierenden Logik gewählt wird, die dann ihrerseits nicht nur Handlungsoperatoren, sondern auch Überzeugungsoperatoren beinhalten wird. In der Literatur finden sich zwei Arten zur Beschreibung von Handlungslogiken. Entweder werden Handlungen mittels Modaloperatoren beschrieben. Oder es werden Ereignis- bzw. Handlungsvariablen eingeführt. Diese Variablen werden mit dem Akteur in Relation gesetzt und es wird beschrieben, dass bestimmte Sachverhalte nach einem solchen Ereignis bzw. nach einer solchen Handlung bestehen. Letztere dynamisch logische Systeme lassen sich z.B. in [47, 87, 145] finden. Sie führen Ereignisvariablen für Basishandlungen ein und ordnen zu, nach welcher Basishandlung welche Sachverhalte bestehen. Hat eine solche Logik noch den Akteur im Blick? Einer Basishandlung liegt ein Ereignis zugrunde. Aber in welchem Verhältnis steht der Akteur, der durch das Ausführen der Basishandlung einen bestimmten Sachverhalt erreicht, zu dem Erreichen des Sachverhaltes?

Dass diese Frage unbeantwortet bleibt, ist einer der Gründe, sich gegen eine solche Konzeption von Handlungen bzw. Handlungsbeschreibungen durch dynamische Logiken zu entscheiden.

Eine Handlungsbeschreibung enthält im Wesentlichen zwei Dinge. Zum einen wird ein während des Ereignisses der Handlung bestehender Sachverhalt beschrieben. Zum

Einleitung

anderen wird zum Ausdruck gebracht, dass der Akteur für das Bestehen dieses Sachverhaltes gesorgt hat. Eine Beschreibung, die beides im Kern enthält, findet sich in modalen Handlungslogiken.

Die Ursprünge von modalen Handlungslogiken sehen viele bereits bei Anselm von Canterbury so z.B. [19, S.19] bzw. [125]. Anselm schreibt, zitiert nach [76, S.124]:

„Omne igitur ‘facere’ dici potest aut ‘facere esse’ aut ‘facere non esse’“.

Alles, was als eine Handlung beschrieben werden kann, ist derart, dass man entweder davon spricht, dass etwas getan wurde, so dass etwas der Fall ist, oder es ist etwas getan worden, so dass etwas nicht der Fall ist.

In der zweiten Hälfte des 20. Jahrhunderts wurden sogenannte modale *Bringing-About*-Logiken diskutiert, für einen Überblick siehe [124, 125]. Ein modaler Handlungsoperator verbindet einen Akteur und eine Variable für einen Sachverhalt und symbolisiert somit einen neuen Sachverhalt, der ausdrückt, dass der Akteur den durch den Operator gebundenen Sachverhalt hervorgebracht hat. Ende der achtziger Jahre wurde von Perloff und Belnap [18] als auch von Horty [83] die Stit-Theorie entwickelt. Das Akronym *Stit* steht hierbei für: ‘See to it that’. Im Deutschen wird von ‘der Akteur sorgte dafür, dass’ gesprochen. Es wurden verschiedene Semantiken und Axiomatisierungen für Stit-Operatoren vorgeschlagen. In der vorliegenden Arbeit wird ein solcher Stit-Operator verwendet, um Handlungen zu beschreiben. Dafür werden kurz die Vor- und Nachteile dieses Operatores im Vergleich mit anderen Stit-Operatoren diskutiert, um zu begründen, warum die Wahl auf diesen Operator fiel.

Wenn die erste Frage nach der Beschreibung der Handlungen in einem logischen System somit beantwortet werden kann, stellt sich die Frage nach der Beschreibung der mentalen Zustände, wie Überzeugungen, wenn eine Beschreibung mit Handlungen in einem einheitlichen System das Ziel ist. Auch wenn der Titel lediglich Überzeugungen erwähnt, so sind Absichten und Wünsche ebenfalls mentale Zustände, die bei Handlungen eine Rolle spielen. Vor allem Absichten scheinen mit den Handlungen eines Akteurs eng verknüpft zu sein.

Ich werde daher nicht nur Überzeugungen, sondern auch Wünsche und Absichten mit modalen Operatoren beschreiben. Hierbei stellt sich wiederum die Frage, welche Operatoren in welcher Semantik sich am besten eignen, um diese mentalen Zustände zu beschreiben. Es finden sich vielfältige Vorschläge. In der Regel werden mentale Zustandsoperatoren durch eine relationale Modallogik beschrieben [81, 64, 66]. Aber auch klassische Modallogiken wurden von vielen zur Beschreibungen für mentale Zustände bereits verwendet, wie z.B. für Überzeugungen in [56], für Wünsche in [139]

und für Absichten in [78]. Durchgesetzt hatten sich in den neunziger Jahren vor allem die BDI-Logiken. Diese beschreiben alle drei Operatoren, sowohl Überzeugungs-, als auch Wunsch- und Absichtoperator in relationalen Semantiken [64, 66, 114]. Diese BDI-Logiken bildeten die Grundlage für viele Erweiterungen, z.B. zur Beschreibung von Emotionen [1], aber auch um in Verbindung mit dynamischen Operatoren Handlungen zu beschreiben [119].

Die Plausibilität einer Beschreibung von Überzeugungs-, Wunsch- und Intentionoperator mittels einer relationalen Semantik zu diskutieren, wird einen Großteil dieser Arbeit ausmachen. Ich werde Axiome, die in diesen Logiken postuliert worden sind, kritisch betrachten. Die Frage ist, ob sich diese Axiome motivieren lassen und ob es möglicherweise eine Alternative zu einer Beschreibung in BDI-Logiken gibt.

Sobald geklärt sein wird, wie der jeweilige mentale Zustand durch Modaloperatoren angemessen beschrieben werden kann, stellt sich die Frage, welche Beziehungen es zwischen Handlungen und dem Vorliegen dieser Zustände gibt. Und welche lassen sich in dieser Beschreibung mittels Operatoren sinnvoll wiedergeben?

Beziehungen zwischen den mentalen Zuständen und Handlungen können nur über verschiedene Interaktionsschemata dargestellt werden. Wenn Interaktionen zwischen den Operatoren schematisch dargestellt werden sollen, stellt sich die Fragen nach den Möglichkeiten, wie ein solches Interaktionsschema zu begründen ist? Zum einen könnte ein Schema auf begriffliche Zusammenhänge zwischen den einzelnen Zuständen und Handlungen zurückgeführt werden. Diese begrifflichen Zusammenhänge können aufgrund normativer oder deskriptiver Aspekte bestehen. Ebenso besteht die Möglichkeit, solche Interaktionsschemata dadurch zu motivieren, dass es von einem externen Standpunkt aus rational ist, dass Handlungen nur mit bestimmten mentalen Zuständen einhergehen bzw. dass bestimmte mentale Zustände sich gegenseitig bedingen oder ausschließen sollten. Welche Schemata sich sinnvoll begründen lassen und welche Verbindungen zwischen Handlungen und Überzeugungen sich nicht in einem logischen System beschreiben lassen, wird in Kapitel 4 ausführlich diskutiert werden.

Im letzten Kapitel wird dann Schritt für Schritt ein logisches System konstruiert, das auf der Stit-Theorie aufbaut und um modale Operatoren zur Beschreibung der mentalen Zustände ergänzt wird. Dabei werden die zuvor motivierten Interaktionsschemata nach und nach berücksichtigt, und die Unterschiede zu BDI-Logiken werden kenntlich gemacht.

Der Aufbau der Arbeit gestaltet sich daher wie folgt. Im ersten Kapitel werden nicht nur die Definitionen der Handlungsoperatoren angegeben, sondern auch die Motivation kurz skizziert, Handlungsbeschreibungen auf Stit-Formulierungen aufzubauen.

Einleitung

Der erste Abschnitt des ersten Kapitels stellt somit die Thesen vor, die Belnap et al. bei der Beschreibung von Handlungen mittels Stit-Operatoren annehmen [19]. In Abschnitt 1.2 werden dann zwei unterschiedliche Interpretationen daraufhin untersucht, welche geeigneter ist, eine Handlung eines Akteurs zu beschreiben. Dies geschieht bereits im Hinblick darauf, dass auch mentale Akte bzw. mentale Vorgänge, die keinen Akt beschreiben, Teil des logischen Systems werden, wenn die Stit-Theorie um mentale Zustandsoperatoren erweitert wird. Den Abschluss des ersten Kapitels bildet die Diskussion eines Kritikpunktes an der Stit-Theorie, dass die Vergangenheit von Geschichtsverläufen zwar durch feste Situationen beschrieben wird, dass jedoch konfligierende Sachverhalte in einer gemeinsamen Vergangenheit von verschiedenen Geschichtsverläufen vorliegen können. Es wird in Abschnitt 1.3 darauf hingewiesen, dass dieses Problem sich lösen lässt, so dass es kein Grund ist, Stit-Theorie an sich in Frage zu stellen.

In Kapitel 2 wird eine These der Stit-Theorie aufgegriffen. Belnap [15] behauptet, dass Handlungsbeschreibungen nicht auf Ereignissen basieren können, die dem Akteur irgendwie zugeschrieben werden. Eine solche Beschreibung von Handlungen charakterisiert einen Akteur als reinen Statist in einem Stück, in dem er eigentlich die Hauptrolle spielen sollte. Belnap [15] kritisiert Davidson [48, Essay 6] scharf für seine Beschreibung von Handlungen. Ich werde einen Kritikpunkt Belnaps untersuchen, inwieweit dies zutreffend ist. Um jedoch zu erklären, worin das Problem besteht, ist es notwendig sich über die Verwendung bestimmter Begriffe zu einigen. Was versteht Davidson und was versteht Belnap unter einer Handlungsbeschreibung? Und was versteht jeder von beiden unter einer Handlung und einem Ereignis? Für den Begriff der Beschreibung wird der Begriff des Sachverhaltes bzw. der Proposition zentral sein. Nicht auf Beschreibungsebene, sondern konkret in der Welt vorliegend sind hingegen Zustände und Ereignisse. Diese drei Begriffe werde ich in Kapitel 2 definieren und voneinander abgrenzen.

Nachdem in Kapitel 2 diese grundlegenden Begriffe erörtert worden sind, wird in Kapitel 3 auf diesen Begriffen aufbauend den Fragen nachgegangen, was eine Handlungsbeschreibung ist und was einer Handlung zugrunde liegt. Das Fazit des Kapitels wird sein, dass jeder Handlung ein Ereignis zugrunde liegt. Aber dieses Ereignis ist nur anhand der Beschreibungen des Ereignisses, des bestehenden Sachverhaltes, dem Akteur als Handlung zuzuschreiben, vgl. Abschnitt 3.1. In Abschnitt 3.2 wird festgehalten, dass es aufgrund der Sachverhalte, die während eines Ereignisses bestehen, verschiedene Klassen von Ereignisbeschreibungen geben kann. Es wird hierbei auf Vendlers Unterscheidung von Verbklassen eingegangen, die jeweils verschiedene Er-

eignisse beschreiben, vgl. [151]. Dabei wird deutlich, dass es Ereignisbeschreibungen gibt, die eine Handlungsbeschreibung eher zulassen als andere, siehe Abschnitt 3.2. Manche Ereignisse können einem Akteur nur dadurch als Handlungen zugeschrieben werden, dass der Zustand, der nach der Handlung vorliegt, als vom Akteur hervorgebracht beschrieben wird. Insbesondere trifft dies auf mentale Ereignisse zu, die durch das Bilden bzw. Hervorbringen von mentalen Zuständen charakterisiert sind. Nichtsdestotrotz kann auch im Falle eines mentalen Aktes, z.B. das Bilden des mentalen Zustandes einer Intention unter Ausübung von Kontrolle, eine Beschreibung mittels Stit-Operator deutlich machen, dass es sich um einen Akt des Akteurs und nicht nur um einen Vorgang im Akteur handelt. In Abschnitt 4.3 wird daher diskutiert, welcher Operator zur Darstellung mentaler Akte am geeignetsten ist und welche der drei zur Diskussion stehenden Typen mentaler Zustände durch mentale Akte gebildet werden und welche nicht.

Im Anschluss wird in Kapitel 4 diskutiert, inwieweit mentale Zustände im Falle einer Handlung vorliegen müssen oder aufgrund von Rationalitätsforderungen vorliegen sollten. Aber nicht nur Handlungen machen das Vorliegen bestimmter mentaler Zustände erforderlich, sondern auch die mentalen Zustände untereinander können in notwendigen Beziehungen zueinander stehen. So wird in Abschnitt 4.1 festgehalten, dass eine Handlungsmotivation sich in den Wünschen und Überzeugungen bzw. den konativen und kognitiven Einstellungen des Akteurs finden lässt. Es ist jedoch nicht möglich zu postulieren, dass es Wünsche bzw. Überzeugungen mit konkretem, die Handlung betreffenden Gehalten gibt, die bei jeder Handlung vorliegen müssten. Somit ist ein Interaktionsschema, das für Handlungen notwendige Überzeugungen oder Wünsche eines Akteurs postuliert, nicht begründbar. Da es sich bei der Bildung von Überzeugungen und Wünschen, wie wir sehen werden, nicht um mentale Akte handeln kann, ist ein Schema, das das Vorliegen als Akt beschreibt, zu verneinen. Weitere mögliche Interaktionen werden in Abschnitt 4.1 ebenfalls angesprochen. Den Abschluss bildet die Diskussion, welche Semantik einem Überzeugungsoperator zugrunde gelegt werden sollte und welche Schemata diesen Operator axiomatisieren.

In Abschnitt 4.2 werden Wünsche und Absichten voneinander abgegrenzt und es wird begründet, warum es sich um unabhängige, mentale Zustände handelt. Hierbei spielt die Aufgabe, welche einem Wunsch und welche einer Absicht bei der Handlungsmotivation zukommt, eine entscheidende Rolle. Im Anschluss daran wird sowohl dem Wunsch- als auch dem Absichtsoperator eine der Aufgabe entsprechende Semantik zugewiesen und verschiedene Schemata diskutiert, die den jeweiligen Operator in einer Logik axiomatisieren sollten. Den Abschluss des Kapitels 4 bildet die Zusam-

Einleitung

menfassung der Interaktion zwischen Handlungs- und Intentionsooperatoren. Zu der Motivation der Handlungsoperatoren und den damit verbundenen Axiomenschemata wird auf die Literatur verwiesen, da diese aus einer bestehenden Logik übernommen wurden. Im Mittelpunkt der Diskussion stehen die von mir vorgeschlagenen Interaktionsschemata zwischen Handlungs- und Absichtsoperator.

Im Anschluss an diese Motivation der Beschreibung von Handlungen und mentalen Zuständen durch Operatoren wird in Kapitel 5 dieser Ansatz von bereits in der Literatur diskutierten Beschreibungen abgegrenzt. Insbesondere auf die unterschiedliche Handhabung der BDI-Operatoren in den BDI-Logiken und in der von mir vorgeschlagenen *bdi-stit* Logik wird näher eingegangen, siehe Abschnitt 5.1. In Abschnitt 5.2 wird die allgemeine *bdi-stit* Logik vorgestellt. Als allgemein wird sie bezeichnet, weil sie noch nicht auf alle Interaktionsaxiome festgelegt ist, die in Kapitel 4 motiviert wurden. Dadurch kann eine BDI-Logik von der allgemeinen *bdi-stit* Logik generalisiert werden. Inwieweit Generalisierung zu verstehen ist, wird in Abschnitt 5.3 angeführt.

In Kapitel 6 wird die allgemeine *bdi-stit* Logik als vollständig und entscheidbar mittels einer vollständigen Axiomatisierung nachgewiesen. Die Vollständigkeit lässt sich ebenfalls anhand eines Tableauealküls zeigen, siehe Abschnitt 6.2. In Abschnitt 6.3 wird ein Ausblick auf eine mögliche Erweiterung der *bdi-stit* Logik um temporalen Operatoren gegeben. Es wird gezeigt, dass diese Erweiterung möglich ist, dass dies jedoch mit gewissen Schwierigkeiten hinsichtlich einer vollständigen Axiomatisierung bzw. der Frage nach der Entscheidbarkeit verbunden ist. In Abschnitt 6.4 wird dann der Intentionsooperator spezifiziert, so dass aus der allgemeinen *bdi-stit* Logik, die *bdi-stit_{dint}* Logik wird. Dieser Intentionsooperators erfüllt die in Abschnitt 4.3 eingeführten Interaktionsaxiome zwischen dem Handlungs- und dem Intentionsooperator. Den Schluss der Arbeit bildet der Vollständigkeitsbeweis der *bdi-stit_{dint}* Logik.

1 Eine Logik für Handlungssätze - Stit-Theorie

Auf die Frage, was eine Handlung als solche beschreibt, gibt es viele und doch keine Antworten, wie Belnap et al. festhalten:¹

„We find, to our dismay, that we are no closer now to a linguistic litmus test for agency than was Aristotle.“ [19, S.3]

Ob ein Satz eine Handlung beschreibt oder nicht, ist anhand des Satzes nicht immer erkennbar. Dennoch versuchen sich Belnap et al. an einem solchen Lackmustest und dies nicht ohne Erfolg. In diesem Abschnitt werde ich die Thesen benennen, die hinter diesem Lackmus Test stehen, und im Anschluss die Umsetzung in die modallogische Sprache der Stit-Theorie aufzeigen.

Um uns der Idee hinter der Stit-Theorie zu nähern, möchte ich mit drei Beispielsätzen beginnen, die [19] entlehnt sind. Welchen Unterschied in Anbetracht der Frage, ob es sich um Beschreibungen von Handlungen Alexanders handelt, erkennt man in den folgenden Sätzen:

1. Alexander bestieg den Thron von Makedonien 336 v. Chr.
2. Alexander hatte den Thron von Makedonien ab 336 v. Chr. inne.
3. Alexander beerbte den Thron von Makedonien 336 v. Chr.

Während intuitiv zugestimmt wird, dass es sich bei Ersterem um eine Beschreibung einer Handlung von Alexander handelt, kann im zweiten Fall festgestellt werden, dass mit diesem Satz keine Handlung ausgedrückt wird. Der zweite Satz beschreibt einen

¹Im Zuge dieser Arbeit verwende ich Anführungszeichen, wie folgt: Doppelte Anführungszeichen verwende ich um innerhalb des Fließtextes deutlich zu machen, dass es sich um zitierte Begriffe, die ich nicht näher erläutere, oder direkt um Zitate handelt. In einfache Anführungszeichen setze ich Bezeichnungen bzw. Beispielsätze, z.B. der Sachverhalte ‘es regnet’, das Ereignis ‘Tod Caesars’, der Begriff ‘Zustand’ aber auch ‘ α sorgt dafür, dass’ oder ‘Alexander besteigt den Thron’.

1. Eine Logik für Handlungssätze - Stit-Theorie

Zustand, der ab 336 v. Chr. vorlag. Ob Alexander den Thron aktiv bestieg oder er ihm lediglich zufiel, geht aus diesem Satz nicht hervor. Mit dem dritten Satz wird nun ebenfalls keine Handlung ausgedrückt. Er unterscheidet sich allerdings von dem zweiten, da er keinen Zustand beschreibt, sondern den Vorgang, dass Alexander 336 v. Chr. auf den Thron kam. Um zu hinterfragen, ob hinter den beschriebenen Sachverhalten sich Handlungen verbergen, schlägt die Stit-Theorie die Betrachtung folgender Sätze vor:

- 1'. Alexander sorgte dafür, dass er den Thron von Makedonien 336 v. Chr bestieg.
- 2'. Alexander sorgte dafür, dass er den Thron von Makedonien ab 336 v. Chr. inne hatte.
- 3'. Alexander sorgte dafür, dass er den Thron von Makedonien 336 v. Chr beerbte.

Offensichtlich ist diesen Sätzen ein Sinn zuzuschreiben, aber ist dies noch der, der die Bedeutung der obigen Sätze erfasst? Belnap et al. halten daher fest [19, S.7]:

Vollständigkeitsthese Für jeden Satz Q lässt sich ein grammatikalischer, sinnvoller Satz bilden, indem dem Satz Q 'α² sorgt dafür, dass' vorangestellt wird.³

Nachdem festgehalten ist, dass für jeden Satz ein sinnvoller, korrespondierender Satz existiert, der mit 'α sorgt dafür, dass' eingeleitet wird, stellt sich die Frage, was mit dieser Einleitung ausgedrückt wird. Die Sätze aus obigem Beispiel machen den Unterschied deutlich. Drückte der 2. Satz noch einen Zustand aus, so beschreibt Satz 2'. eine Handlung, und zwar diejenige Handlung Alexanders, die dazu führte, dass der Zustand, der in Satz 2. wiedergegeben wird, eintritt. Für die Sätze 3. und 3'. gilt selbiges. Der Vorgang, der in Satz 3. beschrieben wird, kann Alexander auch ohne sein Zutun geschehen sein. In Satz 3'. hingegen wird deutlich, dass es sich bei diesem Vorgang, um eine Handlung Alexanders handelt. Belnap et al. postulieren dies in folgender These [19, S.7]:

Akteursthese Ein Satz, der mit der Konstruktion 'α sorgt dafür, dass' eingeleitet wird, drückt eine Handlung von α aus und zwar unabhängig von der im Nebensatz ausgedrückten Proposition.

Es stellt sich die Frage nach der Bedeutung der Sätze. Die Wahrheit des Satzes 1'. lässt

²Griechische Buchstaben α, β, \dots symbolisieren beliebige, jedoch feste Akteure.

³Falls Q ein Hauptsatz ist, ist es im Deutschen notwendig, die Satzstellung von Haupt- auf Nebensatz zu korrigieren. Bsp: 'er hebt den Arm' wird zu 'er sorgt dafür, dass er den Arm hebt.' Für die Imperativsätze siehe [19, S.10, 15f]. Interrogativsätze werden behandelt wie Behauptungssätze.

sich auf die Wahrheit des Satzes ‘Alexander bestieg den Thron ...’ zurückführen und umgekehrt. Währenddessen ist es in den jeweiligen zweiten und dritten Sätzen nicht klar, ob die Aussage, mit Apostroph gekennzeichnet, genau dann wahr ist, wenn die korrespondierende Aussage wahr ist. So folgt aus der Wahrheit der Aussagen hinter 2’ und 3’ jeweils die Wahrheit der Aussagen 2. und 3. Jedoch gilt die Umkehrung nicht. Es wäre möglich, dass Alexander den Thron von Makedonien ab 336 v. Chr. inne hatte, jedoch nicht dafür sorgte, dass es der Fall war. Diese Besonderheit im Falle von Handlungssätzen halten Belnap et al. in folgender These fest, [19, S.7]:

These der Stit-Paraphrase Ein Satz Q drückt genau dann eine Handlung eines Akteurs α aus, wenn der Satz ‘ α sorgt dafür, dass Q ’ eine geeignete Paraphrase für Q ist. Approximativ heißt dies, dass ein Satz Q eine Handlung von α ausdrückt, falls Q und ‘ α sorgt dafür, dass Q ’ äquivalent sind.

Demzufolge kann, immer wenn mit einem Satz Q eine Handlung ausgedrückt werden soll, der Satz mit ‘der Handelnde sorgt dafür, dass Q ’ begonnen werden, ohne dass sich der Wahrheitsgehalt durch Hinzufügen des Präfixes ändert. Die These, das betonen Belnap et al., stellt keine Reduktion oder Definition dessen dar, was eine Handlung sein soll. Sie gibt lediglich die Möglichkeit wieder, an einem Satz zu erkennen, dass er tatsächlich eine Handlung beschreibt. Entscheidet diese Paraphrase auch darüber, wann etwas keine Handlung ist? Sicherlich nicht - die Paraphrase soll lediglich darauf hinweisen, dass der beschriebene Sachverhalt dem Subjekt, das für etwas sorgt, eine aktive Rolle zuschreibt. Es heißt nicht, dass, wenn die Wahrheitsbedingungen dieser Paraphrase und des Ausgangssatzes voneinander abweichen, der zum Ausgangssatz korrespondierende Sachverhalt im Allgemeinen keine Handlung sein kann. Ich berufe mich wieder auf obigen Beispielsatz 2’. ‘Alexander sorgt dafür, dass er den Thron inne hat.’ Dieser Satz beschreibt eine Handlung. Das Komplement des Stit-Satzes ‘Alexander hat den Thron inne’ hingegen drückt keine Handlung aus. Es lässt jedoch die Möglichkeit offen, dass die mit ‘Alexander hat den Thron inne’ ausgedrückte Proposition durch eine Tatsache wahr gemacht wird, die auch eine andere Proposition wahr machen würde, die eine aktive Vorgehensweise Alexanders beinhaltet. Die Stit-Paraphrasierung ist eine exzellente Möglichkeit zu testen, ob es sich bei einer Beschreibung eines Sachverhaltes, um eine Handlung eines bestimmten Akteurs handelt. Jedoch schließt eine ungewohnt klingende Paraphrase nicht aus, dass der durch die Paraphrase beschriebene Sachverhalt nicht auch Teil einer Handlung sein kann.⁴

⁴Eine explizite Definition der Begriffe ‘Sachverhalt’, ‘Ereignis’ usw. erfolgt in Kapitel 2.

1. Eine Logik für Handlungssätze - Stit-Theorie

Auf die Unterscheidung zwischen einem Komplement und dem Satz weisen Belnap et al. hin [19, S.5,12f]. Zum Beispiel ist mit dem Komplement des Satzes ‘es ist verboten, dass Köche auf der Kommandobrücke sind’ der Teilsatz ‘Köche sind auf der Kommandobrücke’ gemeint. Da Verbote sich nicht auf Zustände beziehen, d.h., es nicht verboten ist, dass etwas der Fall ist, sondern nur verboten wird, dass etwas von jemandem getan wird, so dass dieser Fall eintritt („impersonal oughts“), muss sich das Komplement auf eine Handlung beziehen, die von jemanden verrichtet bzw. nicht verrichtet werden soll. Der semantische Inhalt des Komplementes, die ausgedrückte Proposition, ist in Belnaps et al. Terminologie nicht „agentive“, da sie nur einen Zustand wiedergibt, jedoch ist das Komplement als „agentive“ zu verstehen, da es das Komplement eines Verbotes ist.⁵ Belnap et al. plädieren daher dafür, dass in bestimmten Kontexten das Komplement immer agentive sein sollte, was durch eine Stit-Paraphrase jederzeit ausdrückbar ist.

Damit ist nicht gemeint, dass es sich um eine Handlung von dem Subjekt oder den Subjekten (‘Köchen’) handelt, die im Komplement erwähnt werden. Deutlich wird dies, wenn ‘Köche’ durch ‘Hunde’ oder ‘Sprengstoffe’ ersetzt werden [19, S.12]. Belnap et al. verdeutlichen mit diesem Beispiel, dass obiger Satz zwar grammatikalisch korrekt gebildet, aber dennoch in gewisser Weise unklar ist. Eine Handlung wird implizit beschrieben, jedoch tritt kein Handlungsverb auf, das einen Vorgang als Handlung kennzeichnet. Unklar ist auch, für wen das Verbot besteht und welche Handlung verboten ist. Eine allgemeine Anwendung des Präfixes ‘es ist verboten, dass’ nicht auf einen Satz, sondern nominell auf eine Tätigkeitsbeschreibung führt zu keiner Klärung. So könnte der Satz zu ‘es ist verboten, sich als Koch auf die Kommandobrücke zu stellen’ oder ‘es ist verboten, Köche auf der Kommandobrücke zu platzieren’ umformuliert werden. Jedoch bleiben beide Formulierungen vage über den Handelnden. Wenn ‘Köche’ durch ‘Hunde’ ersetzt werden, wird dies deutlich. Das Subjekt, für das die Ausführung der Handlung verboten ist, bleibt unklar. Belnap et al. geben hiermit eine Satzkonstruktion an, deren Komplemente agentive sind, obwohl die Propositionen, die durch die Komplemente beschrieben werden, keine Handlungen beinhalten. Es bedarf der Benennung des Akteurs und der ausdrücklichen Kennzeichnung, dass es sich hierbei um eine Handlung des Akteurs handelt. Beispielsweise, ‘es ist verboten, dass Hans dafür sorgt, dass Köche auf der Kommandobrücke sind’.

Somit schafft eine Stit-Paraphrase des Komplementes hierfür eine vollständige Klärung und zwar auch dann, wenn das Komplement nicht durch die angegebene Propo-

⁵Eine adäquate Übersetzung für ‘agentive’ scheint es im Deutschen nicht zu geben. Daher verwende ich ‘agentive’ ohne Übersetzung.

sition, sondern durch die Satzkonstrukte, die es vervollständigt, als agentive genommen wird. Dies gilt z.B. für deontische Kontexte, aber auch für Imperativkonstruktionen und propositionale Einstellungen, z.B. Intentionen [19, S.13]. Auf den Punkt der Intentionen gehen Belnap et al. nicht genauer ein. Ich werde in Abschnitt 4.3 darauf zurückkommen. Ein Beispiel sei an dieser Stelle genannt: Das Komplement der Satzkonstruktion ‘Alexander beabsichtigt, dass er König der Makedonier wird’ ist agentive, nicht weil ‘König zu sein’ eine Handlung beschreibt, sondern weil die Satzkonstruktion ‘Alexander beabsichtigt’, die durch ‘er wird König der Makedonier’ komplettiert wird, agentive ist. D.h. die Sätze ‘Alexander beabsichtigt, dass er König der Makedonier wird’ und ‘Alexander beabsichtigt, dass er dafür sorgt, dass er König der Makedonier wird’ sind äquivalent.

Belnaps et al. Vorschlag ist, die Stit-Paraphrasierung generell anzuwenden, um Verwirrungen zu vermeiden, ob mit einem Satz tatsächlich eine Handlung ausgedrückt werden soll oder nicht und wer der Akteur ist. Wenn ein Aussage- oder Teilsatz eine Handlung beschreibt, entweder aufgrund der Satzkonstruktion, die er vervollständigen soll, oder weil die ausgedrückte Proposition eine Handlung wiedergibt, oder wenn ein Imperativsatz gegeben ist, schlägt die Stit-Theorie vor, eine Stit-Paraphrase des Satzes zu verwenden [19, S.15]:

Stit-Normalform These In Untersuchungen von Satzkonstruktionen, die in einem Komplement des Satzes eine Handlung angeben wollen, werden Verwechslungen vermieden, ohne dass ein Informationsverlust auftritt, wenn diejenigen Komplemente durchweg durch Stit-Paraphrasen ersetzt werden.

Das Ergebnis einer vollständigen Ersetzung der Komplemente durch die Stit-Paraphrasen bezeichnen Belnap et al. als Stit-Normalform [19, S.15]. Es wird in der Stit-Normalform nicht nur klar und deutlich, dass eine Handlung beschrieben wird, sondern auch wer der Akteur ist.

„A Stit sentence, [α stit Q], since it displays its agent and appropriate declarative complement publicly and obviously, is the appropriate picture of the underlying agentive partly because it remains recognizably the same in any and every context.“ [19, S.16]

Anhand dieser Paraphrase wird deutlich, wie sich eine Handlung eines Akteurs α in einer formalen Sprache der Logik ausdrücken lässt. Belnap et al. verwenden für jeden Akteur α jeweils einen Operator, der auf eine Aussage angewandt wird, α stit Q. In [19] beschreiben sie Handlungen durch verschiedene Modaloperatoren, die auf aussagenlogische Formeln angewandt werden, und in Strukturen interpretiert werden, die

1. Eine Logik für Handlungssätze - Stit-Theorie

sich verzweigende Zeit modellieren. Im Folgenden werde ich kurz darauf eingehen, was eine Branching Time Struktur (BT) ist, welche Operatoren Belnap et al. vorschlagen und warum ich mich für den deliberative Stit-Operator zur Beschreibungen von Handlungen entschieden habe.

1.1 Branching Time

Ein sich verzweigendes Modell der Zeit, welches die Vergangenheit als eindeutig bestimmt, jedoch die Zukunft als unbestimmt annimmt, ist geeignet, um Handlungen modellieren zu können. Die Idee ist, dass zum jeweils aktuellen Zeitpunkt der Handelnde den weiteren Verlauf durch seine unterschiedliche Wahl, Handlungen auszuführen, mitbestimmt. Die getroffene Wahl repräsentiert somit die Handlung.

Von Lamport [91] wurde festgehalten, dass sowohl lineare als auch sich verzweigende Zeitmodelle ihre Daseinsberechtigung haben. Zum einen gibt es unterschiedliche Sichtweisen, ob die Welt deterministisch oder indeterministisch ist und welchen ontologischen Status zukünftige Geschichtsverläufe haben. Zum anderen ist es möglich, Operatoren zu beschreiben, die auf den jeweils anderen Modellen nicht in äquivalenter Form beschreibbar sind. Weitere Vorteile, die den Nachteil der höheren Komplexität einer Branching Time Logik gegenüber einer linearen Zeitlogik überwiegen sollten, wurden von Halpern und Emerson [52] festgehalten.

Bereits Prior [111] entwickelte eine Branching Time Struktur, die eine indeterministische Zeitvorstellung repräsentierte. Betrachtet man indeterministische Modelle, so gibt es zwei Herangehensweisen. Die eine von Prior nach William von Ockham benannte Vorstellung erlaubt es, kontingenten Aussagen in zukünftigen Momente verschiedene Wahrheitswerte zuzuschreiben. Es werden verschiedene Geschichtsverläufe in der Zukunft angenommen, so dass es möglich ist, in verschiedenen Geschichtsverläufen zukünftiger Momente ein und derselben Aussage verschiedene Wahrheitswerte zuzuordnen. Die andere nach Charles Sanders Peirce benannte Herangehensweise verneint die Möglichkeit, Aussagen in der Zukunft unterschiedliche Wahrheitswerte zuzuweisen. Es wird festgelegt, dass die Behauptung, dass etwas stattfinden wird, notwendigerweise wahr oder falsch ist. Es gibt somit keinen Unterschied zwischen, dass etwas der Fall sein wird und dass etwas möglicherweise der Fall sein wird.

Zwei Zeitlogiken von Modellen sich verzweigender Zeit sind CTL , eine Peircesche Logik, und CTL^* , eine Ockhamistische Erweiterung von CTL , die von Emerson, Halpern und Sistla in den 80er Jahren entwickelt wurden [51, 53]. Sowohl CTL als auch CTL^* sowie $PCTL^*$, eine um Vergangenheitsoperatoren erweiterte CTL^* [115], sind

vollständig axiomatisiert.

Für die in der Stit-Theorie untersuchten Handlungsoperatoren ist es möglich, die sich verzweigende Struktur so einfach wie möglich zu halten. Da ich keine Axiomatisierung der zugrunde liegenden Zeitlogik benötige, sondern mich auf die Stit-Operatoren und in den Abschnitten 6.1 und 6.4 auf die Operatoren konzentriere, die die Zuschreibung verschiedener, mentaler Zustände eines Akteurs beschreiben, übernehme ich die vollständige Ockhamistische Logik. Im Gegensatz zu CTL^* , oder im Englischen als „bundled“ (gebündelt) bezeichneten Ockhamistischen Logiken, für die Axiomatisierungen bereits existieren, verzichte ich auf die explizite Angabe von solch gebündelten, aktuell möglichen Geschichtsverläufen zu jedem Zeitpunkt, sondern lasse alle auftretenden Geschichtsverläufe zu. Insbesondere lasse ich eine Interpretation der auftretenden Aussagebuchstaben über beliebige Paare von Geschichtsverläufen und Momenten zu. Dies hat zwar zur Folge, dass es mir leider nicht möglich ist, für die Zeitoperatoren eine vollständige Axiomatisierung anzugeben, hält jedoch die zugrunde liegende Struktur einfach.⁶

Die vollständige Ockhamistische Logik, in der Form wie sie Prior vorschlug, lässt sich, wie folgt, definieren. Es werden zwei zukunftsgerichtete Operatoren F und G eingeführt, wobei F für „irgendwann einmal in der Zukunft“ und G für „immer in der Zukunft“ steht [111]. In späteren Logiken werden die Operatoren durch X und U ersetzt bzw. definiert. Der Operator X drückt aus, dass im nächsten Moment etwas wahr ist, der Operator U ist ein zweistelliger Junktor und besagt, dass die Aussage an erster Stelle solange wahr ist, bis die Aussage an zweiter Stelle wahr wird. Dann lassen sich F und G durch die anderen beiden Junktoren darstellen. Um diese Operatoren in Modellen zu interpretieren, werden die Modelle auf einer Branching Time Struktur entwickelt. Ähnliches kann für die Vergangenheitsoperatoren eingeführt werden. Dabei steht P für irgendwann einmal in der Vergangenheit. Das Pendant zu U ist dann S . Der zweistellige Junktor S besagt, dass in der Vergangenheit eine Formel solange wahr war, bis eine andere sich bewahrheitete. Die nachfolgende Definition einer Branching Time Struktur ist aus [19, S.30] entnommen und geht auf Prior und Thomason zurück [111, 141].

Definition 1 (Branching Time Struktur - BT Struktur) *Eine Branching Time Struktur besteht aus einer nichtleeren Menge M und einer Ordnungsrelation \leq , die reflexiv, transitiv und antisymmetrisch ist, wobei für je drei Momente $m, m', m'' \in M$*

⁶Es ist bisher unklar, wie die vollständige Ockhamistische Logik mit einer beliebigen Interpretation der Aussagebuchstaben durch beliebige Mengen von Moment-Geschichte-Paaren endlich zu axiomatisieren ist. Dies ist jedoch nicht Teil dieser Arbeit.

1. Eine Logik für Handlungssätze - Stit-Theorie

gilt, falls $m' \leq m$ und $m'' \leq m$, dass $m' = m''$ oder $m' \leq m''$ oder $m'' \leq m'$.

Auf einer Branching Time Struktur lassen sich folgende Begriffe einführen, die für die weiteren Ausführungen von Bedeutung sind. Anhand der Ordnungsrelation lassen sich sogenannte *Geschichtsverläufe* (Geschichten) ablesen. Die Relation $<$ auf M , d.h. die Teilmenge von der Relation \leq ohne die Diagonale, wird interpretiert mit: $m < m'$ gdw. m ist *früher als* m' bzw. m' ist *später als* m . Ein *Geschichtsverlauf* h ist eine maximal lineare Teilmenge von M bezüglich der Relation \leq . Die Menge aller Geschichten auf M wird mit H bezeichnet, wobei H_m für ein $m \in M$ die Menge aller Geschichten benennt, die den Moment m beinhalten. Anhand der Struktur lassen sich die *Situationen*, in denen die Formeln im Modell bewertet werden, angeben. Ein Element (m, h) der Menge $M \times H$ wird als eine *Situation* bezeichnet, falls $m \in h$. Aufgrund der gemeinsamen Vergangenheit, die je zwei Geschichten haben, können sie bezüglich eines gegebenen Momentes als *geteilt* oder *ungeteilt* definiert werden. Zwei Geschichten h, h' gelten im Moment m als ungeteilt, solange es einen späteren Moment $m' > m$ mit $m' \in h \cap h'$ gibt. Dies heißt, dass zwei Geschichten in m geteilt sind, obwohl $m \in h \cap h'$ liegen kann.

Eine gebündelte Branching Time Logik entsteht, wenn für jeden Moment angegeben wird, welches die aktuellen Geschichten sind, die zur Auswertung im Modell betrachtet werden. Dies kann z.B. über eine Relation R gegeben sein. Eine gebündelte Struktur könnte, wie folgt, definiert sein: $\mathbb{B} = (M, \leq, R)$, wobei $R \subseteq H \times M \times H$ und für alle $(h, m, h') \in R$ gilt, dass $m \in h, h'$.⁷ Im Falle der ungebündelten, vollständigen Ockhamistischen Logik gilt, dass für alle $m \in M, h, h' \in H_m$ gilt, dass $(h, m, h') \in R$. Auf der gebündelten als auch ungebündelten Struktur lassen sich Erfüllungsbedingungen für die Zeitoperatoren angeben. Da sie nicht hauptsächlich Gegenstand der Untersuchungen dieser Arbeit sind, gehe ich nur im Vergleich der *BDICTL** mit *bdi-stit* in Abschnitt 6.3 kurz darauf ein.

1.2 Stit-Operatoren: Achievement oder Deliberativ

Auf solchen Branching Time Strukturen entwickelten Belnap und Perloff die sogenannte Stit-Theorie [18]. Sie schlagen verschiedene Handlungsoperatoren vor und interpretieren diese durch eine Äquivalenzrelation, die auf den Geschichtsverläufen zu den jeweiligen Momenten bzw. sogenannten *Instants* definiert ist. Ein Instant ist eine

⁷Ein Überblick über temporale Logiken auf gebündelten und ungebündelten BT Strukturen findet sich in [171].

1.2. Stit-Operatoren: Achievement oder Deliberativ

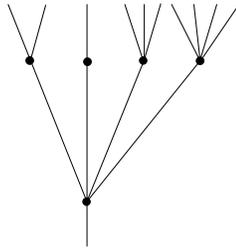


Abbildung 1.1: Branching Time Struktur

maximale Menge an untereinander nicht vergleichbaren Momenten. Um eine sinnvolle Theorie der Handlungen in Form von Wahlmöglichkeiten, bei denen Instants eine Rolle spielen, zu definieren, muss eine Branching Time Struktur mehreren Postulaten genügen. Eine Vorwegannahme ist, dass je zwei Geschichtsverläufe isomorph aufeinander abgebildet werden können. Nur dann ist es möglich, dass die Instants folgenden drei Bedingungen genügen:

- Die Menge aller Instants stellt eine Partitionierung der Menge der Momente dar. Dabei kommt aus jeder Geschichte mindestens ein Moment in jedem Instant vor (Vollständigkeit der Partition).
- Aufgrund der Unvergleichbarkeitsforderung der Elemente innerhalb eines Instants gehört genau ein Moment aus jeder Geschichte zu einem Instant (Eindeutigkeit der Partition).
- Die Partitionierung ist ordnungserhaltend. Für zwei beliebige Instants i_1, i_2 und für alle Geschichten h, h' gilt: Falls $i_1 \cap h \leq i_2 \cap h$, dann $i_1 \cap h' \leq i_2 \cap h'$.⁸

Dass die einzelnen Geschichten isomorph aufeinander abbildbar sind, ist notwendig, die drei Bedingungen sicherzustellen. Die Instants definieren eine Äquivalenzrelation auf den Momenten. Die Äquivalenzklasse, die den Moment m enthält, wird mit $i_{(m)}$ abgekürzt. Falls auf einer Branching Time Struktur (BT) eine Partitionierung in Form der Instants (I) angenommen wird, verdeutlichen Belnap et al. dies, indem sie von einer BT + I Struktur sprechen, siehe [19]. Ich werde diese abkürzende Bezeichnung im Weiteren verwenden.

Um Handlungen von Akteuren zu definieren, wird eine BT Struktur zusätzlich um eine Menge von Akteuren und einer Relation auf den Geschichten erweitert, die die

⁸Da der Schnitt eines Instants und einer Geschichte genau einen Moment enthält, ist diese abkürzende Schreibweise sinnvoll und eindeutig.

1. Eine Logik für Handlungssätze - Stit-Theorie

Wahlmöglichkeiten des jeweiligen Akteurs zu jedem Moment wiedergibt. Die Menge der Akteure wird mit \mathcal{A} bezeichnet, die Wahlmöglichkeiten werden durch die Relation C angegeben. Dabei ist $C \subseteq (\mathcal{A} \times M) \times (H \times H)$.⁹ Die Relation C ordnet jedem Akteur-Moment-Paar eine Äquivalenzrelation auf den Geschichten zu, die durch den Moment verlaufen. Die Äquivalenzklasse, zu der die Geschichte h im Moment m für den Akteur α gehört, wird mit $C_m^\alpha(h)$ angegeben. Die Geschichten die in $C_m^\alpha(h)$ enthalten sind, werden als wahläquivalent zur Geschichte h für den Akteur α im Moment m bezeichnet. Zu einem Moment m ungeteilte Geschichten h, h' sind immer wahläquivalent. Somit legt die Wahl des Akteurs den weiteren Geschichtsverlauf fest bzw. schließt alternative Verläufe aus.

Eine wesentliche Bedingung in der Stit-Theorie an die Wahlmöglichkeiten der Akteure ist die sogenannte *Unabhängigkeit*, vgl. [19, S.217f]. Für eine Anzahl von k unabhängigen Akteuren muss es möglich sein, dass jeder Akteur jede ihm mögliche Handlung auch ausführen kann und nicht durch die Wahl eines anderen Akteurs in seiner Wahl eingeschränkt ist. Für jeden Akteur und für jeden Moment gibt es für jede Kombination der von den Akteuren möglicherweise getroffenen Wahlmöglichkeiten mindestens einen weiteren Geschichtsverlauf, der diese Kombination in dem Moment möglich macht. Da jede Wahlmöglichkeit eines Akteurs aus einer Menge von Geschichten besteht, heißt das, dass der Durchschnitt über alle von k Akteuren beliebig getroffenen Wahlmöglichkeiten zu gegebenem Moment m nicht leer ist. Die *Bedingung der Unabhängigkeit der Akteure* an eine BT Struktur lautet, vgl. Postulat 9 „Independence of Agents“ [19, S.217]:

Für jede Anzahl von Akteuren \mathcal{A} , für jeden Moment m und für jede Funktion σ , die jedem Akteur a auf eine seiner Wahlmöglichkeiten abbildet, $\sigma(\alpha) \in C_m^\alpha$ gilt, dass $\cap \{ \sigma(\alpha) \mid \alpha \in \mathcal{A} \} \neq \emptyset$.

Die jeweils erhaltenen Strukturen werden mit BT + AC bzw. BT + I + AC abgekürzt, wobei A für Akteure und C für die Wahlmöglichkeiten („Choices“) steht. Im Falle des Vorhandenseins von Instants I lässt sich mittels der Äquivalenzrelation C_m^α für alle $\alpha \in \mathcal{A}, m \in M$ eine Äquivalenzrelation auf späteren Instants definieren. Zwei Momente m', m'' sind wahläquivalent für α im Moment m , falls $m'' \in i_{(m')}$ und für alle Geschichten $h' \in H_{m'}, h'' \in H_{m''}$ gilt, dass $h'' \in C_m^\alpha(h')$. Um zu verdeutlichen, ob von der Wahläquivalenz zweier Momente oder Geschichten bzgl. eines Akteur-Moment-

⁹In den Veröffentlichungen von Belnap et al. zur Stit-Theorie wird diese Relation als *Choice*-Funktion bezeichnet, die die Menge der Akteur-Moment-Paare in die Menge der Partitionen der durch den Moment verlaufenden Geschichten abbildet.

1.2. Stit-Operatoren: Achievement oder Deliberativ

Paares gesprochen wird, bezeichnet \underline{C}_m^α die Relation auf den Momenten, $m' \in \underline{C}_m^\alpha(m'')$.

Mit diesen Vorüberlegungen ist es möglich, die im Folgenden definierten Formeln über den eben eingeführten Strukturen zu interpretieren.

Definition 2 *Die Formeln in der Stit-Theorie sind induktiv aufgebaut:*

1. *jeder Aussagebuchstabe $p, q, r, \dots p_1, p_2, \dots$ ist eine Formel.*
2. *falls φ, ψ Formeln sind, so sind auch $\neg\varphi, (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \supset \psi), (\varphi \equiv \psi), \Box\varphi$ und $\Diamond\varphi$ Formeln.*
3. *falls φ eine Formel ist und α eine Akteursvariable, dann sind $\alpha dstit : \varphi$ und $\alpha astit : \varphi$ Formeln.*

Die Formel $\Box\varphi$ wird als historische Notwendigkeit gelesen. Die Formeln $\alpha cstit : \varphi$, $\alpha dstit : \varphi$ und $\alpha astit : \varphi$ werden als ‘der Akteur α sorgt dafür, dass φ ’ gelesen. Während der *dstit*-Operator (deliberative Stit) ein sogenannter instantaner Operator ist, d.h. dass die Auswertung nur Situationen zum aktuellen Zeitpunkt in Betracht zieht, ist der *astit*-Operator (Achievement Stit) zeitlich rückwärtsgerichtet. Der Ausdruck ‘Achievement’ wird hierbei verwendet, um auszudrücken, dass der Akteur zu einem früheren als dem aktuellen Moment eine Entscheidung getroffen hat, mit dem er aktuell das Bestehen von φ in jedem Geschichtsverlauf erreicht hat, der durch den aktuellen Moment verläuft. Die Bewertung des Operators erfolgt somit zu zwei Zeitpunkten. Der Operator ist nicht instantan. Damit tatsächlich ein Erreichen durch den Akteur beschrieben wird, ist zum Zeitpunkt der Entscheidung ein möglicher Geschichtsverlauf nötig, der auf demselben Instant wie der aktuelle Zeitpunkt liegt und der das Bestehen des Sachverhaltes φ nicht zur Folge hatte. Der *astit*-Operator wurde zuerst von Belnap und Perloff vorgeschlagen [18].

Der deliberative Operator beschreibt einen Akteur als abwägend. Der Akteur hat zum aktuellen Zeitpunkt eine Wahl zu treffen. Seine Wahl sichert das Bestehen des Sachverhaltes. Um auszudrücken, dass der Sachverhalt nicht trivialerweise besteht, existiert zum aktuellen Zeitpunkt eine mögliche Situation, in der der Sachverhalt nicht besteht. Diese Form des Stit-Operators wurde unabhängig voneinander von Kutschera in [153] und von Horty später als Reaktion auf den Achievement Stit-Operator vorgeschlagen. Der Unterschied zwischen dem Achievement Stit- und dem deliberativen Stit-Operator zeigt sich in den verschiedenen Interpretationen der Operatoren über der BT + AC bzw. BT + I + AC Struktur.

1. Eine Logik für Handlungsätze - Stit-Theorie

Die Wahrheitsbedingungen sind in [17, 18, 19] angegeben. Die Interpretation der atomaren Formeln erfolgt über eine Bewertungsfunktion v , die die BT $[+ I] + AC$ Struktur zu einem Modell vervollständigt. Sie bildet die atomaren Formeln in die Potenzmenge der Situationen ab und die Akteursvariablen in die Menge der Akteure \mathcal{A} . Das Bild $v(p)$ beinhaltet dabei die Moment-Geschichte-Paare, in denen p wahr ist, für alle atomaren Formeln p .

Definition 3 (Stit Semantik) *Gegeben sei eine BT + I + AC Struktur mit den eben eingeführten Mengen und v eine Bewertungsfunktion auf dieser Struktur. Sei $s = (m, h)$ eine Situation in der BT Struktur, seien α, β Akteursvariablen und seien φ, ψ Formeln gemäß Definition 2, dann gelte:*

$$\begin{aligned}
 \mathcal{M}, s \models \varphi & \quad \text{gdw.} \quad s \in v(\varphi), \text{ falls } \varphi \text{ eine Formel gemäß 2.1 ist.} \\
 \mathcal{M}, s \models \alpha = \beta & \quad \text{gdw.} \quad v(\alpha) = v(\beta). \\
 \mathcal{M}, s \models \neg\varphi & \quad \text{gdw.} \quad \mathcal{M}, s \not\models \varphi. \\
 \mathcal{M}, s \models (\varphi \wedge \psi) & \quad \text{gdw.} \quad \mathcal{M}, s \models \varphi \text{ und } \mathcal{M}, s \models \psi. \\
 \mathcal{M}, s \models \Box\varphi & \quad \text{gdw.} \quad \mathcal{M}, (m, h') \models \varphi \text{ für alle } h' \in H_{(m)}. \\
 \mathcal{M}, s \models \alpha \text{ dstit: } \varphi & \quad \text{gdw.} \quad (+) \text{ für alle } h' \in C_m^\alpha(h) \text{ gilt } \mathcal{M}, (m, h') \models \varphi, \\
 & \quad \quad \quad (-) \text{ es gibt } h'' \in H_m \text{ mit } \mathcal{M}, (m, h'') \not\models \varphi. \\
 \mathcal{M}, s \models \alpha \text{ astit: } \varphi & \quad \text{gdw.} \quad \text{es gibt } w \in M \text{ mit } w < m \text{ und} \\
 & \quad \quad \quad (+) \text{ für alle } m' \in \underline{C}_w^\alpha(m), h' \in H_{m'} \text{ gilt:} \\
 & \quad \quad \quad \mathcal{M}, (m', h') \models \varphi, \\
 & \quad \quad \quad (-) \text{ es gibt } m'' \in i_{(m)} \text{ und } h'' \in H_{m''} \text{ mit} \\
 & \quad \quad \quad \mathcal{M}, (m'', h'') \not\models \varphi.
 \end{aligned}$$

Offensichtlich ist im Falle des *dstit*-Operators die Hinzunahme von Instants überflüssig, so dass es ebenfalls unnötig ist anzunehmen, dass die Geschichten paarweise isomorph wie im oben benannten Sinn sein müssen. Der *dstit*-Operator lässt sich somit über einer BT + AC Struktur interpretieren. Beim *astit*-Operator ist jedoch eine BT + I + AC Struktur notwendig. Es gibt weitere Unterschiede. Bevor darauf eingegangen wird, seien einige Gemeinsamkeiten genannt. Beiden Operatoren ist gemein, dass sie eine Positiv- und eine Negativ-Bedingung als Erfüllungskriterium aufweisen, vgl. [17]. Im Falle einer Handlung erscheint es notwendig, den Akteur als ursächlich zu beschreiben. Um zu beurteilen, dass der Akteur für das Eintreten von φ gesorgt hat, muss es einen Geschichtsverlauf geben, in dem φ nicht eingetreten wäre.¹⁰ Dar-

¹⁰Da ich von φ als Aussagevariable spreche, ist das Eintreten von φ synonym zu verstehen mit dem Bestehen des zur Aussagevariablen bzw. Formel φ korrespondierenden Sachverhaltes, vgl. Kapitel 2. Eine

1.2. Stit-Operatoren: Achievement oder Deliberativ

auf zielt nicht nur die zweite, sogenannte Negativ-Bedingung (–) ab, sondern auch die Positiv-Bedingung (+) sorgt bei beiden Operatoren dafür, den Akteur als Ursache des Eintretens von φ zu beschreiben. Falls der Akteur zum Zeitpunkt der Wahl sich für einen Verlauf entscheidet, dann kann er nur Ursache für die in dem Moment der Auswertung eingetretenen Sachverhalte sein, wenn in jedem Geschichtsverlauf, den er nicht von dem Ausgewählten unterscheiden kann, die Sachverhalte ebenfalls eintreten.

An dieser Formulierung lässt sich bereits der grundlegende Unterschied zwischen dem *astit* und dem *dstit*-Operator festmachen. Der Zeitpunkt der Wahl und der Zeitpunkt, an dem die Handlung als vollzogen erachtet wird, sind bei *astit* notwendigerweise verschieden. Beim *dstit*-Operator fallen beide zusammen. Daher sind die Operatoren nicht wechselseitig definierbar, selbst wenn temporale Operatoren hinzugenommen werden. Seien P und F die temporalen Operatoren ‘irgendwann in der Vergangenheit’ bzw. ‘irgendwann in der Zukunft’. Man würde vermuten, dass die folgenden Beziehungen gelten:

$$\alpha \text{astit}:\varphi \supset P \alpha \text{dstit}:F \varphi,$$

$$\alpha \text{dstit}:F \varphi \supset F \alpha \text{astit}:\varphi.$$

Sie erweisen sich jedoch als nicht allgemeingültig, siehe [17, S.549]. Der Unterschied zwischen den Stit-Operatoren wird in der Interaktion mit den Modaloperatoren deutlich. Ein Ausdruck der Form $\alpha \text{astit}:\varphi$ bedeutet, dass „the present momentary fact that φ is guaranteed by a prior choice of the agent α “ [19, S.33]. Entschied sich der Akteur zu einem früheren Zeitpunkt, eine Handlung zu vollziehen, und wird deren Vollendung erst in dem aktuellen Moment ausgewertet, so ist es für die Zuschreibung der Handlung zu dem Akteur notwendig, dass er in allen Geschichtsverläufen, die zu dem aktuellen Moment möglich sind, die Handlung vollzogen hat, sonst hätte er sie durch seine frühere Entscheidung nicht verursacht. Ein notwendiges Axiom für den *astit*-Operator ist somit:

$$\alpha \text{astit}:\varphi \supset \Box \varphi.$$

Für den *dstit*-Operator ist dieses Axiom offensichtlich unerfüllbar. Aus der Negativ-Bedingung folgt, dass es mindestens einen anderen Geschichtsverlauf parallel zu der aktuellen Situation gibt und mindestens in einer der aktuell möglichen Situationen

Korrespondenztheorie zwischen in einer Situation bestehenden Sachverhalten und wahren Aussagen (Propositionen) wird angenommen. Aussagen werden mit Propositionen gleichgesetzt.

1. Eine Logik für Handlungssätze - Stit-Theorie

der Fall ist, dass φ nicht der Fall ist, obwohl er aktuell wahr ist. Wenn eine notwendigerweise wahre Tatsache in der aktuellen Situation vorliegt, kann keine durch den *dstit*-Operator beschriebene Handlung des Akteurs den aktuellen Sachverhalt verursacht haben. Das zugehörige Axiom lautet:

$$\alpha \text{ dstit}:\varphi \supset \neg\Box\varphi.$$

Wenn angenommen wird, dass der *dstit*-Operator über einer BT + I + AC Struktur interpretiert wird, ergibt sich somit aus den eben genannten Beziehungen folgendes Interaktionsaxiom zwischen den Operatoren:

$$(\alpha \text{ dstit}:\varphi \wedge \alpha \text{ astit}:\varphi) \supset \perp.$$

Dieser Unterschied in der Interaktion mit den historischen Modaloperatoren ist die Grundlage dafür eine Logik, die Interaktionen zwischen Handlungen und mentalen Zuständen beschreibt, auf der Semantik des *dstit*-Operators zu konstruieren. Dafür möchte ich im Folgenden kurz argumentieren und darauf in Abschnitt 4.3 genauer eingehen.

Eine Diskussion von Interaktionen zwischen dem Handlungs- und dem Überzeugungsoperator sollte die Debatte um den doxastischen Voluntarismus berücksichtigen. Es wird diskutiert, ob eine bestimmte Überzeugung zu haben durch eine Handlung hervorgebracht werden kann oder nicht, siehe hierfür Abschnitt 4.1. Eine inhaltlich gerichtete Kontrolle bei der Überzeugungsbildung, d.h. Kontrolle darüber sich direkt zu entscheiden, eine bestimmte Proposition zu glauben, wird i.A. zurückgewiesen. Hierfür vergleiche z.B. [4, 102].

Es scheint unplausibel anzunehmen, dass der Akteur eine Situation auswählen könnte, in der der Sachverhalt, dass eine bestimmte Proposition φ von ihm geglaubt wird, auf einem Instant eine historische Notwendigkeit ist, die in einem anderen Geschichtsverlauf auf demselben Instant falsch ist. Selbst wenn er Kontrolle darüber hätte, durch seine Wahl direkt sicherzustellen, dass er in allen kommenden möglichen Geschichtsverläufen die Überzeugung hätte, dass φ , scheint es unplausibel, dass er eine Überzeugungsbildung in einem aktual unmöglichen Geschichtsverlauf vermeiden könnte.¹¹

¹¹Hiermit ist nicht gemeint, dass er sich entscheidet, etwas zu tun, so dass die Überzeugung, dass φ , aus der Handlung resultiert. Beispielsweise kann jemand sicherstellen zu glauben, dass das Licht an ist, indem er den Lichtschalter getätigt. Jedoch ist mir unklar, welche Wahl er vollziehen sollte, um bei eingeschaltetem Licht nicht zu glauben, dass das Licht an ist. Es ist keine Antwort zu sagen, dass er das Licht ausschalten könnte. So würde er indirekt dafür sorgen, die Überzeugung zu haben, aber nicht direkt, vgl. Feldman [58, S.671].

1.2. *Stit-Operatoren: Achievement oder Deliberativ*

Im Falle der Bildung mentaler Zustände scheint es nur natürlich anzunehmen, dass die Kontrolle zur Bildung des mentalen Zustandes die Kontrolle, den Zustand nicht zu bilden, impliziert, auch wenn bei Handlungen, die sich nicht auf die Bildung / Wandlung / Beeinflussung mentaler Zustände beziehen, dies nicht notwendigerweise der Fall ist.¹²

Der mentale Zustand liegt im Subjekt vor oder nicht vor. Wenn er nicht vorliegt, gibt es zwei Möglichkeiten: Entweder das Subjekt hat Kontrolle über das Vorliegen des Zustandes, es hat diese ausgeübt und der Zustand tritt ein oder er tritt nicht ein. Dann ist das Nichteintreten ebenfalls eine Wahl des Akteurs. Dies hätte zur Folge, dass für den Akteur in der Situation, die mit der aktuellen auf demselben Instant liegt, das Nichtbesitzen der Überzeugung eine notwendige Wahrheit ist. In ihr hat der Akteur hervorgebracht, dass er die Überzeugung nicht hat. Oder er hat keine Kontrolle ausgeübt oder keine gehabt. Dann kann weder das Eintreten noch das Nichteintreten als seine Handlung gezählt werden. Der Akteur kann zu einem gegebenen φ nicht ohne Weiteres die Überzeugung, dass φ , bilden. Der doxastische Voluntarismus, verstanden als inhaltsgerichtete, direkte Kontrolle über unsere Überzeugungen, ist unplausibel. Es gibt gute Gründe den doxastischen Voluntarismus als nicht akzeptabel zu betrachten, siehe [4, 102, 162]. Daher sollten Aussagen der Form $\alpha \text{ astit}:\varphi$ bzw. $\alpha \text{ dstit}:\varphi$, wobei φ ausdrückt, dass der Akteur α eine bestimmte Proposition glaubt, nicht erfüllbar sein.

Plausibel scheint hingegen zu sein, dass eine Überzeugung eines Akteurs eine historische Notwendigkeit ist, da sie auf den bis dato aufgetretenen und dem Akteur vorliegenden Evidenzen gebildet wird. Wenn in einem Moment eine Überzeugung notwendigerweise eintritt und alle durch den Moment verlaufenden Geschichten eine Wahlzelle des Akteurs zu einem früheren Moment bilden, dann gibt es zwei Möglichkeiten: Die erste ist, für diesen Fall zu postulieren, dass zu allen Momenten desselben Instants die Überzeugung notwendigerweise wahr ist. Dies scheint unplausibel, da in unterscheidbaren, nicht parallelen Geschichtsverläufen unterschiedliche Propositionen in der Vergangenheit wahr sein könnten und so auch unterschiedliche Evidenzen vorliegen können, anhand derer die Akteure ihre Überzeugungen bilden.

Die zweite Möglichkeit wäre zu akzeptieren, dass die Überzeugungsbildung, dass φ , durch den *astit*-Operator beschrieben wird, was einer Handlungsbeschreibung gleich käme und somit im Widerspruch zum direkten doxastischen Involuntarismus stände. Das heißt aufgrund der Interaktion des *astit*-Operators mit der historischen Notwendigkeit und der Intuition, dass Überzeugungen historische Notwendigkeiten sind, bie-

¹²Auf diese Punkte komme ich in Abschnitt 3.3 zurück.

1. Eine Logik für Handlungssätze - Stit-Theorie

tet sich für eine Modellierung von Interaktionen zwischen Handlungs- und Überzeugungsoperatoren der *dstit*-Operator im Vergleich zum *astit*-Operator an. Eine umständlichere Handhabung des letzteren im Vergleich mit dem *dstit*-Operator bei einer Axiomatisierung bzw. dass im Falle des *astit*-Operators weitere Anforderungen der BT Strukturen bestehen, wie die Isomorphie der Geschichtsverläufe und die Definition der Instants, lassen die Wahl des *dstit*-Operators ebenfalls sinnvoll erscheinen.¹³

Es gibt einen weiteren offensichtlichen Unterschied zwischen dem Besitz einer propositionalen Einstellung, wie z.B. einer Überzeugung, und einer Handlung bei der Beschreibung durch Operatoren. Ein Operator, der das Vorliegen eines mentalen Zustandes¹⁴ wie den einer Überzeugung oder eines Wunsches beschreibt, interagiert mit den aktuellen Fakten und historisch möglichen Sachverhalten nicht in der Art wie es bei Handlungen der Fall ist. Ein Handlungsoperator¹⁵ muss in einem gewissen Sinn faktiv sein. Ob eine Handlung ausgeführt wurde, wird u.a. an dem Bestehen des Sachverhaltes bewertet, der das Resultat der Handlung ist. Die Symbolkette ' $\alpha dstit:\varphi$ ' drückt zwar lediglich aus, dass der Akteur, der durch α symbolisiert wird, den Sachverhalt, der durch φ benannt wird, hervorgebracht hat, und somit nicht *wie* er ihn hervorgebracht hat. Aber er drückt aus, dass φ . Bei einer propositionalen Einstellung wie Überzeugungen ist dies nicht notwendigerweise der Fall. Der propositionale Inhalt einer Überzeugung kann in einer Situation aktuell falsch sein. Ich spreche dann von wahren bzw. korrekten Überzeugungen, wenn der Überzeugungsinhalt im Moment des Vorliegens der Überzeugung wahr ist. Inwieweit die Sachverhalte, die in aktuellen bzw. aktuell möglichen Situationen bestehen, zur Auswertung von Operatoren, die Überzeugungen und Wünsche beschreiben, herangezogen werden sollten, werde ich in den Abschnitten 4.1 und 4.2 diskutieren.

¹³Eine Axiomatisierung des *dstit*-Operators wurde von Xu angegeben [19, 167, 168]. Ich werde in Abschnitt 6.1 darauf zurückkommen. Eine vollständige Axiomatisierung für den *astit*-Operator unter Annahme bestimmter Einschränkungen der BT + I + AC Struktur findet sich in [19].

¹⁴Im Falle einer Überzeugung wird es zumeist als äquivalent betrachtet, vom Besitz der doxastischen Einstellung und dem Vorliegen des mentalen Zustandes der Überzeugung zu sprechen. Ich würde dies für Wünsche ebenfalls postulieren. Ein Akteur hat die propositionale Einstellung eines Wunsches φ genau dann, wenn ein mentaler Zustand des Wünschens, dass φ , im Akteur vorliegt.

¹⁵Auf die Problematik, dass Handlungen nicht auf Propositionen operieren, im Sinne wie Sachverhalte Propositionen korrespondieren, sondern Handlungen nur durch diese charakterisiert werden, werde ich in Kapitel 3 beschreiben. Dafür ist es notwendig, die Begriffe 'Sachverhalt', 'Proposition' und 'Handlung' genau zu definieren, siehe Kapitel 2.

1.3 Das Problem der konfligierenden Historie

Ein anderes Problem der eben definierten BT Strukturen, ist, dass Situationen in ungeteilten Geschichten verschiedene, sich möglicherweise widersprechende Formeln erfüllen können. Gegeben sei folgendes Modell $M = \{a, b, c, d, e\}$ mit $\leq = \text{trcl}^{16}\{(a, b), (a, c), (b, e), (b, d)\}$. Die Interpretationsfunktion für einen beliebigen Aussagebuchstaben p sei durch $v(p) = \{(a, (\{a, b, d\}, \leq)), (b, (\{a, b, d\}, \leq))\}$ definiert. Die Aussage p ist in der Situation $(a, (\{a, b, d\}, \leq))$ erfüllt und in $(a, (\{a, b, e\}, \leq))$ nicht erfüllt, obwohl die Geschichten $(\{a, b, d\}, \leq)$ und $(\{a, b, e\}, \leq)$ zum Moment a nicht geteilt sind.

Mehrere Autoren sehen darin ein Problem, da dies auch für einen beliebigen Akteur und ununterscheidbare Geschichten zutrifft. Wenn zwei Geschichten nicht geteilt sind und somit eine gemeinsame Vergangenheit besitzen, dann scheint es nur natürlich anzunehmen, dass in der gemeinsamen Vergangenheit auch dieselben Aussagen zutreffen und somit für Aussagen über die Vergangenheit gilt, dass sie notwendigerweise wahr bzw. notwendigerweise falsch sind.

Ein Lösungsvorschlag ist [62], die Aussagevariablen in jedem Moment zu bewerten. In allen Situationen, die in einem Moment liegen, ist die Bewertung der propositionalen Aussagevariablen dieselbe. Nachgerade sind alle atomaren Aussagen historisch notwendigerweise wahr oder falsch in jeder Situation. Da jeder Operator der deliberativen Stit-Theorie in einem Moment bewertet wird, wären ohne Berücksichtigung temporaler Operatoren alle Aussagen notwendigerweise wahr oder notwendigerweise falsch. Alle *dstit*-Formeln somit unerfüllbar. Somit löst die Auswertung der atomaren Formeln in den Momenten das oben beschriebene Problem zwar hinreichend, allerdings ist diese Lösung nicht zwingend und scheint ad hoc zu sein.

Zanardo beschreibt in [171] einen Vorschlag, den er Ming Xu zuschreibt. Eine Motivation für Xus Lösung ist die folgende. Das eigentliche Problem der konfligierenden Historie ist nicht, dass es unplausibel ist, dass es verschiedene Geschichtsverläufe durch ein und denselben Moment gibt, in denen Formeln verschieden bewertet werden, sondern dass die unterschiedliche Bewertung der Formeln für die Geschichtsverläufe in ein und demselben Moment unplausibel ist, wenn sie dieselben zukünftigen Situationen beinhalten. Somit scheint es notwendig zu sein, dass in verschiedenen Geschichten, solange sie nicht geteilt sind, die Formeln in den Momenten mit demselben Wahrheitswerten belegt werden. Dies gilt jedoch nicht notwendig auch für alle zukünftigen Momente. Die Eindeutigkeit der gemeinsamen Vergangenheit sollte zudem nicht nur für atomare, sondern für alle Formeln gelten, die in einer gemein-

¹⁶Die Abkürzung steht für den transitiven und reflexiven Abschluss der nachfolgenden Menge.

1. Eine Logik für Handlungsätze - Stit-Theorie

samen Vergangenheit wahr sind. Im obigen Lösungsansatz, in dem nur die Momente eine Rolle bei der Interpretation der atomaren Formeln spielen, würde dies bedeuten, dass Zukunftsoperatoren in der ungeteilten Vergangenheit dann mit notwendigerweise wahr bzw. notwendigerweise falsch bewertet werden. Dies entspräche einer Peirceschen und keiner Ockhamistischen Logik mehr.

Eine weitere Überlegung macht Xus Vorschlag plausibel, wenn die Frage gestellt wird, wann sich eine Geschichte teilt. Dies ist der Fall, wenn die darunter liegende Zeitstruktur durch die Wahl der Relation \leq dies vorgibt. Eine Intuition könnte sein, dass sich die Geschichten teilen, sobald sich widersprechende Sachverhalte in den Geschichten zu einem bestimmten Zeitpunkt bestehen. Sobald zu einem Moment in einer Geschichte die Aussage φ und in der anderen Geschichte, die durch denselben Moment verläuft, die Aussage $\neg\varphi$ gilt, sollten die Geschichten geteilt sein. Im spätesten Moment der Schnittmenge zweier Geschichten werden diese bereits als geteilt erachtet, da sie verschiedene Formeln erfüllen.

Folgt man dieser Argumentation, lässt sich das Problem der konfligierenden Historie mit einer geeigneten Wahl der Interpretationsfunktion der Aussagebuchstaben vermeiden. Diese Strategie scheint plausibel und weniger ad hoc. Um der Dissonanz der gemeinsamen, verschiedenen Vergangenheit entgegenzuwirken, ist für die Interpretationsfunktion die folgende Eigenschaft zu berücksichtigen: Wenn für alle $h, h' \in H, m \in M$ ein $m' \in h \cap h'$ mit $m' > m$ existiert, dann soll für alle atomaren Formeln gelten, dass $(m, h) \in v(p)$ gdw. $(m, h') \in v(p)$. Solange die Geschichten nicht geteilt sind, werden die atomaren Formeln in den Situationen der ungeteilten Geschichten gleich bewertet. Es genügt, dies für die atomaren Formeln zu fordern, solange die Junktoren, die über der BT Struktur interpretiert werden, im aktuellen Moment ausgewertet werden. Gemeint ist damit, dass sich die Erfüllbarkeitsdefinition nur auf den Moment bezieht, zu dem die Formel bewertet wird. Daraus folgt, dass nicht nur die atomaren Formeln, sondern alle Formeln in den ungeteilten Geschichten gleich bewertet werden. Demnach wäre dies zutreffend, solange die Sprache Junktoren wie Konjunktion, Disjunktion, historische Notwendigkeit und den deliberativen Stit-Operator enthält.¹⁷

Um zu garantieren, dass die Bewertungsfunktion eben genannte Eigenschaft besitzt, wäre folgender Algorithmus zielführend. Für jeden Moment m wird eine Äquivalenzrelation auf den Geschichten H_m eingeführt. Für zwei Geschichten $h', h \in H_m$ gilt,

¹⁷Für den Achievement Stit-Operator wäre die Gleichbewertung von Formeln in ungeteilten Geschichten im selben Moment ebenfalls kein Problem, da die Wahrheitsdefinition zwar auf einen früheren Moment Bezug nimmt, jedoch nicht auf die Bewertung einer Formel in dem früheren Moment zurückgreift.

1.3. Das Problem der konfligierenden Historie

sie sind geschichtsäquivalent, wenn sie nicht geteilt sind, d.h., dass es einen späteren Moment als m gibt, der in beiden Geschichten liegt. Dass dies eine Äquivalenzrelation ist, ist offensichtlich, siehe auch [19].

Lemma 4 Sei (M, \leq) eine BT Struktur, wobei $m \in M$ und H_m die Menge der Geschichten, die durch m verlaufen, bezeichnet. Dann wird durch folgende Relation für jedes $m \in M$ eine Äquivalenzrelation auf H_m für alle h, h' definiert: $h \equiv_m h'$ gdw. es existiert $m' \in M$ mit $m' > m$ und $m' \in h \cap h'$.

BEWEIS: Reflexivität und Symmetrie sind trivial. Angenommen es gilt für die Geschichten $h \equiv_m h'$ und $h'' \equiv_m h'$. Es existieren $m', m'' \in M$ mit $m', m'' > m$ und $m' \in h \cap h'$ und $m'' \in h' \cap h''$. Da $m', m'' \in h'$ sind, sind m' und m'' vergleichbar. Das Minimum von beiden liefert die Behauptung $h \equiv_m h''$. \square

Die Äquivalenzrelation definiert eine Partitionierung auf den Geschichten zu jedem Zeitpunkt, so dass über der Menge der Situationen ebenfalls eine Äquivalenzrelation definiert werden kann, $(m', h') \equiv (m, h)$ gdw. $m = m'$ und $h \equiv_m h'$. In [19] heißen die Äquivalenzklassen unmittelbare Möglichkeiten.¹⁸ Analog bezeichne ich Paare von Momenten und einer zugehörigen unmittelbaren Möglichkeit als eine unmittelbar mögliche Situation.

Wird die Bewertungsfunktion über eine Hilfsfunktion v' angegeben, die die atomaren Formeln nicht in die Potenzmenge der Situationen abbildet, sondern in die Potenzmenge der unmittelbar möglichen Situationen $S' = \{(m, [h]_{\equiv_m}) \mid (m, h) \in S\}$, so lässt sich die Bewertungsfunktion $v : \mathcal{L} \rightarrow \mathcal{P}(S)$ mit $v(p) = \{(m, h) \mid (m, [h]_{\equiv_m}) \in v'(p)\}$ angeben. Die so definierte Bewertungsfunktion erfüllt offensichtlich die oben genannte Eigenschaft. Es wäre möglich, statt der Interpretation über den Situationen die Auswertung der Formeln in einem Stit-Modell über der eben definierten Menge der unmittelbar möglichen Situationen vorzunehmen.

In [172] beschäftigt Zanardo sich mit ununterscheidbaren Funktionen I , die zu einer BT Struktur hinzugenommen werden und jedem Moment eine Äquivalenzrelation auf der Menge der durch den Moment verlaufenden Geschichten H_m zuordnen. Die Auswertung der Formeln erfolgt in Moment-Äquivalenzklasse-Paaren über der (M, \leq, I) Struktur. Diese Struktur bezeichnet Zanardo als I -Tree. Beispiele für die Wahl der Funktion I sind z.B. die Choice-Funktion für einen festen Akteur, aber auch die trivialen Möglichkeiten wie $I_p(m) = H_m \times H_m$ und $I_o(m) = \{(h, h) \mid h \in H_m\}$. Die Wahl I_o führt zu einer zur darunter liegenden BT Struktur äquivalenten Beschreibung. Die

¹⁸In [16] hatte sie Belnap als elementare Möglichkeiten bezeichnet.

1. Eine Logik für Handlungssätze - Stit-Theorie

Wahl I_p führt zu einer Perceischen Logik [172, S.314]. Zanardo zeigt, falls die ununterscheidbare Funktion I mit $I(m) = \{(h, h') \mid h \equiv_m h'\}$ gewählt wird, dass dies die minimale Forderung ist, so dass alles Vergangene notwendigerweise der Fall war: $P\varphi \supset \Box P\varphi$ ist gültig, solange φ keinen F -Operator enthält. Da Ming Xus Vorschlag äquivalent zu diesem ist, ist er ebenfalls eine Minimalforderung an eine BT Struktur, damit die Vergangenheit nicht kontingent ist. Es ist zu beachten, dass im Gegensatz zu dem Lösungsvorschlag, die atomaren Formeln in den Momenten zu bewerten, der unvereinbar mit dem *dstit*-Operator ist, eine Formel der Form $\Box P\alpha \text{ dstit} : \varphi$ durchaus erfüllbar ist.

Belnap et al. unterstreichen in Postulat 8 „No choice between undivided histories“ [19, S.462], dass Akteure zwischen ungeteilten Geschichten nicht wählen können. Aus der Ungeteiltheit zweier Geschichten folgt auch ihre Ununterscheidbarkeit. Somit könnte in einem Rahmen, in dem die Bewertung der Formeln nicht über die Situationen, sondern über die unmittelbar möglichen Situationen erfolgt, der damit einen Spezialfall eines I -Trees darstellt, die Choice-Funktion für einen festen Akteur über diesem I -Tree als Abbildung der Momente in die Menge der Äquivalenzrelationen über den Äquivalenzklassen zur Ungeteiltheit wählen, d.h. $C_m^\alpha([h]_{\equiv_m}) = \{[h']_{\equiv_m} : \text{für einige } h' \in H_m\}$. Dieser Rahmen $(M, \leq, I) + AC$ hätte Postulat 8 sofort zur Folge, wohingegen bei einer $BT + AC$ Struktur, die Wahl von C das Postulat explizit berücksichtigen sollte.

Somit wäre eine Bewertung über den unmittelbar möglichen Situationen einer Bewertung über einer $BT + AC$ Struktur äquivalent, die zudem der Bedingung der Ungeteiltheit der Vergangenheit und Postulat 8 genügt. Daher verwende ich im Weiteren wie in der Literatur üblich die Menge der beliebigen Situationen über einer BT Struktur als Bewertungskontexte in einem Modell, zumal Vergangenheitsoperatoren in dieser Arbeit nur am Rande betrachtet werden. Für die Beschreibung von Handlungen und mentalen Zuständen durch instantane Operatoren ist das Problem der konfligierenden Geschichten nicht von Bedeutung. Da sich dies durch entsprechende Wahl der Bewertungsfunktion bzw. durch Konstruktion eines I -Trees beheben ließe, wobei die ununterscheidbare Funktion I die Ungeteiltheit beschreibt, stellt dies auch im Falle einer temporalen Erweiterung einer *dstit*-Logik kein Problem dar. Aus Gründen der Übersichtlichkeit halber verzichte ich jedoch auf diese Konstruktion. Für die Diskussion von Interaktionen zwischen Operatoren, die mentale Zustände und Handlungen beschreiben, sind die genaue Definition dessen, was ein Zustand, was eine Handlung ist, von größerer Bedeutung. Daher werde ich im folgenden Kapitel Begriffe, wie Zustand, Ereignis und Sachverhalt genauer untersuchen.

2 Zustände, Sachverhalte und Ereignisse

Überzeugungen, Wünsche und Intentionen werden im Allgemeinen¹ als mentale Zustände angesehen. Sie werden als propositionale Einstellungen beschrieben. Bevor ich Beziehungen zwischen diesen Einstellungen näher untersuchen kann, halte ich fest, was unter Zuständen und Ereignissen verstanden wird. Zwischen mentalen und nicht-mentalenen Zuständen scheint es Unterschiede zu geben, z.B. hinsichtlich der Kontrolle, die Akteure über ihre die jeweiligen Zuständen besitzen. Jedoch lassen sich ebenso viele Gemeinsamkeiten finden, die einem Zustand generell inhärent sind. Er besteht bzw. liegt vor zu einem bestimmten Zeitpunkt. Er ist weder dynamisch noch statisch. Diese Attribute treffen auf Ereignisse zu. Hierbei scheint es ebenfalls Unterschiede zwischen mentalen Ereignissen, wie das Aufgeben, Formen und Fortbestehen von mentalen Zuständen, und nicht-mentalenen Ereignissen zu geben. Ereignissen ist es unmöglich, dass sie punktuell wie Zustände sind, sondern sich über Zeiträume hinweg erstrecken. Um Handlungen und mentale Zustände in Beziehungen zu setzen, stellt sich die Frage, wie diese zu beschreiben sind. Da ich die These vertreten werde, dass Handlungen ein Ereignis zugrunde liegt, werde ich die Begriffe ‘Zustand’ und ‘Ereignis’ im mentalen als auch im nicht-mentalenen näher untersuchen.

Der grundlegende Begriff des Zustandes sollte nicht zwischen einem mentalen Zustand versus einem nicht-mentalenen Zustand variieren. Auch ein mentaler Zustand sollte die Kriterien, die der Begriff ‘Zustand’ ohne Einschränkung auf das Mentale beinhaltet, erfüllen. Das Gleiche gilt für Ereignisse. Für die Konstruktion einer Logik die

¹*Im Allgemeinen* verwende ich, weil einige Autoren bestreiten, dass mentale Zustände generell nicht angemessen als propositionale Einstellung charakterisiert werden können, z.B. [20], bzw. dass Wünsche und Intentionen zwar Einstellungen aber keine propositionalen Einstellungen sind, vgl. Thagard in [140]. Es gibt weiterhin auch Einwände gegen die These, dass es sich bei Intentionen überhaupt um mentale Zustände handelt, hierfür siehe [7, 142]. Thompson [142] argumentiert, dass es sich bei einer Intention, etwas zu tun, nicht um einen mentalen Zustand handeln kann, da diese nicht statisch ist. Ich werde dafür argumentieren, dass ein Zustand generell nicht statisch ist, so dass dieses Argument nicht greift.

2. Zustände, Sachverhalte und Ereignisse

Handlungen und Überzeugungen beschreibt, stellt sich die Frage, wie Handlungen und mentale Zustände einheitlich beschrieben werden können (im Sinne von in einem gemeinsamen System beschreibbar). Beide werde ich durch modallogische Operatoren erfassen. Dafür sollte, die Frage beantwortet werden, wann etwas als eine Handlung beschrieben wird und was einen Handlungssatz kennzeichnet. Während man bei Überzeugungen von mentalen Zuständen spricht, ist es eine weitverbreitete Ansicht, dass Handlungen eher Ereignisse zugrundeliegen. Daher möchte ich in diesem und dem folgenden Kapitel klären, was einen Zustand von einem Ereignis unterscheidet, was ein Ereignis bzw. einen Zustand charakterisiert und wie dies auf Beschreibungsebene unterschieden wird bzw. ob es immer dazu führt, dass Handlungen und mentale Zustände verschieden beschrieben werden müssen.

Beginnen werde ich zur Motivation mit der Frage, was eine Handlung beschreibt. Im Anschluss daran werde ich in diesem Kapitel die Begriffe ‘Zustand’, ‘Ereignis’, ‘Sachverhalt’ und ‘Situation’ in Geschichtsverläufen diskutieren und definieren, um mich im anschließenden Kapitel 3 auf Beschreibungsebene der Beantwortung zweier Fragen zu widmen. Zum einen werde ich die motivationale Frage beantworten, was eine Handlung von einer Zustandsbeschreibung unterscheidet, und zum anderen werde ich mich der erstgenannten Frage stellen, wie können Handlungen und mentale Zustände einheitlich beschrieben werden.

Belnap hat aufgezeigt, z.B. in [15], dass man sich von zwei Seiten den Fragen nähern kann, was ein Handlungssatz ist und wie er sich zusammensetzt. Belnap stellt auf die eine Seite ‘Actions-as-Events’ mit ihrem Vertreter Davidson und auf die andere Seite die Weiterentwicklung der Modallogik von Handlungssätzen. Während Davidson betrachtet, welche Komponente(n) einen Satz als Handlungssatz auszeichnen, stellt die andere Seite die Frage, inwieweit sich Handlungssätze in den Kontexten, in denen sie auftreten, als solche zu erkennen geben.

Davidson [48, Essay 6] geht davon aus, dass ein Handlungssatz ein Ereignis beschreibt, dem der Handlungssatz mehrere Eigenschaften bzw. Beziehungen zu anderen Einzeldingen zuschreibt, auf die die Teile des Satzes referieren. So analysiert Davidson die logische Struktur eines Satzes, indem er jedem grammatischen Teil (Subjekt, Objekt, Prädikat/Verb, adverbialen Bestimmungen etc.) ein ein- bzw. mehrstelliges Relationssymbol und entsprechende Konstanten zuordnet, diese in die Relationssymbole einsetzt, und die Relationssymbole durch Konjunktionen verbindet.

Jedes Verb, welches eine Handlung ausdrückt, hat nach Davidson eine sogenannte Ereignis-Stelle (Event Place) [48, Essay 6]. Das Verb, das die Art des Ereignisses vorgibt, wird durch ein Relationssymbol dargestellt, dessen eine Stelle eine Ereignis-

variable einnimmt. Da Davidson dem Verb eine ausgezeichnete Stellung zuschreibt, lässt sich seine Darstellung eines Handlungssatzes am ehesten anhand der abhängigen und unabhängigen Ergänzungen des Verbes bzw. der weiteren Angaben erklären. Ist die Valenz eines Verbes innerhalb eines Satzes n , so wird das Relationssymbol, welches Davidson für das Verb einführt, $n + 1$ Stellen aufweisen, wobei die $n + 1$ -ste Stelle die Ereignisvariable besetzt. Die obligatorischen Ergänzungen des Verbes werden durch Individuenkonstanten symbolisiert, die die restlichen n Stellen auffüllen, die z.B. das Handlungssubjekt bzw. das Objekt symbolisieren. Fakultative Ergänzungen können durch Individuenkonstanten als Besetzung von Stellen des Relationssymboles des Verbes oder ebenso wie weitere Ergänzungen, d.h. weitere freie Angaben, behandelt werden. Diese Angaben, die fast ausschließlich adverbiale Bestimmungen sind, werden ebenfalls als entsprechende Relationssymbole und Individuenkonstanten eingeführt, die in die Relationssymbole eingesetzt werden, wobei jedes Relationssymbol eine zusätzliche Stelle für die Ereignisvariable bereithält. Die Relationssymbole für das Verb und die restlichen Angaben des Satzes werden durch Konjunktionen verbunden. Die in jedem Konjunktionsglied auftretende freie Ereignisvariable wird durch einen Existenzquantor gebunden.

Die logische Struktur eines Satzes (Davidson spricht nur von Handlungssätzen) lässt sich daher durch eine Existenzaussage darstellen, die in der Prädikatenlogik erster Stufe formuliert ist. Zum Beispiel 'I flew my spaceship to the morning star' würde Davidson in die logische Form $\exists x(Flew(I, myspace\ ship, x) \wedge To(morning\ star, x))$ bringen, vgl. [48, Essay 6]. Die Bedeutung des Satzes setzt sich somit aus der Bedeutung der einzelnen Bestandteile des Satzes und der gemeinsamen Beziehung zu einem darunterliegenden Ereignis (Underlying Event) zusammen. Bei Parson [103] werden Objekt und Subjekt selbst durch zweistellige Relationssymbole dargestellt, die eine Stelle für die Individuenkonstante, die das Handlungssubjekt symbolisiert und die zweite Stelle für die Ereignisvariable besitzen. Der Beispielsatz 'I fly my spaceship to the morningstar' lautet dann $\exists e(Fly(e) \wedge Subject(I, e) \wedge Object(myspace\ ship, e) \wedge To(morning\ star, e))$. Auf die Tempushandhabung von Parson möchte ich an der Stelle nicht eingehen, daher habe ich den Beispielsatz von Davidson ins Präsens gesetzt.

Diese Satzanalyse ist viel beachtet worden. In der Linguistik gibt es viele Befürworter dieser Ereignissemantik. Es finden sich aber auch Befürworter, die sowohl in der Linguistik als auch in der Philosophie beheimatet sind, vgl. z.B. [80, 103].

Nach Belnap weist diese Analyse vor allem dann Erklärungsmängel auf, wenn Fragen nach dem Verneinen von Handlungssätzen, dem aktiven Unterlassen von Handlungen hinsichtlich eines bestimmten Akteurs oder nach der Intention eines Akteurs,

2. Zustände, Sachverhalte und Ereignisse

eine Handlung zu vollziehen, gestellt werden bzw. wenn es sich um deontische Kontexte handelt. Da ein Satz zwar verneint werden kann, ein Ereignis jedoch nicht, stellt sich die Frage, wie Davidsons Analyse logischer Formen von Handlungssätzen diesem Problem begegnet. So wäre der Satz ‘I did not fly to the morning star’ je nach Kontext zu $\neg\exists x(\text{Flew}(I, \text{myspace ship}, x) \wedge \text{To}(\text{morningstar}, x))$ oder zu $\exists x(\text{Flew}(I, \text{myspace ship}, x) \wedge \neg\text{To}(\text{morningstar}, x))$ bedeutungsgleich.²

Belnap legt dar, dass der Ansatz, die logische Form von Handlungssätzen mittels modallogischer Operatoren anstelle von Davidsons Beschreibung in Logik 1. Stufe anzugeben, vorzuziehen ist. Der Satz $p = \text{‘I flew to the morning star’}$ würde mittels Stit-Paraphrase für den Akteur I lauten, ‘I see to it that I flew to the morning star’, so dass die Übersetzung $I \text{ dstit} : p$ und die Negation dementsprechend $\neg I \text{ dstit} : p$ kein Problem darstellen würden. Außerdem ist es in der Stit-Theorie von Vorteil, dass ein Satz der Form $\alpha \text{ dstit} : p$ verneint werden kann, in dem Sinne dass α nicht dafür gesorgt hat, dass p , bspw. weil β dafür gesorgt hat, so dass p dennoch wahr ist. Es kann aber auch durch die Möglichkeit der Mehrfachanwendung modaler Operatoren geklärt werden, dass jemand etwas nicht nur nicht getan hat, sondern aktiv dafür gesorgt, es nicht zu tun, $\neg\alpha \text{ dstit} : p$ versus $\alpha \text{ dstit} : \neg\alpha \text{ dstit} : p$.

Belnap meint den Grund zu kennen, warum Davidsons Ansatz fehlschlägt, „I have in mind the tendency of the program to play down the question of how agentive constructions embed in larger constructions“ [15, S. 778]. Da Davidsons Ansatz für eine logische Form von Handlungssätzen nach eigener Angabe auf den grammatikalischen bzw. logischen Funktionen der einzelnen Teile des Satzes aufbaut, vergisst er Belnap zufolge die andere Seite, die ein solcher kompositionaler Ansatz benötigt. „Davidson 1967 sets out to show how ‘the meanings of action sentences depend on their structure’ (Anm. d. Verf. [48, Essay 6 S. 105]), but does *not* begin with the aim of showing how the meanings of sentences that *contain* action sentences depend on *their* structure“ [15, S. 778].

Dies ist natürlich kein Grund anzunehmen, dass, falls Ereignisse eine eigene (ontologische) Kategorie sind, die Klasse der Handlungen nicht in diese Kategorie fal-

²Kritik, dass diese logische Darstellung auch für bestimmte Modifikatoren, die in Sätzen auftreten, nicht adäquat ist, kam z.B. von Stalnaker und Thomason [137]. Sie bemerkten, dass diese Analyse bei Sätzen mit bestimmten Attributen wie z. B. ‘partway’ oder auch anderen adverbialen Bestimmungen ebenfalls nicht zutreffend ist [137]. Die Negation und die Handhabung von einigen adverbialen Bestimmungen ist in Davidsons Analyse ein Problem. Ein neuerer Ansatz, der die logische Form eines Satzes ebenfalls als eine Existenzaussage über Eventualitäten (Ereignisse/Zustände) in der Prädikatenlogik 1. Stufe darstellt, findet sich beispielsweise in [103]. Parson behandelt auf S. 121f auch das Problem ‘partway’ und andere Probleme einer Ereignissemantik bei der Handhabung von adverbialen Bestimmungen.

len kann. Problematisch ist es anzunehmen, dass sich Beschreibungen für Handlungen bzw. die Klassifikation von Handlungssätzen allein aufgrund der Zuordnung der Handlungen in die Kategorie 'Ereignis' darstellen lassen. Nur weil sie in diese (ontologische) Kategorie fallen, erweist sich eine nur auf dieser Einordnung basierende Beschreibung nicht automatisch als vollständig.

Nun hat Davidson nicht gesagt, dass Handlungen eine eigene ontologische Subklasse von Ereignissen sind, sondern dass Ereignisse Handlungen sind, wenn es eine Beschreibung des Ereignisses gibt, unter der der Handelnde sein Tun absichtlich vollzieht [48, Essay 3]. D.h., ob es sich um eine Handlung handelt, wird auf der Beschreibungsebene entschieden. Dennoch versucht Davidson in Essay 6, die logische Form eines Handlungssatzes daraus abzuleiten, dass es ein Ereignis gibt, auf das die Beschreibung zutrifft. Er behauptet weiter, dass eine Handlung an dem Verb des Satzes bzw. an den einzelnen Teilen des Satzes erkennbar ist, und dass ein Handlungsverb sich durch ein mehrstelliges Prädikat beschreiben lässt, das eine Ereignis-Stelle aufweist. Ich möchte am Beispiel von 'es ist verboten, dass' aufzeigen, dass die einzelnen Bedeutungen der Teile eines Satzes noch nicht die Bedeutung des gesamten Satzes vermögen darzustellen. Dies ist einer der Kritikpunkte, die auch Belnap gegen Davidson vorbringt.

Der Satz 'es ist verboten, dass Hunde auf der Kommandobrücke sind' verweist implizit auf Handlungen, vgl. „the Restricted Complement Thesis“ z.B. festgehalten [15, S.787] bzw. [17, S.615]. Die These besagt, dass die Komplemente von deontischen Operatoren auf Stit-Sätze, die ausdrücken, dass ein bestimmter Akteur etwas vollbringt, eingeschränkt werden sollten. Demzufolge kann nur verboten werden, was man tun bzw. unterlassen kann. Es ist nicht verboten etwas zu sein, sondern dieses Sein nicht zu verhindern. Ein Grund hierfür liegt in einem anerkannten Prinzip:³ Sollen impliziert Können. Immer wenn ein Akteur etwas tun soll, muss es innerhalb seiner Möglichkeiten liegen, dies auch zu tun, wobei es für „sein Können“ erforderlich ist, eine Form von Kontrolle zu haben. Kontrolle über einen Zustand hat man jedoch nur über das Hervorbringen, Beibehalten oder Beenden des Zustandes.⁴ Ein Zustand ist somit nicht verboten, sondern das Hervorbringen, das Nichtverhindern oder auch das Fortbestehenlassen des Zustandes sind verboten. Obwohl 'dass Hunde auf der Kommandobrücke sind' eine Zustandsbeschreibung ist, drückt der Satz 'es ist verboten, dass ...' implizit aus, dass es verboten ist dafür zu sorgen, dass dieser Zustand, dass Hunde auf der Kommandobrücke sind, herbei geführt wird.

³Dieses Prinzip wurde u.a. von Kant vertreten. Eine neuere Verteidigung von '„Cannot“ implies „Not Ought“' findet sich in [84].

⁴Ich werde in den folgenden Abschnitten den Begriff 'Zustand', wie ich ihn fasse, näher erklären.

2. Zustände, Sachverhalte und Ereignisse

Ich werde den Satz ‘es ist verboten, dass p ’ auf Davidsons Analyse hin untersuchen. Wenn ein Handlungssatz einem Ausdruck ‘es ist verboten, dass’ folgen muss und sich die Bedeutung des Ausgangssatzes aus den Bedeutungen der Teile des Satzes und der Beziehung zu einem Ereignis zusammensetzen lässt, dann muss in Übereinstimmung mit Davidsons Analyse es der Fall sein, dass sich die Bedeutung des Satzes aus der Bedeutung des Ausdruckes ‘es ist verboten’, dessen Ergänzung dem Nebensatz ‘dass p ’ und der Beziehung zu einem Ereignis zusammensetzen.

Nehme man an, die Handlung und somit das Verb, welches die Ereignis-Stelle aufweist, sei im ersten Teilsatz ‘es ist verboten, dass ...’ zu finden. Wenn der Satz insgesamt ein Handlungssatz wäre, würde nach Davidson die logische Struktur lauten $\exists x \text{Verboten}(p, x)$, *Verboten* hätte eine Ereignisstelle und müsste den Satz als eine Beschreibung einer Aktivität auszeichnen. Nun ist ein Verbot kein Ereignis. Ein Ereignis in die Relation *Verboten* einzusetzen, ist daher nur sinnvoll, wenn die grammatikalisch notwendige Ergänzung p von *Verboten*, die selbst ein Satz ist, ein Handlungssatz ist, dessen Verb eine Ereignisstelle besitzt, so dass die Stelle in *Verboten* nur die Ergänzung zu dem mit p beschriebenen Ereignis ist. Der erste Teil des Satzes ‘es ist verboten, dass ...’ drückt allein also noch keinen Handlungssatz aus.⁵

Wenn p ein Handlungssatz ist, wäre die logische Form von p nach Davidson $\exists x(\text{Sein}(\text{Hunde}, x) \wedge \text{Auf}(\text{Kommandobrücke}, x))$. Die Aktivität müsste an dem Verb ‘Sein’ abzulesen sein. Dagegen würde Davidson selbst argumentieren, dass ‘Sein’ kein Handlungsverb ist und es keine Stelle für eine Ereignisvariable x offen lässt. Demzufolge drückt die logische Form von p keinen Handlungssatz aus. Da beide Teilstrukturen aber nicht durch die logische Form eines Handlungssatzes wiedergegeben werden können und das im Beispielsatz auftretende Relationssymbol *Sein* aufgrund des Verbes ‘Sein’ eigentlich keine Ereignisstelle x aufweisen sollte, kann es auch nicht die gemeinsame Beziehung zu einem Ereignis sein, die implizit eine Handlung beschreibt. Nach Davidson handelt es sich bei oben genanntem Verbot somit nicht um

⁵Ich bin jedoch nicht der Meinung, dass nur deshalb das Verb ‘Verboten’ keine Ereignisstelle haben kann. Es hat sie durchaus. Die Ereignisvariable wird jedoch nicht durch einen Existenzquantor in einer Existenzaussage gebunden. Nach meinem Verständnis wird sie im Falle, dass ein Zustand verboten wird, eher durch eine Allaussage wiedergegeben. Jedes Ereignis, was zu p führt, ist verboten, wobei ich hier ein weiteres Prädikat annehme und p als eine Zustandsaussage betrachte, $\forall x(\text{Cul}(x, p) \supset \text{Verboten}(x))$, wobei *Cul*(x, p) symbolisiert, dass x in p kulminiert, d.h. p ist der Zustand, der nach dem Ereignis x vorliegt, siehe [103]. Nehme man an, einer Handlung liegt ein Ereignis zugrunde, dann ist demzufolge jede Handlung verboten, die zu dem durch p beschriebenen Zustand führt. Allerdings ist auch in diesem Fall die Beziehung der Handlung zu einem Ereignis nicht ausreichend, um den Satz als Handlungssatz zu kennzeichnen.

einen Satz, der explizit oder implizit eine Handlung beschreibt.

Nun könnte man argumentieren, dass eine andere Beschreibung des Verbotes durchaus in eine logische Form eines Handlungssatzes gebracht werden kann, z.B. ‘es ist verboten, dass Hunde auf die Kommandobrücke geführt werden’. Da Davidson jedoch eine vollständige Analyse von Handlungssätzen erzielen wollte, ist die Frage, ob es mögliche Formulierungen (Verben) gibt, die mit seiner Analyse konform gehen, nicht relevant.

In diesem Zusammenhang zielt Belnaps Kritik darauf ab, dass es die Verbindung zwischen den Teilen eines Satzes sein kann, die deutlich macht, dass der gesamte Satz einen Handlungssatz beinhaltet und dass dies nicht notwendigerweise an einem Verb oder einem anderen Teil des Satzes bzw. dessen Beziehung zu einem darunterliegenden Ereignis festzumachen ist.⁶ Belnap steht der These ‘Actions-as-Events’ generell kritisch gegenüber. Die Existenzaussage eines Ereignisses reicht nicht aus, um eine Handlungsbeschreibung folgen zu lassen. Eine Handlung eines Akteurs ist es u.U. auch dann eine Handlung, wenn

Eine Analyse eines Handlungssatzes als Existenzaussage über Ereignisse wirft viele Fragen und Probleme auf. Im Folgenden verteidige ich, dass es sich bei Handlungen durchaus um Ereignisse handelt, nur dass die Analyse eines Handlungssatzes als Existenzaussage über Ereignisse nicht hinreichend ist. So werde ich in den Abschnitten dieses Kapitels die Begriffe *Zustand* und *Ereignis* untersuchen. Um Zustände und Ereignisse zu charakterisieren, verwende ich den Begriff des *Sachverhaltes*, Abschnitt 2.1. Dabei gehe ich auf die Beziehung zwischen den Begriffen ‘Sachverhalt’ und ‘Zustand’ sowie auf die Abgrenzung des Sachverhaltes von Propositionen ein. In den Abschnitten 2.2 und 2.3 stelle ich Unterschiede und Gemeinsamkeiten der Begriffe ‘Zustand’ und ‘Ereignis’ heraus. Im anschließenden Kapitel 3 werde ich zeigen, dass Handlungen Ereignisse zugrunde liegen, jedoch für eine Beschreibung als Handlung die Existenzaussage eines Ereignisses nicht hinreichend ist und stelle eine These der Stit-Paraphrasierung für Ereignisse vor.

2.1 Zustände und Sachverhalte

Auf den Begriff ‘Zustand’ trifft weder die Eigenschaft ‘statisch’ noch die Eigenschaft ‘dynamisch’ zu. Ein *Zustand* zeichnet sich dadurch aus, dass die Wahrheitsbedingung

⁶Auf die Unterteilung von Verben in Handlungsverben und Nicht-Handlungsverben gehe ich in Kapitel 3 ein.

2. Zustände, Sachverhalte und Ereignisse

für eine Zustandsbeschreibung in einem Zeitpunkt eines Geschichtsverlauf festgemacht wird. Ein Zustand liegt in einer Situation vor. Dann ist dies der Zustand in der Situation *s*. Für die Frage, ob ein Zustand vorliegt oder nicht, ist nur ein ausgezeichneter Zeitpunkt vonnöten. Eine Zeitspanne, die für eine Zuschreibung der Eigenschaft ‘statisch’ bzw. ‘dynamisch’ erforderlich wäre, ist es nicht nötig, um einen Zustand zu charakterisieren oder zu identifizieren. Etwas wird genau dann als statisch beschrieben, wenn über einen bestimmten Zeitraum keine Änderung eintritt. Dies impliziert, dass eine Zuschreibung als statisch nicht mit der Zuschreibung, dass dies ein Zustand sei, zusammenfallen kann.

Meist wird ‘etwas ist ein Zustand’ und ‘etwas ist statisch’ gleich gesetzt, z.B. Vendler [151]. Dabei handelt es sich meines Erachtens um ein Missverständnis, das aus der Tatsache resultiert, dass sprachlich nicht zwischen dem Bestehen eines Zustandstyps über einem Zeitraum und einem Zustand unterschieden wird. So kann der Satz ‘die Sonne scheint’ einen Zustand als auch das Anhalten des Zustandes beschreiben. Es ist lediglich aus dem Kontext ersichtlich, ob dieser Satz sich punktuell auf einen bestimmten Zeitpunkt bezieht ‘die Sonne scheint jetzt gerade’ oder ob er das Anhalten des Zustandes beschreibt, dass die Sonne über einen bestimmten Zeitraum hinweg scheint, z.B. ‘heute scheint die Sonne den ganzen Tag’. Dass Sätze sowohl eine Zustandsbeschreibung als auch eine Beschreibung des Anhaltens des Zustandes sein können, zeigt, dass der Sinn eines Satzes nicht der Zustand ist, sondern der Zustand etwas in einer bestimmten Situation ist, dass durch Sätze beschrieben werden kann.

Ist der ausgedrückte Sinn eines Satzes eine Proposition und korrespondiert jeder Proposition ein Sachverhalt, der in einer Situation besteht oder nicht besteht, dann wird ein Zustand durch eine Menge von Sachverhalten charakterisiert, die in der Situation bestehen, in der der Zustand vorliegt. Dafür ist es unumgänglich, den Begriff des Sachverhaltes und der Proposition zu definieren, so dass zwischen der Beschreibung eines Zustandes und dem Objekt, das beschreibt, unterschieden werden kann.

Sachverhalte (SVH) und Sachverhaltseigenschaften (SVH-Eigenschaften)

Ein *Sachverhalt* ist als eine notwendig⁷ existierende, abstrakte Entität, die in gewisser

⁷In diesem Kapitel bezieht sich ‘notwendig’ auf eine metaphysische Modalität. Etwas ist notwendig, wenn es in allen Geschichtsverläufen (Welten) der Fall ist. In den anderen Kapiteln 1 und folgende werde ich weitere Modalitäten betrachten. Zum einen wird die historische Möglichkeit und Notwendigkeit definiert. Sie unterscheidet sich von der hier verwendeten Notwendigkeit, die alle Geschichtsverläufe und alle Situationen in diesen Geschichtsverläufen betrachtet. Im Falle der historischen Notwendigkeit ist die Menge der relevanten Geschichtsverläufe insoweit eingeschränkt, dass nur die Geschichts-

Weise zu anderen Sachverhalten in Relationen steht. Wie Plantinga festhält, können Sachverhalte andere Sachverhalte inkludieren bzw. präkludieren. Ein Sachverhalt inkludiert genau dann einen anderen, wenn es nicht möglich ist, dass ersterer besteht, aber letzterer nicht besteht. Entsprechend präkludiert ein Sachverhalt einen anderen Sachverhalt genau dann, wenn letzterer nicht besteht, sobald ersterer besteht. Eine mögliche Welt bei Plantinga ist ein Sachverhalt S , der maximal in dem Sinne ist, dass jeder andere, *nicht-transiente* Sachverhalt T entweder von S inkludiert oder präkludiert wird.⁸ Ein nicht-transienter Sachverhalt ist ein Sachverhalt, der, wenn er besteht, zu jedem Zeitpunkt besteht. Sein Bestehen innerhalb einer Welt bzw. eines Geschichtsverlaufs ist *zeitinvariant*. Es ist nicht erforderlich, für das Bestehen des nicht-transienten Sachverhaltes einen Zeitpunkt oder ein Zeitintervall zu ergänzen. Der Sachverhalt beinhaltet einen Zeitpunkt oder ein Zeitintervall, falls es nötig ist bzw. ist die Angabe eines Zeitintervalles unnötig.⁹ Nicht-transiente Sachverhalte sind z.B. mit Zeitangaben fixierte Sachverhalte, wie ‘der Sachverhalt, dass die olympischen

verläufe betrachtet werden, die einen bestimmten Zeitpunkt passieren und eine gemeinsame Vergangenheit haben. Für dieses Kapitel bedeutet ‘notwendigerweise’, dass alle Geschichtsverläufe betrachtet werden.

⁸Plantinga vertritt somit eine aktualistische Repräsentation von möglichen Welten, der als „ersatzism“ bzw. „actualism“ bzw. „moderate realism“ bezeichnet wird. Es existiert nur das, was aktual existiert. Mögliche Welten existieren aktual als abstrakte Objekte. Ein weiterer Vertreter, der ebenfalls mögliche Welten als aktual existierende, abstrakte Objekte definiert, ist Adams [2]. Adams fasst mögliche Welten, „world-stories“, als maximale, konsistente Menge von Propositionen auf. Er bezeichnet dies als eine Art des „soft-actualism“. Stalnaker in [132] hingegen kennzeichnet mögliche Welten als Eigenschaften, die die aktuelle Welt haben könnte. In [133] verteidigt Stalnaker den moderaten Realismus, dass mögliche Welten existieren und dass die Ansicht, dass „actual“ ein indexikalischer Ausdruck ist, von einem semantischen Standpunkt jedoch nicht von einem metaphysischen Standpunkt korrekt ist. Er verneint jedoch die Lewische These, dass mögliche Welten „things of the same sort as the actual world“ sind [93, S.86]. Da für Stalnaker die aktuelle Welt genau das ist, was in der Realität existiert und nicht mehr: „the moderate realist believes that the only possible worlds there are — the ways things might have been — are (like everything else that exists at all) elements of our actual world“ [134, S.50]. In der Einleitung zu [136] fasst er das kurz zusammen: Eine mögliche Welt ist eher ein möglicher Zustand, den die Welt haben könnte. Dies setzt einen Zustandsbegriff voraus, der unabhängig von Welt ist, und ich nicht vertreten würde. Eine mögliche Welt ist eine Option/Alternative oder Möglichkeit, wie die Welt auch beschaffen sein könnte. Stalnaker sagt weiter, dass mögliche Welten keine konkreten Objekte noch konkrete Situationen sein können.

⁹Ich unterscheide mich hierin von Pollock, auf den die Formulierung ‘transient’ zurückgeht. Er sieht lediglich nicht-transiente Sachverhalte als notwendig existierend an. Transiente Sachverhalte existieren für ihn nur in Welten, in denen sie bestehen [109]. Ich unterscheide jedoch existieren und bestehen. Sachverhalte existieren auch dann, wenn sie nicht bestehen. Sachverhalte existieren unabhängig von ihren Konstituenten.

2. Zustände, Sachverhalte und Ereignisse

Sommerspiele 2000 in Sydney stattfinden', oder Sachverhalte, die keinen Zeitindex benötigen, wie der Sachverhalt, 'dass $2 + 1 = 3$ ', vgl. z.B. [107, Essays 2,5 und 8].

Allerdings individuiert diese Definition Sachverhalte über deren Bestehen in möglichen Welten (Geschichtsverläufen). Somit sind alle notwendigerweise bestehenden Sachverhalte identisch. Dies ist unplausibel. Daher möchte ich eine Ansicht Zaltas aufgreifen, vgl. [169], der einen Sachverhaltsbegriff einführt, der Sachverhalte über runde Quadrate 'runde Quadrate sind rund' zulässt und es ermöglicht, zwischen Sachverhalten wie '2+1=3' und 'alle geraden Zahlen sind durch zwei teilbar' zu unterscheiden.¹⁰

Jedem Sachverhalt ordnet Zalta eine korrespondierende Sachverhaltseigenschaft zu, so dass ein Sachverhalt ein Objekt charakterisieren kann, indem das Objekt diese Eigenschaft enkodiert („encode“) [169]. Nach Zalta gibt es zwei Beziehungen zwischen einem Objekt und einer Eigenschaft. Ein Objekt kann eine Eigenschaft enkodieren („encode“) oder instanziiieren („exemplify“). Eigenschaften, die ein Objekt instanziiieren kann, sind ihm external zu- oder abschreibbar. Eigenschaften, die ein Objekt enkodieren kann, sind ihm intrinsisch. Nicht-abstrakte Dinge instanziiieren Eigenschaften. Ein abstraktes Objekt, das durch das Haben von Eigenschaften bestimmt ist, diese aber nicht erfüllen muss, enkodiert diese Eigenschaften. Wenn ein abstraktes Objekt Eigenschaften enkodiert, sind es diese Eigenschaften, die das Objekt dann individuiieren. Beispielsweise ist das runde Quadrat von Mally [98] enkodiert (bestimmt durch) die Eigenschaften 'rund zu sein' und 'ein Quadrat zu sein', ohne dass es diese Eigenschaften instanziiiert (erfüllt). Wenn ein abstraktes Objekt eine Eigenschaft enkodiert, dann enkodiert es sie notwendigerweise. Abstrakte Objekte sind dann identisch, wenn sie dieselben Eigenschaften enkodieren. Nicht-abstrakte, konkrete („concrete“ in [169]) Dinge, wie z.B. Hans, London, ein bestimmter Tisch etc. werden über das Instanziiieren von Eigenschaften individuiert. Somit sind sie genau dann identisch, wenn sie die dieselben Eigenschaften instanziiieren. Bei Zalta können sie keine Eigenschaften enkodieren. Ich würde hingegen sagen, dass sie genau die Eigenschaften instanziiieren, die sie enkodieren. In Mallys Terminologie formuliert, erfüllen sie genau die Eigenschaften, durch die sie bestimmt werden.

Für die Enkodierung von Eigenschaften ist es möglich, dass ein abstraktes Objekt

¹⁰Zalta [169] bezieht sich in seinem Aufsatz auf Barwise und Perrys Situationssemantik [14]. Barwise spricht von dem „Granularity Problem“ und bemerkt in [13], dass es in möglichen Welten unmöglich ist, „to make the kind of fine-grained distinctions that appear to be needed for accounts of propositional attitudes or of information“. Dies trifft jedoch nur zu, wenn man mögliche Welten sich als in irgendeiner Weise konsistente Repräsentationen vorstellt und keine unmöglichen Objekte zulässt.

sowohl die Eigenschaft F nicht enkodiert als auch die Eigenschaft nicht- F nicht enkodiert. Für Instanziierung von Eigenschaften ist diese Unvollständigkeit, wie Zalta es nennt, nicht möglich, [169, S.395ff]. Instanziiert ein Objekt, die Eigenschaft ‘rund’ nicht, dann instanziiert es die Eigenschaft ‘nicht-rund’. Es ist jedoch möglich, dass ein abstraktes Objekt die Eigenschaft ‘ F ’ enkodiert und ‘ F ’ nicht instanziiert. Ein rundes Quadrat enkodiert die Eigenschaften ‘rund zu sein’ und ‘quadratisch zu sein’. Jedoch instanziiert es die Eigenschaften ‘rund zu sein’ und ‘quadratisch zu sein’ nicht.

Das Objekt ‘Sherlock Holmes’ beispielsweise enkodiert die Eigenschaften ‘Mann zu sein’, ‘Detektiv zu sein’ etc. und instanziiert die Eigenschaften ‘Romanfigur von Doyle zu sein’, ‘nicht real zu sein’ etc. Aber er enkodiert auch Eigenschaften, deren Negation er instanziiert. Z.B. die Eigenschaft ‘kein Mann zu sein’ wird durch das Objekt ‘Sherlock Holmes’ instanziiert und die Eigenschaft ‘ein Mann zu sein’ enkodiert es. Dies ist nur für abstrakte Objekte möglich. Abstrakte Objekte, wie die Zahl 4, wird bestimmt durch die Eigenschaften ‘eine natürliche Zahl zu sein’ und ‘durch vier teilbar zu sein’ etc. Sie erfüllt diese Eigenschaften auch. In welcher Beziehung steht die Zahl 4 mit der Eigenschaft ‘mordlustig zu sein’. Sie erfüllt diese Eigenschaft nicht, d.h. sie erfüllt die Eigenschaft ‘nicht mordlustig zu sein’. Aber wird die Zahl 4 durch die Eigenschaft ‘nicht mordlustig zu sein’ bestimmt? Meines Erachtens wird sie genausowenig durch die Eigenschaft es nicht zu sein wie es zu sein bestimmt. Die Eigenschaft spielt bei der Individuierung der Zahl 4 keine Rolle. Sie wird möglicherweise durch die Eigenschaft bestimmt ‘sowohl die Eigenschaften ‘nicht mordlustig zu sein’ und ‘mordlustig zu sein’ nicht zu enkodieren’.¹¹

Weiterhin nimmt Zalta an, dass für jede Gruppe von Eigenschaften ein abstraktes Objekt existiert, das genau diese Eigenschaften und keine anderen enkodiert. Für Zalta scheint diese Annahme sehr plausibel. Sie postuliert jedoch eine unübersehbar große Anzahl abstrakter Objekte. Allerdings garantiert die Aussage auch, dass für notwendigerweise bestehende Sachverhalte (SVH) jeweils ein anderes abstraktes Objekt existiert, das die zugehörige Sachverhalteigenschaft (SVH-Eigenschaft) enkodiert, vgl. [169]. Zu dem Sachverhalt ‘dass ein sprechender Esel existiert’ gehört die SVH-Eigenschaft ‘so beschaffen zu sein, dass ein sprechender Esel existiert’. Somit sind notwendigerweise bestehende Sachverhalte nicht notwendigerweise identisch, da sie durch verschiedene Objekte enkodiert werden können. Aber welche Objekte enkodieren bzw. erfüllen SVH-Eigenschaften?

¹¹Dies steht mit Leibniz Identitätsdefinition nicht im Widerspruch.

Sachverhalte in Situationen und Geschichten

Sachverhalte bestehen in Situationen, in Welten und in Geschichtsverläufen. Wenn ein Sachverhalt in einem Geschichtsverlauf besteht, so existiert eine Situation in diesem Geschichtsverlauf, die die zugehörige Sachverhaltseigenschaft erfüllt. Ebenso besteht ein Sachverhalt in einer Situation, wenn die SVH-Eigenschaft von der Situation erfüllt wird. Wenn Objekte Eigenschaften erfüllen, sind diese Objekte auch die ersten Kandidaten, für die Antwort auf die Frage, welche Objekte diese Eigenschaften enkodieren. Situationen bzw. Geschichtsverläufe erfüllen Sachverhaltseigenschaften nicht nur, sie enkodieren sie auch. Da Erfüllen und Enkodieren nicht notwendigerweise zusammenfällt, gibt es möglicherweise eine Situation, die die SVH-Eigenschaft zu '2+1=3' enkodiert, aber die SVH-Eigenschaft zu dem Sachverhalt 'alle geraden Zahlen sind durch zwei teilbar' nicht enkodiert, auch wenn sie möglicherweise beide erfüllt sind. Werden Sachverhalte durch SVH-Eigenschaften individuiert und sind Situationen die Objekte, die diese Eigenschaften enkodieren, ohne diese notwendigerweise zu instanziiieren, dann spreche ich abkürzend davon, dass eine Menge von Sachverhalten die Situation vollständig bestimmt. Die Menge an zugehörigen SVH-Eigenschaften werden durch die Situation enkodiert und umgekehrt individuiieren SVH-Eigenschaften (und damit auch die Menge an Sachverhalten) die Situation.

Mit diesen eben beschriebenen Beziehungen zwischen Objekt und Eigenschaft ist eine Situation als ein Objekt definierbar, welches SVH-Eigenschaften enkodiert und diese Eigenschaften auch instanziiieren kann, wobei SVH-Eigenschaften die Situation vollständig individuiieren. Die SVH-Eigenschaft zu dem Sachverhalt 'die olympischen Sommerspiele finden in Sydney statt' wäre dann die Eigenschaft 'so beschaffen zu sein, dass die olympischen Sommerspiele in Sydney stattfinden'. Die SVH-Eigenschaft zu dem Sachverhalt 'es regnet' ist 'so beschaffen, dass es regnet'. Situationen können diese Eigenschaften instanziiieren.

In einer Situation, die eine bestimmte SVH-Eigenschaft instanziiiert, besteht der zugehörige Sachverhalt. Situationen und Welten sind nach Zalta ähnlich in ihrer Struktur. Eine Situation wird über eine Menge von SVH-Eigenschaften individuiert. Im Falle der Welt nimmt Zalta an, dass es eine maximale und in gewisser Weise mögliche Menge an SVH-Eigenschaften gibt, die die Welt enkodiert. Die SVH-Eigenschaften sind nach Zalta das, was die Welt identifiziert [169, S.400]. Dann ist es möglich, dass in zwei Welten, dieselben nicht-transienten Sachverhalte bestehen, aber aufgrund dessen, dass die Situationen verschiedene SVH-Eigenschaften enkodieren, die Welten (Geschichtsverläufe) verschieden sind.

2.1. Zustände und Sachverhalte

Dass eine Menge von Sachverhalten möglich ist, kann man wie folgt verstehen:¹² Nach Plantingas Definition von Inklusion und Präklusion kann eine Menge an Sachverhalten als möglich definiert werden, wenn kein Sachverhalt der Menge durch die anderen präkludiert wird bzw. im Falle eines einzelnen Sachverhaltes, wenn der einzelne Sachverhalt bestehen kann. Es gibt zudem für den Zeitpunkt in der Situation eine maximale und mögliche Menge an Sachverhalten, die in dieser Situation bestehen und diese individuieren. Für einen Zeitpunkt maximal bedeutet, dass jeder zu diesem Zeitpunkt bestehende Sachverhalt durch die Teilmenge inkludiert wird. Präklusion und Inklusion von Sachverhalten sind durch das Bestehen von Sachverhalten in Situationen definiert, somit durch das Instanzieren von SVH-Eigenschaften, die bestimmte Situationen erfüllen. Erfüllbarkeit von SVH-Eigenschaften ist konsistent und im Gegensatz zum Enkodieren durch Situationen ist sie auch vollständig. Für jede Situation und Sachverhaltseigenschaft gilt, dass die Situation die SVH-Eigenschaft entweder erfüllt oder sie erfüllt die Eigenschaft, diese Eigenschaft nicht zu erfüllen.¹³

Die Sachverhalte, die in einer Situation bestehen, müssen jedoch nicht maximal sein, d.h. dass sie nicht notwendigerweise alle zeitinvarianten (nicht-transienten) Sachverhalte präkludieren oder inkludieren, sondern es können zeitabhängige Sachverhalte sein, die in einer Situation bestehen. Ein Sachverhalt 'es regnet in Bochum' ist ein Sachverhalt, dessen Bestehen zeitabhängig bzw. situationsabhängig ist. Hingegen ist der Sachverhalt 'es regnet am 9.4.2011 um 5:17 Uhr in Bochum' zeitinvariant in dem Sinne, dass er bei festgelegtem Geschichtsverlauf besteht bzw. nicht besteht. Sein Bestehen ist somit nur abhängig vom Geschichtsverlauf und nicht vom Zeitpunkt, der die Situation im Geschichtsverlauf bestimmt. Eine maximal mögliche Menge an Sachverhalten oder ein maximaler Sachverhalt ist dann so definiert, dass er jeden zeitinvarianten (situationsunabhängigen) Sachverhalt entweder präkludiert oder inkludiert. Er stellt somit eine mögliche Welt bzw. einen möglichen Geschichtsverlauf dar, wie von Pollock [109] und Plantinga [106, 107, Essay 9 S.189, Essay 10 S.195] vorgeschlagen.¹⁴

Somit kann umgekehrt ein Geschichtsverlauf als eine Menge von geordneten Zeitpunkt, d.h. von Situationen, aufgefasst werden. Die Sachverhalte, die entstehen, indem

¹²Dabei bedeutet 'möglich' nicht, dass die Menge der Sachverhalte möglicherweise existiert. Jeder Sachverhalt existiert notwendigerweise. 'Möglich' bezieht sich auf das Bestehen des Sachverhaltes.

¹³Beim Enkodieren einer Eigenschaft war es möglich, dass das Objekt weder die Eigenschaft noch die Eigenschaft, nicht diese Eigenschaft zu enkodieren, enkodiert.

¹⁴Im Hinblick auf die BT Struktur spreche ich statt wie Pollock und Plantinga von Welten von Geschichtsverläufen, insbesondere deswegen, weil ich den Begriff der Welt in Kapitel 5 für die Definition von Nachbarschaftssystemen anderweitig verwende.

2. Zustände, Sachverhalte und Ereignisse

den in den Situationen bestehenden Sachverhalten der jeweilige Zeitpunkt hinzugefügt wird, sind dann zeitinvariant. Zusammengefasst stellen sie eine maximal mögliche Menge an Sachverhalten dar, die jeden zeitinvarianten Sachverhalt inkludieren bzw. präkludieren. Aufgrund der Maximalität der bestehenden Sachverhalte in den Situationen ist die Menge der durch die Situationen und Zeitpunkte zeitinvarianten Sachverhalte ebenfalls maximal.

Eine Frage ist dabei, in welche ontologische Kategorie Objekte fallen, wenn sie die Eigenschaft 'zu sein' zwar enkodieren jedoch nicht erfüllen. Ein Sachverhalt besteht in einer Situation, wenn die Situation die zugehörige SVH-Eigenschaft instanziiert. Betrachte man die heutige Situation zu diesem Zeitpunkt in diesem Geschichtsverlauf, dann besteht der Sachverhalt 'Sherlock Holmes ist ein Mann' nicht, obwohl die zugehörige SVH-Eigenschaft möglicherweise von Situationen enkodiert und instanziiert wird. Das Objekt 'Sherlock Holmes' instanziiert in der heutigen Situation jedoch die Eigenschaft 'Mann zu sein' nicht. Somit instanziiert die Situation die SVH-Eigenschaft nicht und der Sachverhalt besteht nicht. Ist es notwendigerweise der Fall, dass abstrakte Objekte bestimmte Eigenschaften enkodieren, ist es hingegen kontingent, welche Eigenschaften ein abstraktes Objekt erfüllt. Der Sachverhalt 'Sherlock Holmes ist eine Romanfigur' besteht jedoch, da die zugehörige SVH-Eigenschaft durch die heutige Situation erfüllt wird. Es ist nicht erforderlich, dass 'Sherlock Holmes' die Eigenschaft 'konkretes, nicht abstraktes Objekt zu sein' instanziiert, um die Eigenschaft 'Romanfigur zu sein' zu instanziiieren. Es ist möglich, Sachverhalte als bestehend in Situationen zu definieren, obwohl die Objekte, die zu dem Sachverhalt gehören, nicht in der Welt der konkreten Dinge instanziiert sind.

Das abstrakte Objekt 'Sherlock Holmes' wird durch (enkodiert) eine Menge an Eigenschaften bestimmt, die das Objekt nicht notwendigerweise in jeder Situation instanziiert. Die SVH-Eigenschaft zu dem Sachverhalt 'dass diese Eigenschaften durch das Objekt instanziiert werden' muss nicht in der Situation instanziiert sein. Jedoch der Sachverhalt 'dass diese Eigenschaften das Objekt bestimmen' besteht und zwar notwendigerweise.

Linsky und Zalta unterscheiden daher nicht nur zwischen abstrakten und konkreten, sondern außerdem zwischen konkreten, im aktuellen Geschichtsverlauf realisierten Dingen und kontingenterweise nicht-konkreten Dingen, die in anderen Geschichtsverläufen, wenn dieser der aktuelle Geschichtsverlauf wäre, zu den in diesem Geschichtsverlauf konkreten Dingen gehören würden. Sie sind nicht abstrakt, sondern kontingenterweise nicht-konkret. Sie werden durch das Instanziiieren bestimmter Eigenschaften in den Geschichtsverläufen, in denen sie konkret sind, individuiert und

nicht durch das Enkodieren von Eigenschaften. Ein Beispiel von kontingenterweise nicht-konkreten Dingen wäre z.B. meine Schwester, die ich möglicherweise in einem anderen Geschichtsverlauf gehabt hätte.

Abstrakte Dinge, die notwendigerweise existieren und durch die Eigenschaften, die sie enkodieren, individuiert werden, sind bspw. Sachverhalte, Propositionen, Zahlen, Zustandstypen, Ereignistypen.¹⁵ Da ich annehme, dass abstrakte Dinge durch das Enkodieren von Eigenschaften individuiert werden, enkodieren sie Eigenschaften notwendigerweise. Für konkrete und kontingenterweise nicht-konkrete Dinge gilt, da sie genau die Eigenschaften enkodieren, die sie instanziierten, dass sie Eigenschaften kontingenterweise enkodieren und erfüllen. Somit können sie auch in verschiedenen Geschichtsverläufen verschiedene Eigenschaften enkodieren und erfüllen.

Die im aktuellen Geschichtsverlauf kontingenterweise nicht-konkreten Dinge können somit die Eigenschaft 'konkret zu sein' in Situationen anderer Geschichtsverläufe instanziierten. Dies können abstrakte Dinge nicht. Dies ist der Ausweg den Zalta und Linsky vorschlagen [94, 95], um dem Problem der nicht-existierenden Objekte zu entgehen. Diese Unterscheidung in konkrete, kontingenterweise nicht-konkrete und abstrakte Objekte ermöglicht, dass ein Objekt eine Eigenschaft in einer Welt instanziierten kann, ohne selbst in dieser Welt realisiert zu sein.

Plantingas Begriff der „individual essence“ hat dieselbe Funktion, siehe [107, Essay 5,7,8]. Einem Objekt kommt Existenz notwendigerweise zu, auch wenn das Objekt kontingenterweise existiert. Jedes Objekt hat seine „individual essences“ [107, Essay 8 S.158], die wie Eigenschaften generell bei Plantinga notwendigerweise existierend sind und die ein Objekt in einer Welt genau dann instanziiert, wenn es notwendigerweise der Fall ist, dass, wenn diese Welt die aktuelle Welt wäre, das Objekt in ihr existiert. Eigenschaften sind dabei nicht Mengen von Objekten, die diese Eigenschaften instanziierten, sondern sie sind notwendigerweise existierende Entitäten. Sie können weltindiziert sein. Die Eigenschaft 'weise zu sein' kommt Sokrates in diesem aktuellen Geschichtsverlauf h (Welt) zu [107, Essay 7 S.141]. Angenommen es gibt Geschichtsverläufe, in denen Sokrates die Eigenschaft 'weise zu sein' nicht erfüllt, erfüllt er jedoch in allen Geschichtsverläufen, die Eigenschaft 'weise in diesem Geschichtsverlauf h zu sein'. Dies ist ein Grund für Plantinga, warum Eigenschaften in allen Welten existieren müssen.

Somit kann Plantinga, die Eigenschaft in einer Welt zu instanziierten, von der Instanziiierung des Objektes in der Welt trennen. Jedoch ist es unplausibel, warum die

¹⁵Zustände und Ereignisse selbst sind jedoch entweder konkret oder kontingenterweise nicht-konkret.

2. Zustände, Sachverhalte und Ereignisse

Eigenschaft 'weise zu sein im Geschichtsverlauf h ' im Geschichtsverlauf h eine andere Eigenschaft als 'weise zu sein' ist. Ich würde erwarten, 'dass Sokrates in h weise ist' ein anderer Sachverhalt als derjenige ist, 'dass Sokrates in h weise ist'. Die Eigenschaft 'weise zu sein' ist meines Erachtens jedoch in beiden Sachverhalten dieselbe. Es ändert sich die Situation und somit die SVH-Eigenschaften, die die jeweilige Situation in den Geschichten individuieren, jedoch ändert sich nicht die Eigenschaft 'weise zu sein'.

Anstelle die Eigenschaften mit Weltindizes zu versehen, ist es meines Erachtens unproblematisch, Sachverhalte in anderen Geschichtsverläufen über Geschichtsverläufe zuzulassen. Der Sachverhalt 'dass Sokrates weise ist' ist offensichtlich ein anderer Sachverhalt als 'dass Sokrates in h weise ist'. Ersterer besteht kontingenterweise, wohingegen letzterer notwendigerweise besteht bzw. nicht besteht. Plantinga erreicht eine Trennung von der Existenz des Objektes und der Instanziierung von Eigenschaften in einer Welt jedoch zu dem Preis, dass 'weise zu sein in h ' für verschiedene Geschichten h verschiedene Eigenschaften sind. Zalta und Linsky ermöglichen die Trennung von Instanziierung von Eigenschaften und Enkodieren von Eigenschaften und somit die Aussage der Existenz des Objektes, ohne dass das Objekt im aktuellen Geschichtsverlauf konkret realisiert sein muss (Trennung von existent und konkret zu sein).

Zalta und Linsky [94, 95] halten jedoch fest, dass bestimmte Eigenschaften „concreteness-entailing“ sind. Unter „concreteness-entailing“ (Konkretheit erfordern) verstehen sie, dass ein Objekt, das bspw. die Eigenschaft 'Schwester zu sein' instanziiert, diese dann und nur dann instanziiert, wenn es selbst konkret ist. Die Eigenschaften, die Konkretheit erfordern, können bei Zalta und Linsky nur durch konkrete Objekte erfüllt sein. Eine Eigenschaft 'Mann zu sein' erfordert Konkretheit, jedoch die Eigenschaft 'eine Romanfigur zu sein' erfordert dies nicht. So kann eine Situation die SVH-Eigenschaft 'so beschaffen sein, dass Sherlock Holmes eine Romanfigur ist' instanziiieren, ohne dass Sherlock Holmes ein konkretes Objekt in dieser Situation sein muss. Die SVH-Eigenschaft des Sachverhaltes 'Sherlock Holmes ist ein Mann' kann jedoch in einem Geschichtsverlauf, wo die Eigenschaft 'Mann zu sein' Konkretheit erfordert, durch eine Situation enkodiert werden. Sie kann nur nicht instanziiert werden, da Sherlock Holmes die Eigenschaft 'Mann zu sein' nicht instanziiieren kann, wenn er nicht konkret ist. Wenn die Eigenschaft 'Mann zu sein' Konkretheit erfordert, ist dies nur möglich, wenn Sherlock Holmes ein konkretes Objekt ist. Grund hierfür ist, dass die Menge an Sachverhaltseigenschaften, die in einer Situation instanziiert sind, in gewisserweise möglich also nicht inkonsistent sein können. Die SVH-Eigenschaften 'Mann zu sein ist Konkretheit erfordern', 'Sherlock Holmes ist ein Mann' und 'Sher-

lock Holmes ist ein abstraktes Objekt' bestehen nicht alle in einer Situation, da sie einander präkludieren.¹⁶

Die konkrete Existenz der Objekte des Sachverhaltes ist wiederum ein Sachverhalt, der relevant für das Bestehen des Sachverhaltes sein kann, wenn die Eigenschaften, die der Sachverhalt dem Objekt zuschreibt, Konkretheit erfordert. Jedoch ist er irrelevant für die Existenz des abstrakten Objektes des Sachverhaltes. Auf der anderen Seite enthält ein Sachverhalt Objekte, die selbst abstrakt sind und die in einer Situation oder einem Geschichtsverlauf konkret sein können, jedoch nicht müssen, damit die Situation den Sachverhalt instanziiert.

Zusammengefasst ist eine Situation durch eine Menge an Sachverhalten und deren SVH-Eigenschaften vollständig bestimmt. Einige Sachverhalte sind dabei situationsabhängig, also transient. Sie sind nicht nur durch Zeitpunkte, sondern durch mögliche Geschichtsverläufe indiziert. Dass ein Sachverhalt zu einem festen Zeitpunkt in einem Geschichtsverlauf besteht, ist dann ein anderer Sachverhalt, der nicht-transient ist, und dessen Bestehen bzw. Nichtbestehen notwendigerweise ist. Ein Geschichtsverlauf ist eine zeitliche Abfolge von Situationen. Eine Situation bzw. ein Geschichtsverlauf wird durch SVH-Eigenschaften individuiert, die die jeweilige Situation bzw. das Paar aus Situation und Geschichtsverlauf enkodiert. In dem aktuellen Geschichtsverlauf zu gegebenem aktuellen Zeitpunkt (aktuelle Situation) existiert eine Menge von Sachverhalten, die bestehen und somit die *Tatsachen* in diesem Geschichtsverlauf darstellen. Im Folgenden werde ich auf Beispiele für Tatsachen zurückkommen. Die möglichen Geschichtsverläufe in der aktuellen Situation beinhalten Situationen, die durch Sachverhaltseigenschaften individuiert werden, deren Sachverhalte noch nicht bestehen, bis zu dem Zeitpunkt, dass die Situation eintritt und somit aktual wird, und die zugehörigen Sachverhaltseigenschaften durch die Situation instanziiert werden. Für eine Tatsache ist somit das Bestehen eines Sachverhaltes in einer vergangenen bzw. in der aktuellen Situation im konkreten Geschichtsverlauf notwendig. Eine Übersicht über einen Teil der in diesem Abschnitt diskutierten Begriffe findet sich in Tabelle 2.1.

¹⁶Nach Linsky und Zalta ist es notwendigerweise der Fall ist, dass ein Objekt konkret ist, wenn es die Eigenschaft *F* instanziiert, wenn diese Konkretheit erfordert. Dies lässt zwei Möglichkeiten zu, entweder der Sachverhalt 'das runde Quadrat ist rund' besteht, dann erfordert 'rund' nicht Konkretheit. Oder man ist der Meinung, dass 'rund' Konkretheit erfordert, dann bestehen Sachverhalte der Form 'das runde und nicht-runde Quadrat ist rund' nicht. Es kann somit Situationen geben, die den Sachverhalt 'das runde Quadrat ist nicht-rund' erfüllen, aber nur dann wenn in dieser Situation der folgende Sachverhalt nicht erfüllt ist: Die Eigenschaft 'nicht-rund zu sein' ist nicht identisch mit der Eigenschaft 'rund zu sein nicht zu erfüllen'. Ich möchte jedoch solche Situationen im Folgenden unbeachtet lassen. Die Regel der doppelten Negationsbeseitigung wird beibehalten.

2. Zustände, Sachverhalte und Ereignisse

Zustand, Sachverhalt oder Proposition

Ein Zustand ist aus zweierlei Gründen kein Sachverhalt. Zustände existieren nicht notwendigerweise. Ein Zustand existiert innerhalb eines Geschichtsverlaufs, nur solange er vorliegt. Der Zustand, dass ich überzeugt bin, dass φ , existiert, wenn ich überzeugt bin, dass φ . Der Sachverhalt, dass ich überzeugt bin, dass φ , existiert unabhängig davon, ob er möglicherweise besteht oder nicht besteht. Der zweite Unterschied zwischen Sachverhalten und Zuständen, wie ich sie verstehe, ist, dass Sachverhalte wesentlich feinkörniger als Zustände sind.

Mit Feinkörnigkeit meine ich Folgendes: Mehrere Sachverhalte können in einer Situation in einem Geschichtsverlauf bestehen, die ein und denselben Zustand in diesem Geschichtsverlauf charakterisieren. ‘Der Erfinder der Brille sitzt unter dem Baum’ und ‘Benjamin Franklin sitzt unter dem Baum’ charakterisieren bei fixierter Situation im heutigen Geschichtsverlauf denselben Zustand. Die Sätze drücken jedoch verschiedene Propositionen aus, denen auch verschiedene Sachverhalte korrespondieren. Es hätte zu einem früheren Zeitpunkt einen möglichen anderen als den heutigen Geschichtsverlauf geben können, in dem Benjamin Franklin nicht dieselbe Person wie der Erfinder der Brille ist. In diesem Geschichtsverlauf können beide Personen zu demselben Zeitpunkt unter einem Baum sitzen. Dann bestehen in diesem alternativen Geschichtsverlauf die zwei Sachverhalte bzw. sind in diesem Geschichtsverlauf die beiden oben genannten Propositionen wahr. Aber es existieren zwei verschiedene Zustände in dem Geschichtsverlauf, die durch diese Sachverhalte charakterisiert werden. Nämlich liegt der Zustand vor, in dem sich Benjamin Franklin unter dem Baum sitzend befindet, und gleichzeitig liegt der Zustand vor, in dem sich die andere Person, der Erfinder der Brille, unter dem Baum sitzend befindet.

Ein Sachverhalt ist keine Proposition. Eine Proposition ist wahrheitsfähig, ein Sachverhalt nicht. Eine Proposition ist entweder wahr oder falsch. Ob ein Sachverhalt besteht oder nicht besteht, hängt möglicherweise von dem Geschichtsverlauf und der Situation ab, u.U. vom Zeitpunkt bzw. Zeitintervall, in denen das Bestehen des Sachverhaltes betrachtet wird. Eine Proposition beinhaltet nach meinem Verständnis bereits den Zeitpunkt bzw. Zeitraum, in dem sie als wahr oder falsch bewertet wird, d.h. sie ist zeitinvariant und innerhalb eines Geschichtsverlaufs wahr oder falsch. Ihr korrespondiert ein Sachverhalt, der nicht-transient ist.

Der Satz ‘die letzten olympischen Sommerspiele fanden 2000 in Sydney statt’ drückt in der heutigen Situation im heutigen Geschichtsverlauf eine wahre Proposition aus. Der Sachverhalt, der der Proposition korrespondiert, kann wiederum als eine

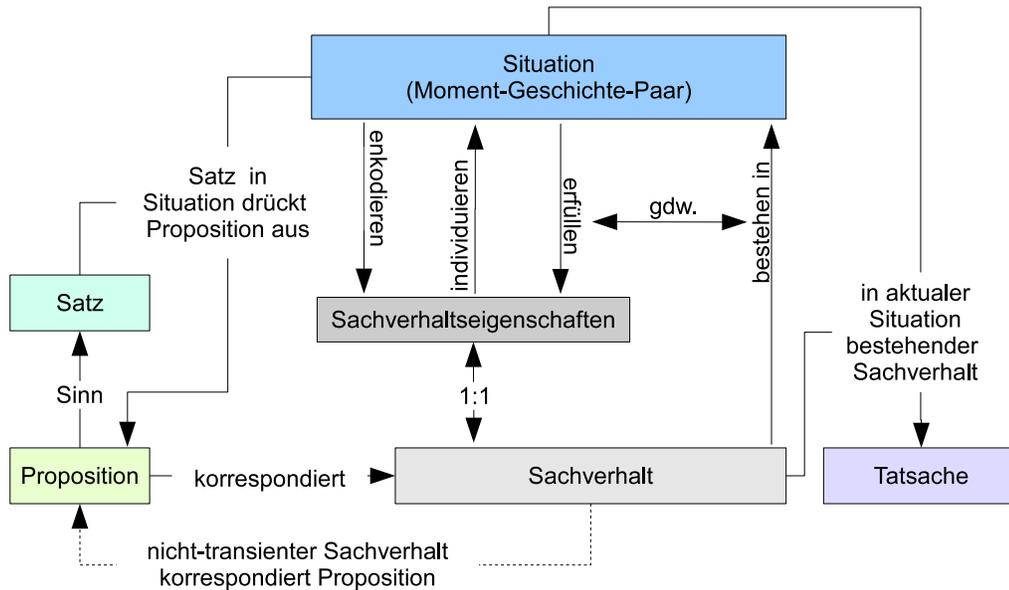


Abbildung 2.1: Sachverhalt und Situation

Zusammensetzung mehrere Sachverhalte angesehen werden, 'einen Zeitpunkt 2000 hat es gegeben', 'es hat Sommerspiele gegeben' etc. Die einzelnen Teilsachverhalte des Sachverhaltes sind vielfältig. Entfernt man einen Teilsachverhalt, erhält man möglicherweise wiederum einen Sachverhalt. Bei Propositionen ist dies nicht der Fall. Entfernt man von der Proposition, die durch den obigen Satz ausgedrückt wird, den Zeitpunkt 2000, so ist die Proposition nicht mehr wahrheitsfähig bzw. ist ihr kein eindeutiger Wahrheitswert (wahr oder falsch) mehr zuordbar, d.h. es ist keine Proposition mehr. Entfernt man vom Sachverhalt 'dass die letzten olympischen Sommerspiele 2000 in Sydney stattfanden' das Objekt '2000', so erhält man hingegen wiederum einen Sachverhalt, der jedoch nicht mehr zeitinvariant ist, da es Situationen gibt, in denen er besteht, und Situationen innerhalb desselben Geschichtsverlaufs, in denen er nicht besteht. Der erhaltene Teilsachverhalt ist situationsabhängig (transient).

Der Satz 'die letzten olympischen Sommerspiele fanden in Sydney statt' drückt dennoch eine Proposition aus, weil der Äußerungskontext des Satzes den Bewertungszeitpunkt impliziert. In diesem Fall beinhaltet er sowohl den Geschichtsverlauf als auch den Zeitpunkt der Sommerspiele. Wurde der Satz am 9.4.2003 im heutigen Geschichtsverlauf geäußert, dann bezieht sich 'die letzten olympischen Sommerspiele' auf das konkrete Ereignis der Sommerspiele 2000 in Sydney.

Eine Proposition ist wahr, wenn der Sachverhalt, der ihr korrespondiert, in dem Ge-

2. Zustände, Sachverhalte und Ereignisse

schichtsverlauf und dem Zeitpunkt besteht. Das Verhältnis von Sachverhalt, Proposition, Situation und Bewertung lässt sich in der aussagenlogischen Modallogik wiederfinden, in der Sachverhalte, Propositionen, Welten und Wahrheitswerte dargestellt werden.

In diese Arbeit symbolisieren Formeln Sachverhalte, die durchaus transient sein können. Eine Formel wird in einem Modell in einer Situation bzw. Welt bewertet. Wird eine Formel in einer Situation als wahr bewertet, wird durch das Modell und dessen Bewertungsfunktion, die den atomaren Teilen des Sachverhaltes ihre Referenz zuweist, ein Bestehen des komplexeren Sachverhaltes in der Situation zugeordnet. Es ist somit erst durch die Hinzunahme der Situation/Welt und des Modells möglich, dem durch den Formelausdruck symbolisierten Sachverhalt eine Proposition und somit einen Wahrheitswert zuzuordnen. Es ergibt sich folgendes Bild in der Stit-Theorie: $\mathcal{M}, (m, h) \models \varphi$ symbolisiert eine Proposition. Der Sachverhalt, der durch φ symbolisiert wird, besteht in der Situation (m, h) im Geschichtsverlauf h . Sei φ eine negierte atomare Formel p , dann ist das Bestehen von φ ($\neg p$) weiter durch die Situation zu spezifizieren. Dem Modell \mathcal{M} , welches weitere Informationen über die die Situation (m, h) charakterisierenden SVH-Eigenschaften besitzt, kann das Bestehen bzw. Nichtbestehen von p entnommen werden, so dass der Aussage $\mathcal{M}, (m, h) \models p$ eine Proposition korrespondiert, die verneint werden kann. Dann wird ersichtlich, ob der Sachverhalt φ in der Situation im Geschichtsverlauf (m, h) besteht.

Jedoch korrespondiert nicht jedem Sachverhalt eine Proposition. Transienten Sachverhalten, die situationsabhängig sind, korrespondiert keine Proposition, da ihr Bestehen an das Bestehen anderer Sachverhalte geknüpft ist. Beispielsweise der Sachverhalt, dass die Sonne scheint, ist ein Sachverhalt der situationsabhängig ist. Die Spezifikation, die bei einem Satz über den Äußerungs- und Bewertungskontext erfolgt, ist bei einem Sachverhalt die Situation. Ein Sachverhalt kann 'x ist grün' sein. Dieser Sachverhalt ist situationsabhängig und kann erst ein Korrespondent einer Proposition sein, wenn 'x' spezifiziert ist, d.h. 'x' auf ein festes Objekt in der spezifizierten Situation referiert, das dann grün oder nicht grün ist.

Sei es der Fall, dass es zu einem Zeitpunkt in einer Situation keine grünen Gegenstände gibt. Der Sachverhalt 'dass es keine grünen Gegenstände zu dem gegebenen Zeitpunkt gibt' besteht. Dann besteht der Sachverhalt 'x ist grün' nicht, denn sein Objekt 'x' kann nicht spezifiziert werden, da kein solches abstraktes oder nicht-abstraktes Objekt existiert. Dies folgt aus der Forderung, dass in einer Situation eine maximale und mögliche Menge an SVH-Eigenschaften instanziiert ist. Das Bestehen des Sachverhaltes 'x ist grün' wird von anderen genannten Sachverhalten präkludiert.

Der Sachverhalt 'x ist nicht grün' besteht ebensowenig, da 'x' nicht spezifiziert werden kann.

Gibt es jedoch grüne Objekte zu dem Zeitpunkt, d.h. der Sachverhalt 'dass es grüne Objekte gibt' besteht in der Situation, ist das Bestehen des Sachverhaltes 'x ist grün' an das Bestehen anderer Sachverhalte gekoppelt. Ist 'x' durch ein Objekt in dieser Situation spezifiziert und bestehen zwei weitere Sachverhalte 'dass dieses x zu dem Zeitpunkt rot ist' und 'dass alles, was rot ist, nicht grün ist', so präkludieren diese Sachverhalte den Sachverhalt 'x ist grün', obwohl 'x' durch ein Objekt in der Situation spezifiziert werden kann. Dabei wird der Sachverhalt 'es gibt etwas, das grün ist' nicht präkludiert. Demzufolge bedeutet Präklusion nicht automatisch die Instanziierung der Negation des Sachverhaltes. Denn 'x ist nicht grün' ist im zweiten Beispiel instanziiert im ersten Beispiel jedoch nicht notwendigerweise.

Ist die Negation einer Proposition definiert, ist dies bei einem Sachverhalt nicht der Fall. Die Proposition, dass der Sachverhalt besteht, kann negiert werden, ohne dass dies impliziert, dass es eine Negation eines Sachverhaltes gibt, die besteht. Es ist möglich, dass der Sachverhalt 'x ist grün' nicht besteht und der Sachverhalt 'x ist nicht grün' ebensowenig besteht. Dies kann natürlich nur der Fall sein, solange keine Proposition einem Sachverhalt korrespondiert. Falls der Sachverhalt 'x ist grün' einer Proposition korrespondiert, referiert 'x' auf ein festes Objekt in der Situation und dieses müsste dann entweder die Eigenschaft 'grün zu sein' oder 'nicht grün zu sein' instanziiieren. Ein Sachverhalt kann somit nur negiert werden, wenn ihm eine Proposition zugeordnet wird. Das bedeutet, dass es eine Situation geben muss, in der der Sachverhalt besteht bzw. nicht besteht, so dass die Proposition, die dem Bestehen des Sachverhaltes in dieser Situation zugeordnet wird, negiert wird. Eine Art von Negation eines beliebigen Sachverhaltes ist somit im Allgemeinen erst möglich, wenn eine Situation gegeben ist.

Ein analoges Beispiel, wie transienter Sachverhalt und Situation sich zueinander verhalten, ist somit, wie sich ein indexikalischer Satz und der Äußerungskontext zueinander verhalten: 'Dies ist grün'. Der Satz drückt nur dann eine Proposition aus, wenn das Ding, worauf 'dies' verweist, Teil der Proposition ist.¹⁷ Wird der Kontext des indexikalischen Ausdruckes offengelassen, drückt er keine Proposition mehr aus. Ähnlich verhält es sich mit situationsabhängigen Sachverhalten und Situation. Wird ein situationsabhängiger (transienter) Sachverhalt ohne Situation bewertet, ist

¹⁷Die Referenz eines indexikalischen Ausdruckes ist kontextabhängig (context-sensitive). Sie haben eine Funktion, die zu jedem Kontext das referierte Objekt liefert (in Kaplans Terminologie wäre dies der „Character“ [88, p.505f]).

2. Zustände, Sachverhalte und Ereignisse

sein Bestehen bzw. Nichtbestehen nicht definiert. Der Sachverhalt existiert jedoch unabhängig davon. Er korrespondiert jedoch keiner Proposition. Er besitzt immer noch eine Art Funktion, die, wenn die Situation fest ist, besagt, ob der Sachverhalt in dieser Situation besteht oder nicht besteht. Somit kann einem transienten Sachverhalt nur in einer Situation eine Proposition korrespondieren. Durch die Fixierung der Situation erhält man einen nicht-transienten Sachverhalt.

Ein nicht-transienter Sachverhalt wäre dann der Sachverhalt, dass dieser situationsabhängige Sachverhalt in dem Geschichtsverlauf in der Situation besteht bzw. nicht besteht. Somit ist das Bestehen eines situationsabhängigen Sachverhaltes wiederum an das Bestehen andere Sachverhalte geknüpft. Es ist dann hinreichend, wenn Plantinga sagt, dass ein Geschichtsverlauf (Welt) durch nicht-transiente Sachverhalte beschrieben wird [107].

Die Trennung von Existenz und Bestehen von Sachverhalten macht es möglich, die Identität von Sachverhalten nicht über ihr Bestehen innerhalb von Geschichtsverläufen zu definieren, sondern über die Enkodierung von SVH-Eigenschaften durch Situationen. Für eine Übersicht der Begriffe und Relationen vergleiche Abbildung 2.1. Zwei Sachverhalte sind identisch, wenn die Situationen, die die zugehörigen SVH-Eigenschaften enkodieren (ohne sie notwendigerweise zu instanziierten, d.h. ohne dass die Sachverhalte in den Situationen bestehen müssen) identisch sind.

Ein Sachverhalt ist nicht auf einen Zeitpunkt festgelegt. Häufig ist ein Zeitintervall vonnöten, um das Bestehen des Sachverhaltes zu gewährleisten. ‘Die Sommerspiele fanden in Sydney statt’ ist ein transienter Sachverhalt, der ein bestimmtes Ereignis benötigt, um als bestehend in einer Situation zu gelten. Der mit dem Satz (und damit fixiertem Äußerungs- und Bewertungskontext) ‘es gastieren am 21.9.2000 die Sommerspiele in Sydney’ ausgedrückten Proposition korrespondierende Sachverhalt ist nicht hinreichend, um das Bestehen des Sachverhaltes zu gewährleisten, der mit der Proposition korrespondiert, die durch den Satz ‘die Sommerspiele fanden in Sydney statt’ ausgedrückt wird. Letzterer Sachverhalt charakterisiert einen Zustand während des Ereignisses der Sommerspiele. Erst genannter Sachverhalt charakterisiert einen Zustand nach dem Ereignis.

Eine *Tatsache* ist in dem aus heutiger Sicht aktuellen Geschichtsverlauf der bestehende Sachverhalt, der der Proposition korrespondiert, die in dem Satz ‘die Sommerspiele 2000 fanden in Sydney statt’ ausgedrückt wird. Diese Proposition ist wahr in dem aktuellen Geschichtsverlauf und wird für immer wahr sein. Ein in dem Geschichtsverlauf bestehender Sachverhalt, der, wenn er einmal besteht, zu jedem späteren Zeitpunkt besteht, ist eine Tatsache. Insbesondere ist eine Tatsache somit zeitin-

2.1. Zustände und Sachverhalte

variant, da sie die Situation beinhaltet, in der sie bestand. Genauer gesagt, sei der Geschichtsverlauf und Zeitpunkt durch eine aktuelle Situation festgesetzt, so ist jeder in diesem Geschichtsverlauf bis zu dem betrachteten Zeitpunkt bestehende nicht-transiente Sachverhalt eine Tatsache. Der Sachverhalt 'die Sommerspiele 2000 fanden in Sydney statt' wird dann immer bestehen. Wohingegen die Sachverhalte 'die Sommerspiele fanden in Sydney statt' (nicht als Satz aufgefasst, somit ohne Bewertungskontext) bestehen oder nicht bestehen kann, je nachdem was das Objekt 'die Sommerspiele' zu dem jeweiligen Sachverhalt beiträgt.

Ein Sachverhalt kann somit charakteristisch sowohl für ein Ereignis als auch für einen Zustand sein. Umgekehrt können konkrete Ereignisse und Zustände durch eine Menge von bestehenden Sachverhalten charakterisiert werden. Der Sachverhalt 'dass die Sonne scheint' kann für ein Ereignis 'Sonnenschein', wenn das Bestehen des Sachverhaltes über einen bestimmten Zeitraum betrachtet wird, charakteristisch sein. Wird nur ein Zeitpunkt betrachtet, in diesem Moment scheint die Sonne, so trägt der Sachverhalt auch als Charakteristikum des Zustandes bei, dass die Sonne in diesem Moment scheint.

Um im Folgenden weniger umständliche und leichter verständliche Sätze zu konstruieren, erlaube ich mir folgende Abkürzungen. Da jeder Proposition ein Sachverhalt korrespondiert und Propositionen Sinngehalte von Sätzen wiedergeben, werde ich formulieren, dass auch Propositionen bzw. Sätze Zustände bzw. Ereignisse charakterisieren. So verwende ich kurz, der Sachverhalt 'es gastieren die Sommerspiele in Sydney'(Satz) bzw. die Proposition 'es gastieren die Sommerspiele in Sydney', wobei mit Sachverhalt derjenige gemeint ist, der der Proposition korrespondiert, die durch den Satz 'es gastieren die Sommerspiele in Sydney' ausgedrückt wird. Die vorangehenden zwei Sätzen lauten dann: Der Zustand, der durch (den Sachverhalt) 'es gastieren die Sommerspiele in Sydney' charakterisiert wird, garantiert nicht das Bestehen des Sachverhaltes 'die Sommerspiele fanden in Sydney statt'.

Es gibt sicherlich Einwände derart, dass, wenn es Sachverhalte gibt, die keiner Proposition korrespondieren, es Zustände (Ereignisse) geben könne, die nur durch Sachverhalte charakterisiert werden, denen keine Proposition korrespondiert. Dies stellt insoweit kein Problem dar, als dass es die Begriffe des Zustandes und Sachverhaltes, so wie ich sie gebrauche, nicht beeinträchtigt. Da für jede Situation gilt, dass sie die SVH-Eigenschaft instanziiert oder nicht, besteht ein Sachverhalt oder er besteht nicht. Besteht er nicht, besteht der (ein anderer) Sachverhalt, dass der Sachverhalt

2. Zustände, Sachverhalte und Ereignisse

in dieser Situation nicht besteht.¹⁸ So ist es möglich, auch wenn einem Sachverhalt keine Proposition korrespondiert, durch die Angabe von nicht-transienten Sachverhalten, das Bestehen der den Zustand charakterisierenden Sachverhalte durch andere nicht-transiente Sachverhalte zu prä- oder inkludieren. Der Sachverhalt, dass ein bestimmtes Objekt grün ist, ist sicherlich situationsabhängig, z.B. weil das Objekt seine Farbe von Situation zu Situation wechselt oder weil andere Sachverhalte ausschließen, dass es grün ist. D.h. auch wenn ein Zustand durch eine Menge situationsabhängiger Sachverhalte charakterisiert werden kann, gibt es zwei Möglichkeiten, die Sachverhalte durch nicht-transiente Sachverhalte zu ergänzen.

Falls es unmöglich ist (diese Annahme scheint mir unplausibel), die Sachverhalte konsistent durch nicht-transiente Sachverhalte zu ergänzen, deren Bestehen die den Zustand charakterisierenden Sachverhalte inkludieren, würde es sich um einen Zustand handeln, der nicht durch Propositionen charakterisiert werden kann und somit nicht durch Sätze beschreibbar ist. Über Zustände, die nicht charakterisiert werden können und somit auch nicht durch Sätze beschreibbar sind, kann man wenig aussagen und sind uninteressant. Ich kann Wittgenstein nur zustimmen: „Wovon man nicht sprechen kann, darüber muss man schweigen“ [165, Satz 7]. Daher gehe ich davon aus, dass die Zustände, die ich im Folgenden betrachte, z.B. der Zustand nach einer vollzogenen Handlung, die mentalen Zustände wie Wünsche, Überzeugungen und Intentionen, propositional erfasst werden können. So ist die Unterscheidung zwischen Sachverhalt und Proposition nicht von entscheidender Bedeutung.

Desgleichen trifft im Weiteren auch auf Ereignisse und ihre Charakterisierung durch Sachverhalte zu. Wenn offensichtlich ist, dass der Satz S eine Zustands- bzw. Ereignisbeschreibung ist, dann charakterisiert die Proposition φ bzw. der der Proposition korrespondierende Sachverhalt den Zustand bzw. das Ereignis. Dies vorausgesetzt, spreche ich im Folgenden auch von dem Zustand bzw. dem Ereignis, dass φ . Ereignisse sind ebenfalls wie Zustände nur dann existent, wenn sie eintreten. Nichtsdestotrotz gibt es Zustands- und Ereignistypen, die immer wieder vorliegen bzw. eintreten. Eine feste Menge von möglicherweise auch transienten Sachverhalten, die zu wiederkehrenden Zeitpunkten immer wieder vorliegen können, charakterisieren dann einen solchen Zustandstyp. Entsprechend kann eine Menge von transienten Sachverhalten, deren Bestehen sich über ein Zeitintervall erstreckt und deren SVH-Eigenschaften innerhalb dieses Intervalles instantiiert werden können, einen Ereignistyp kennzeichnen.

¹⁸Es ist nicht notwendig, dass eine Negation des Sachverhaltes geben muss.

2.2 Ein Zustand und das Vorliegen des Zustandes

In diesem Abschnitt mache ich auf einen Punkt aufmerksam, der den Unterschied zwischen einem Ereignis und einem Zustand verdeutlicht. Die Frage ist, wenn ein Zustand in einem Moment besteht, wie kann er dann über einen längeren Zeitraum andauern? Die Antwort, die ich gebe, ist, dass nicht der Zustand andauert, sondern ein Ereignis stattfindet, das durch den Sachverhalt charakterisiert wird, dass der Zustand zu jedem Zeitpunkt des Ereignisses vorliegt. Ich werde somit zwischen einem Zustand in einem Moment und dem Vorliegen des Zustandes unterscheiden.

Die Zustände, die durch die Sachverhalte ‘die Sommerspiele gastieren in Sydney’ und ‘die Sommerspiele fanden in Sydney statt’ charakterisiert werden, lassen die Frage nach dem Bestehen der Zustände offen. Ersterer lag sicherlich während der Sommerspiele letzterer danach vor. Der Zeitraum in dem ersterer vorlag, war sicherlich der Zeitraum, in dem die Sommerspiele stattfanden. Die olympischen Sommerspiele waren insgesamt jedoch ein Ereignis kein Zustand. Jedoch ist auch das Vorliegen des Zustandes, das die Sommerspiele 2000 in Sydney gastieren, nicht identisch mit dem Ereignis der Sommerspiele 2000 in Sydney. Demzufolge scheinen wir, über drei verschiedene Entitäten zu sprechen, dem Zustand, dem Vorliegen des Zustandes und dem Ereignis, währenddessen der Zustand unter anderem vorlag.

Dass es notwendig ist, zwischen dem Zustand und dem Vorliegen des Zustandes zu unterscheiden, macht z.B. auch Löbner [96] unfreiwillig deutlich.¹⁹ So vertritt Löbner, dass Zustände verschieden von Ereignissen sind. Während Zustände sich auf Zeiten beziehen (Zeiten bei Löbner sind sowohl Zeitpunkte als auch Zeitintervalle) und Propositionen zugeordnet werden, sind Ereignisse abstrakte Individuen, denen Zeiträume, in denen sie sich ereignen, zugeordnet werden.

Die fehlende Unterscheidung bei Löbner²⁰ zwischen einem Zustand zu einer Zeit und dem Vorliegen eines Zustandes über eine Phase und sein Vorschlag diese Phase zudem, als ein Ereignis zu charakterisieren, haben kontraintuitive Konsequenzen.

¹⁹Eine ganze Reihe von Ansätzen, die sich auf Vendler [151] beziehen, existieren in der Literatur und unterscheiden nicht zwischen dem Anhalten eines Zustandes und dem Zustand, ohne dass dieselben Konsequenzen wie bei Löbner auftreten. Sie sehen einen Zustand nicht als punktuell an, sondern immer als Ereignis über eine Zeitspanne hinweg. Sie bezeichnen als Zustand das Bestehen eines Sachverhaltes über einen gewissen Zeitraum. Die Unterscheidung, die für mich den grundlegenden Unterschied zwischen einem Ereignis und einem Zustand ausmacht, etwas ist punktuell und etwas, das über einem bestimmten Zeitraum auftritt, wird dabei nicht getroffen.

²⁰Löbner wollte in diesem Aufsatz keine Begriffsanalyse für Zustand und Ereignis erstellen, sondern analysiert anhand seiner Definitionen „Tempora“ und „Aspekte“ von Ausdrücken.

2. Zustände, Sachverhalte und Ereignisse

Wiewohl Löbner meint, dass Zuständen bestimmte Ereignisse zugeordnet werden können, als auch selbst welche charakterisieren, ist bei ihm ein Zustand die Menge von Zeiten, in denen der Zustand vorliegt. Allein diese Definition ist zirkulär. Löbner stellt weiter fest, dass für jeden Zustand ein Ereignistyp angegeben werden kann, der in dem Auftreten einer Phase des Zustandes besteht [96, S.176]. Weiterhin kann ein Zustand, da durch eine Proposition gegeben, wahr oder falsch sein und sogar negiert werden (im Sinne von: Der Zustand liegt nicht vor). In [96] auf S.169f wird das Problem deutlich.

Er schließt: Im Satz ‘als meine Wohnung ausgeraubt wurde, war ich in Stuttgart’ drückt der Hauptsatz einen Zustand und der Nebensatz ein Ereignis aus. Demzufolge legt das Zeitintervall des Ereignisses fest, dass der Zustand während der Ereigniszeit bestand, also das Ereignis zeitlich in die Phase des Zustandes fiel. Und nur dann kann der Satz entweder wahr oder falsch sein. Die problematische Folge wird bei einem Satz der folgenden Form deutlich. Angenommen Hans war von Montag bis Mittwoch in Köln. Nach Löbner ist der Satz ‘von Dienstag bis Freitag war Hans in Köln’ weder wahr noch falsch, da die Phase des Zustandes von Montag bis Mittwoch bestand und somit von Dienstag bis Freitag nicht in die Phase des Zustandes fällt.

Dies ist unplausibel. Zum einen ist der Satz offensichtlich falsch, denn ‘von Dienstag bis Freitag in Köln zu sein für Hans’ ist ein anderer Sachverhalt als ‘von Montag bis Mittwoch in Köln zu sein für Hans’, die auch unabhängig voneinander bestehen oder nicht bestehen. Selbst wenn man Löbner recht gäbe und den Satz als weder wahr noch falsch anerkennen würde, widerspricht dies selbst seiner eigenen Theorie. Das Prädikat ‘in Köln zu sein’ ist nach seiner Definition ein Zustandsprädikat über Zeit und Situationen. Demzufolge müsste es in jedem Zeitpunkt bzw. Zeitraum, der in eine positive bzw. negative Phase des Zustandes fällt und auf den das Prädikat angewendet wird, entscheidbar sein, ob die Aussage wahr oder falsch ist. Es ist durchaus plausibel anzunehmen, dass es eine vollständige Zerlegung des Zeitintervalles [Dienstag, Freitag] in Phasen gibt, ‘Hans ist in Köln’ und ‘Hans ist nicht in Köln’, d.h. die von Löbner erwähnte Übergangsphase, in der der Zustand weder vorliegt noch nicht vorliegt, ist nicht vorhanden. Das kann nach Löbner jedoch nicht sein, da für Löbner der Satz ‘von Dienstag bis Freitag war Hans in Köln’ weder wahr noch falsch ist, und somit in diese Übergangsphase fallen müsste. Da jedoch diese Übergangsphase nicht existiert, muss der Satz entweder wahr oder falsch sein.

Die Schlussfolgerung, die ich ziehe ist, dass ‘Hans war von dann bis dann in Köln’ nur ein Ereignis aber kein Zustand sein kann. Das Problem besteht bei Löbner, weil es sich erstreckende Phasen des Zustandes gibt, die der Zustand selber sind. Unterscheidet man das Anhalten des Zustandes (Phase) vom Zustand selbst und ist erst genanntes

2.2. Ein Zustand und das Vorliegen des Zustandes

ein Ereignis löst sich das Problem. Dann ist das Ereignis ‘von Dienstag bis Freitag in Köln zu sein’ nicht eingetreten, auch wenn das Ereignis von ‘Montag bis Mittwoch in Köln zu sein’ eingetreten ist, und es möglich ist, dass der Zustand ‘am Dienstag zu einem bestimmten Zeitpunkt in Köln zu sein’ vorlag. Jedoch lässt sich aus dem Vorliegen des Zustandes zu einem Zeitpunkt keine Aussage darüber ableiten, wie lange dieser Zustand insgesamt vorlag. In keinem Fall ist dieser Zustand dasselbe wie das Ereignis ‘Vorliegen des Zustandes’.

Es gibt Ansätze, die festhalten, dass das Anhalten des Zustandes und der Zustand gleichzusetzen sind, aber beides keine Ereignisse sind. Dies ändert nichts an dem Problem, den Wahrheitswert eines Satzes S_1 ‘Z liegt in t_3 vor’ mit $t_3 \in (t_1, t_2)$ und den Wahrheitswert von S_2 ‘Z liegt von (t_1, t_2) vor’ entweder miteinander zu identifizieren, was dazu führt, dass S_2 wahr und falsch sein kann, oder so wie Löbner zu behaupten, dass solche Sätze weder wahr noch falsch sind, was ebenso unplausibel ist.

Desgleichen gilt, wenn gefragt wird, ob S_3 ‘von Dienstag bis Freitag war Hans in Köln’ und S_4 ‘von Mittwoch bis Donnerstag war Hans in Köln’ wahr bzw. falsch sind. Wenn S_3 eine Zustandaussage und wahr ist, dann ist nicht- S_3 falsch nach Löbner. Die Aussage S_4 müsste, da die in S_4 betrachtete Phase in die Phase des Zustandes von S_3 fällt, wahr sein. Wenn S_4 eine Zustandaussage ist, muss aber die Negation des Satzes demzufolge falsch sein, siehe [96]. Dies ist er nicht, da die Schnittmengen der Phase mit S_3 und nicht- S_3 beide nicht leer sind. Zu schließen, dass S_4 demzufolge keinen Zustand beschreibt S_3 aber sehr wohl, ist wiederum unplausibel.

Bei der Identifizierung des Zustandes mit der Phase tritt ein Problem auf, wenn Teilintervalle der Phase betrachtet werden. Wenn ein Zustand über ein Intervall hinweg anhält und der Zustand mit der Phase des Anhalten identifiziert wird, dann scheint es dem Begriff inhärent zu sein, dass im jeden Teilintervall der Phase derselbe Zustand wie in der gesamten Phase vorliegt. Einen Teilzustand des Zustandes von dem Zustand zu unterscheiden, ist begrifflich unsinnig. Genauso unsinnig ist jedoch die Phase als identisch mit der Teilphase zu betrachten. Korrekter wäre zu sagen, dass ein und derselbe Zustandstyp in jedem Zeitpunkt $t \in (t_0, t_1)$ instanziiert wird. Dieses Anhalten des Zustandes ist eine andere Entität, die nicht der Zustand selber ist. Dann würden die Sätze S_3 und S_4 zwei verschiedene Ereignisse desselben Ereignistyps beschreiben und eindeutige Wahrheitswerte besitzen.

Vendler et al. vertreten, dass jeder Zustand ein Ereignis ‘state’ ist, bei dem sich nichts ändert und ein Verb andeutet, welcher Sachverhalt konstant beibehalten wird.²¹

²¹Auf die Klassifikationen, die Vendler vornimmt, gehe ich in Abschnitt 3.2 ausführlich ein.

2. Zustände, Sachverhalte und Ereignisse

Mit einem Ereignis ‘state’ identifiziert Vendler das, was ich als das Vorliegen des Zustandes oder als die Phase des Zustandes bezeichnen würde. Auf ihn trifft somit die Gleichsetzung von einem Ereignis wird als ‘state’ (als Phase des Zustandes) bezeichnet und ‘ist statisch’ zu. Dieser Gleichsetzung würde ich zustimmen, da ein Ereignis die Definition von ‘statisch’ erfüllt, das durch das Vorliegen eines Zustandes beschrieben werden kann, der sich in diesem Zeitraum nicht ändert. Ich werde also im Weiteren zwischen ‘state(s)’ im Vendlerschen Sinne, d.h. ein Ereignis, dass ein Zustand vorliegt, und ‘Zustand’ unterscheiden.²²

2.3 Zustände und Ereignisse

Nachdem ich charakterisiert habe, was einen Zustand ausmacht und welche Rolle Sachverhalte und Propositionen in der Beschreibung von Zuständen spielen, werde ich nun den Begriff des Ereignisses näher betrachten. Dieser ist deswegen von Bedeutung, da ich im folgenden Kapitel die These unterstützen möchte, dass es sich bei den Handlungen eines Akteurs um Ereignisse handelt. Demzufolge stehen in diesem Abschnitt folgende Fragen zur Diskussion: Was ist ein Ereignis? Und wie kann es charakterisiert und beschrieben werden?

Da ich argumentiert habe, dass für den Begriff ‘Zustand’ eine Attributierung mit statisch oder dynamisch nicht sinnvoll ist, scheint der nächstgelegene Begriff, dem solche Attribute zukommen, der des ‘Ereignisses’ zu sein. Die Unterscheidung zwischen Zuständen und Ereignissen ist offensichtlich notwendig und in der Linguistik, wenn auch in unterschiedlichen Auffassungen, üblich. Davidson hält fest, dass es Verben gibt, die eine Ereignisstelle besitzen, und solche, die keine Stelle für eine Ereignisvariable besitzen, [48, Essay 6]. In [89] unterscheidet Katz zwischen „stative verbs“ und „event verbs“. In [99] findet sich die Unterscheidung zwischen „event main clause“ und „stative main clause“. Löbner [96] argumentiert ebenfalls dafür. Er stellt die statische Theorie, die Sätze als Zustandsaussagen liest, und die dynamische Theorie gegenüber. Perfektive Aussagen beschreiben Ereignisse. Imperfektive Aussagen beschreiben Zustände oder Aussagen, die einen progressiven oder iterativen Aspekt besitzen [96, S.165]. Bei Löbner fallen auch die in Vendlers Terminologie als ‘Activities’ bezeichneten andauernden ‘Aktivitäten’, wie z.B. joggen, unter Zustände. Er

²²Korrekterweise müsste ich statt dem Vorliegen des Zustandes über einem Zeitintervall von dem Instanzieren ein und desselben Zustandstyps in jedem Zeitpunkt des Intervalles durch konkrete Zustände sprechen, die sich nur im Zeitpunkt unterscheiden. Kurz werde ich jedoch im Weiteren nur noch vom Vorliegen des Zustandes sprechen.

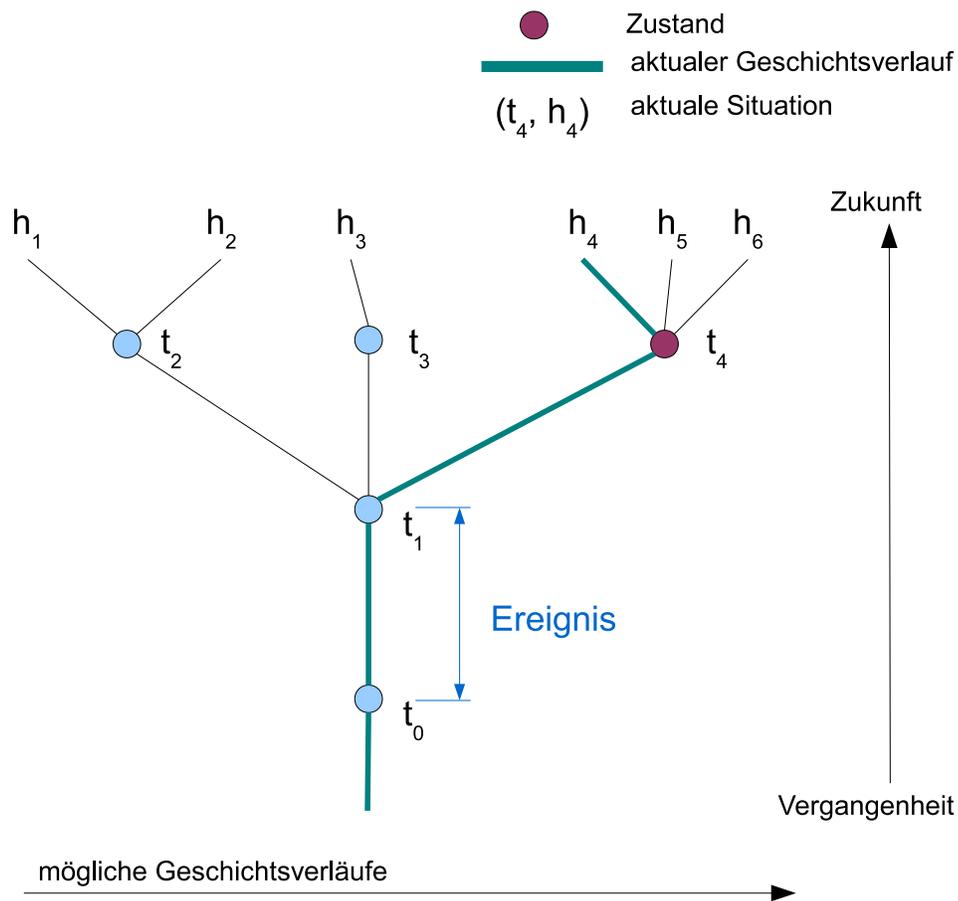


Abbildung 2.2: Zustand und Ereignis als konkrete Zeitpunkte bzw. Zeitintervalle

unterscheidet dann zwischen ‘dynamischen’ und ‘statischen’ Zuständen. Dies ist meines Erachtens eine viel zu weit gefasste Definition des Begriffes ‘Zustand’.

Eine meines Erachtens zutreffende Charakterisierung von Zuständen und Ereignissen findet sich bei Parson:

„Events and Processes (as opposed to States) would *take time*, whereas States (as opposed to Events and Processes) would *last* through time.“
[103, S.37]

Hieran wird meines Erachtens deutlich, dass einem Ereignis der Zeitraum, den es einnimmt, essentiell ist, wohingegen bei einem Zustand das Anhalten des Zustandes nicht an den Zeitraum, über den der Zustand besteht, gebunden ist. Es ist für den Zustand nur essentiell, in dem Zeitpunkt vorzuliegen, zu dem er betrachtet wird.

2. Zustände, Sachverhalte und Ereignisse

Meiner Intuition nach wird der Unterschied beider Begriffe durch das jeweilige Verhältnis zur Zeit deutlich.

„It is commonly believed, at least among philosophers, that there are such entities as ‘times’, that our knowledge of these entities is what enables us to identify and individuate events, and that the concept of *a time* is essential to the concept of *an event*.“ [44, S.413]

Auch wenn Chisholm sich selbst nicht unter diese Philosophen zählt,²³ für die die Komponente Zeit entscheidend bei der Differenzierung und Individuierung von Ereignissen (und Zuständen) ist, zähle ich mich dazu. Da ich unter einem Zustand etwas verstehe, das durch das Bestehen eines Sachverhaltes zu einem festen Zeitpunkt vollständig charakterisiert wird, unabhängig davon wie lange er vorliegt bzw. wodurch er zustande kam, spielt der Zeitraum, in dem er vorliegt, keine Rolle. Dies oder jenes ist im Moment *m* der Fall. Er beinhaltet dabei keine Änderung und kein Geschehen und demzufolge auch keine Handlung. Ob ein Zustand vorliegt, ist, wenn es entscheidbar ist, punktuell entscheidbar. Der Begriff ‘Ereignis’ ist oft diskutiert worden. Eine Übersicht dieser Diskussion würde den Rahmen dieser Arbeit sprengen und ist nicht gewollt. Da ich im Rahmen dieser Arbeit auf der Beschreibungsebene verbleibe, werde ich nur kurz auf den Begriff ‘Ereignis’ eingehen, den ich im Weiteren zugrunde legen werde.

Ein *Ereignis* ist etwas, das der althochdeutschen Bedeutung von ‘ir ougen’ folgend besagt, dass etwas vor die Augen tritt bzw. offenbar wird, vgl. [70, S.45f]. Dies bedeutet, dass es ein *Vor*, als es sich noch nicht ereignete, und ein *Danach*, als es eingetreten war, geben muss. Mit einem Ereignis ist somit immer eine Zeitspanne assoziiert, in der das Ereignis eintritt, siehe z.B. auch Löbner [96, S.167] oder von Wright [154, S.28f].²⁴ Ereignisse scheinen daher auf den ersten Blick vom Typ her dynamisch zu

²³Chisholm erklärt in seinem späteren Aufsatz Ereignisse als eine bestimmte Art von „states“, die keine Zeitkomponente haben und daher auch nicht durch sie individuiert werden, sondern durch ihre ‘Substrate’ und die Eigenschaften, die diese haben. Das Substrat muss ein kontingentes Ding und der Inhalt, der besagt, dass das Substrat die Eigenschaft hat, darf ebenfalls nicht notwendigerweise erfüllt sein, [44]. Demzufolge ist es nach Chisholm nicht notwendig, die Zeit als seiend oder nicht seiend zu betrachten, da Ereignisse unabhängig von der Ontologie der Zeit sind. In ‘Person and Object’ [43] fasst er allein Sachverhalte als grundlegende Kategorie auf und beschreibt Propositionen, Ereignisse und Fakten als jeweils eine Art von Sachverhalt. Dies trifft nicht meine Einteilung in die verschiedenen Kategorien. Jedoch charakterisieren bei mir die zu bestimmten Zeitpunkten bestehenden bzw. nicht-bestehenden Sachverhalte Propositionen, Zustände und Ereignisse.

²⁴Moens und Steedman, die ich im Weiteren noch häufiger zitieren werde, formulieren: „All of the abo-

Ereignis	charakterisiert durch	eine Menge von Sachverhalten, die in den Situationen eines konkreten Zeitintervalles des aktualen Geschichtsverlaufs bestehen,
Zustand	charakterisiert durch	eine Menge von Sachverhalten, die in einer konkreten Situation des aktualen Geschichtsverlaufs bestehen.
Ereignistyp	charakterisiert durch	eine Menge von instanziierten Sachverhalten, die in einem möglicherweise mehreren Zeitintervallen innerhalb eines Geschichtsverlaufs instanziiert werden können.
Zustandstyp	charakterisiert durch	eine Menge von Sachverhalten, die in einer oder mehreren Situationen instanziiert sind.

Tabelle 2.1: Zustand und Ereignis - Sachverhalte in Situationen und Geschichten

sein.²⁵ Meines Erachtens begeht man jedoch einen Fehler, wenn man aufgrund der Betrachtung eines Zeitraumes schließt, dass auch notwendigerweise eine Änderung innerhalb dessen stattgefunden haben muss. Das heißt, dass nicht notwendigerweise eine Änderung eines der auftretenden Zustände in dem betrachteten Zeitintervall vor sich gehen muss, um als ein Ereignis zu zählen. Das Anhalten eines Zustandes über einem bestimmten Zeitraum ist ebenso ein Ereignis.

Wenn definiert wird, wann etwas als ein Ereignis betrachtet wird, sind notwendige Kriterien, dass es durch einen Zeitraum und eine Menge von Sachverhalten charakterisiert wird, die zu den jeweiligen Momenten des Zeitraumes bestehen. Hierin liegt der Unterschied zu einem Zustand, der ebenfalls durch eine Menge von Sachverhalten und einem Zeitpunkt charakterisiert wird. Scheinbar wiederkehrende Ereignisse sind meines Erachtens Ereignisse ein und desselben Typs. Man kann zwischen einem Ereignistyp und einem Ereignis unterscheiden. Eine Ereignisbezeichnung ‘Mord’ assoziiert eine unbestimmte Zeitspanne, den gewaltsamen Tod eines Menschen und dass es

ve categories describe what common sense suggests we call events—that is, happenings with defined beginnings and ends“ und weiter „we distinguish these ‘hard-edged’ categories from a class of indefinitely extending states of affairs, which, equally commonsensically, we call states“ [99, S.17]. Bis auf die Tatsache, dass ich ein ‘unbestimmtes ausgedehntes Anhalten eines Sachverhaltes’ nicht Zustand nennen würde, pflichte ich ihnen bei.

²⁵Es wird vertreten, dass ein Ereignis nur dann vorliegt, wenn eine Veränderung stattfindet. Z.B. charakterisieren der Zustand ‘Davor’ und der Zustand ‘Danach’ bei Von Wright das Ereignis [154, S.36f], das er als *transition* oder *change* (von Sachverhalt oder Prozess zu Sachverhalt oder Prozess) beschreibt.

2. Zustände, Sachverhalte und Ereignisse

mindestens einen Täter und ein Opfer sowie mögliche andere Sachverhalte gibt. Ohne Inanziierung der Sachverhalte ist diese Ereignisbezeichnung eine Bezeichnung für einen Ereignistyp, der durch eine Menge von Sachverhalten charakterisiert wird. Aber erst das eingetretene konkrete Ereignis der ‘Mord an Caesar’ legt die Zeitspanne und Zeitpunkte fest, in denen die zum Mord an Caesar gehörenden Sachverhalte bestehen.

Nach meiner Auffassung, die ebenfalls von Moens und Steedman [99] u.v.m. vertreten wird, ist die Annahme, dass Ereignisse Subereignisse haben können, unproblematisch. Werden Ereignisse durch Sachverhalte charakterisiert, die innerhalb des Zeitraumes des Ereignisses bestehen, so können Teilereignisse durch das Bestehen von Teilmengen der Sachverhalte und möglicherweise weiteren Sachverhalten charakterisiert werden. Mit dem ‘Mord an Caesar’ wird im heutigen Geschichtsverlauf ein Ereignis bezeichnet, das bspw. das Teilereignis beinhaltet, das durch ‘Brutus ersticht Caesar’ charakterisiert wird.

Diese Ereignisse sind nicht identisch. Sie können durch andere Mengen von Sachverhalten charakterisiert werden. Aber ein Teil des Mordes an Caesar im heutigen Geschichtsverlauf ist die Tatsache, dass Brutus Caesar erstach. Teilereignisse werden jedoch nicht nur durch die Angabe eines Teilzeitintervalles individuiert, da sie auch gleichzeitig eintreten können, selbst wenn sie von ihrem Zeitintervall her in das Intervall des Gesamt ereignisses fallen. Brutus stach zu, während Cassius Caesar festhielt. Dies sind möglicherweise zeitgleiche Teilereignisse der Ermordung Caesars. Die Teilereignisse kann man auch nicht pauschal anhand von Propositionen individuiieren. Immerhin können verschiedene Propositionen ‘Longinus hielt Caesar’ und ‘Cassius hielt Caesar fest’ dasselbe Teilereignis charakterisieren. Ähnliche Beispiele führt Davidson an [48, Essay 9]: ‘Meyer climbed Kibo’ und ‘Meyer climbed the highest mountain in Africa’. Es sind verschiedene Propositionen bzw. Sachverhalte, die jedoch in demselben konkreten Ereignis wahr sind bzw. bestehen. Es ist somit nicht sinnvoll, ein Ereignis durch Aufzählen aller Zustände bzw. aller Sachverhalte zu kennzeichnen, die während dieses Ereignis vorliegen bzw. bestehen.

Meines Erachtens ist es ebenfalls nicht möglich, das Ereignis, welches zwischen m_0 und m_1 eintritt, als Paar von Tatsachen, die zu den Zeitpunkten m_0 und m_1 vorliegen, aufzufassen, siehe [154], da nicht klar ist, inwieweit diese Tatsachen selbst zum Ereignis gehören. Zudem gilt: Wenn Caesar den Arm hebt, ist dies ein Ereignis. Die Tatsache zuvor, Caesar hat den Arm unten, und die Tatsache danach, Caesars Arm ist oben, müsste nach von Wright das Ereignis individuiieren, aber ob er ihn hob oder er gehoben wurde, geht daraus nicht hervor. Dies könnten jedoch zwei verschiedene Ereignisse sein, da der Sachverhalt, der das Ereignis des Armhebens charakterisiert,

im Bestehen den Sachverhalt präkludiert, der ein Ereignis charakterisiert, bei dem der Arm gehoben wurde. So muss ein Ereignis sich durch mehr als nur den Anfangs- und Endzustand des Ereignisses auszeichnen.

Nun könnte jemand auf die Idee kommen, Ereignisse als eine Aneinanderreihung von Zuständen zu allen Zeiten, die zwischen Anfangs- und Endpunkt des Ereignisses liegen zu erachten. Dies ist jedoch in mehrerlei Hinsicht unsinnig. Zum einen wird Zeit von den meisten Philosophen und Physikern als kontinuierlich erachtet, auch wenn Modelle, die Zeit modellieren, häufig annehmen, dass Zeit als diskret dargestellt werden kann. Der Grund liegt meist in einer einfacheren Handhabung des Modells. Nimmt man die Zeit als kontinuierlich an, hätte man eine überabzählbare Menge von Zuständen, die einandergereiht werden müssten, so dass niemand mehr ein Ereignis erfassen könnte, was der ursprünglichen Bedeutung des Wortes 'es offenbart sich' widerspräche. Zum anderen wird die Unsinnigkeit an Paradoxien deutlich, wie z.B. Zenos Pfeilparadox [118]. Das Ereignis 'der Pfeil fliegt von A nach B' wird nicht durch die Menge der Zustände, die der Pfeil zu den verschiedenen Zeitpunkten seines Fluges annimmt, hinreichend erfasst. Wenn die Menge der Zustände zu den Zeitpunkten das Ereignis charakterisiert, würde der Pfeil in jedem Zustand während des Fluges keine Bewegung vollziehen, da eine Bewegung eine Zeitspanne voraussetzt. Erst die Betrachtung über eine Zeitspanne offenbart die Bewegung, also das Ereignis selbst.

Des Weiteren stände die Frage im Raum, welche Zustände, die zwischen diesen zwei Zeitpunkten vorliegen, Teil des Ereignisses sind. Es sind offensichtlich nicht alle, da die Zustände zu dem Zeitpunkt auch den Zustand beinhalten, dass Sand in der Sahara ist, der für das Ereignis des fliegenden Pfeils in Griechenland nicht charakteristisch ist. Die Menge aller Zustände zu den Zeitpunkten des Ereignisses ist somit weder hinreichend noch notwendig, um das Ereignis zu charakterisieren.

Caesar ist an einem Märztag gestorben, die Vögel zwitscherten, nachdem ein Messer in seiner Brust steckte, etc. Es werden mehrere Zustände coexistent beschrieben. Es war März. Es gab Vögel. Diese könnte man als Teilzustände zu dem gesamten Weltzustand zu einem bestimmten Zeitpunkt zusammenfassen. Es coexistieren weitere Ereignisse. Caesar starb. Vögel zwitscherten. Es wurde ein Messer in eine Brust gestochen. Die Ereignisse der Stich des Messers in die Brust und Caesars Tod können unter dem Ereignis, das mit 'Ermordung Caesars' bezeichnet wird, subsumiert werden. Ich bestreite, dass sie mit ihm identisch sind. Sie sind Teile des Ereignisses.

Mein Ansatz ist festzulegen, dass ein Ereignis durch Teilereignisse des Ereignisses abgesteckt werden kann, die wiederum durch Sachverhalte charakterisiert werden. Dies klingt nach einem Regress. Die Ermordung Caesars hat viele Teilereignisse, z.B.

2. Zustände, Sachverhalte und Ereignisse

das Zustechnen des Brutus, die Senatsversammlung, der Tod Caesars. Diese haben aber ebenfalls Teilereignisse, die durch Sachverhalte charakterisiert werden. Das Ereignis, das mit ‘Zustechnen des Brutus’ bezeichnet wird, wird z.B. durch folgende Sachverhalte charakterisiert, ‘Brutus Oberarmmuskeln kontrahiert’, ‘das Messer dringt in die Brust ein’ etc. Um das Ereignis der Ermordung Caesars zu charakterisieren, muss nicht festgehalten werden, welcher Oberarmmuskel von Brutus in diesem Zeitintervall kontrahierte. Der Sachverhalt ‘Brutus trieb das Messer in Caesars Brust’ scheint jedoch charakterisierend für das Ereignis der Ermordung Caesars im heutigen Geschichtsverlauf. Es gibt somit Sachverhalte, die das Teilereignis charakterisieren, die für die Charakterisierung des darüberliegenden Ereignisses nicht relevant bzw. nicht charakteristisch sind.

Es gibt möglicherweise für jedes Ereignis eine von der Anzahl her unbestimmte Menge an Teilereignissen. Aber diese sind nur bis zu einer gewissen Tiefe, im Sinne von Teilereignis des Teilereignisses etc. relevant für das zu individuierende Ereignis. Jedes Ereignis e besteht aus Teilereignissen e_1, \dots, e_n , die für das Ereignis e basal sind, jedoch selbst wieder Teilereignisse besitzen können, die für das jeweilige Ereignis e_i basal und verschieden von e_i sein können aber für die Charakterisierung bzw. Individuierung von e nicht relevant sind. Demzufolge bricht der Regress ab. Eine Benennung bzw. Identifizierung aller essentiellen Teilereignisse muss nicht möglich sein, und ist es in den meisten Fällen auch nicht. Dass die Unterscheidung in essentielle versus kontingente Attribute zu einem Ereignis nicht möglich ist und auch bisher kein Versuch einer Definition einer Individuierung von Ereignissen gefruchtet hat, hält Bennett in [21] fest. Dies ist jedoch kein Grund anzunehmen, dass Teilereignisse unplausibel sind, sondern es zeigt lediglich, dass nicht geklärt werden kann, wo die Grenze zwischen Ereignissen verläuft.

Es gibt jedoch Ansätze, die die Ansicht, dass es sich bei Ereignissen um durch ein Zeitintervall bestimmte Entitäten und bei Zuständen um durch Zeitpunkte bestimmte Entitäten handelt, nicht teilen. Es ist ihrer Ansicht nach nicht notwendig, dass Ereignisse sich innerhalb eines Zeitintervalles ereignen müssen, z.B. Bennett [21]. Diese Autoren betrachten bestimmte Ereignisse als punktuell. Die Standardbeispiele für punktuelle Ereignisse sind ‘flash’, ‘hiccuped’ oder ‘burst’, vgl. [21, S.13] bzw. [99].

Diese Beispiele zeigen meines Erachtens deutlich, dass sie zwar meist nur kurze Zeiträume einnehmen, aber es sich jeweils um Zeiträume handeln muss und sie daher nicht punktuell sind. Sie können nicht durch das Bestehen eines Sachverhaltes zu einem Zeitpunkt charakterisiert werden. Um zu erkennen, dass es blitzt, muss es ein Davor geben, als es noch nicht blitzte, einen weiteren Zeitpunkt (vielmehr ist dies

ebenfalls ein Intervall), in dem der Blitz aufflackert, und ein Danach, an dem der Blitz verschwunden ist. Würde man lediglich einen Zeitpunkt innerhalb des beschriebenen Intervalls betrachten, also eine Momentaufnahme, wäre es schwer zu erkennen, dass es blitzt. Eine Eigenschaft des Blitzes ist es, schnell wieder zu verschwinden. Da die Momentaufnahme (Zustand) aber entweder das Davor (es blitzt noch nicht), das Während (es ist ein heller Schein am Himmel) oder das Danach (es hat aufgehört zu leuchten) festhält, ist es nicht möglich, dass gleichzeitig das Verschwinden des Blitzes und der Blitz selbst zu dem Zeitpunkt vorliegen können. Auch wenn das Zeitintervall, in dem das Ereignis 'Blitz' eintritt, sehr klein ist, muss es ein Zeitintervall und kein Zeitpunkt sein, an dem das Ereignis eintritt. Dieselben Argumente kann man in den Fällen 'hiccuped' und 'burst' anbringen.

Das Problem liegt im folgenden Punkt. Über das Vorliegen eines Zustandes zu einem bestimmten Zeitpunkt kann nur entschieden werden, wenn der Zustand über eine Zeitspanne vorliegt. Nur ist eine Bestimmung des Zeitintervalls für das Vorliegen des Zustandes irrelevant bzw. in den meisten Fällen nicht möglich. Punktuelle Sachverhalte, die sich dadurch auszeichnen, von einem Moment zum anderen eine Änderung zu beschreiben, die sofort rückgängig wird, wie z.B. ein Blitz, ein Schluckauf etc. können keine Zustände charakterisieren, da ihr Vorliegen, wie eben erläutert, nicht in einem Zeitpunkt festgemacht ist. Der Umkehrschluss, alles als Zustand zu bezeichnen, was sich über einen unbestimmten Zeitpunkt ausdehnt [99, S.17], halte ich für verkehrt. Dass die Ausdehnung des Intervalles nicht bestimmt werden kann, heißt nicht, dass es nicht ein endliches Intervall gibt und das Vorliegen eines scheinbar ewig ausgedehnten Zustandes als Ereignis betrachtet werden kann.²⁶

Meine Auffassung von Zustand und Ereignis kann durch folgende Metapher verdeutlicht werden. Wenn ein Ereignis, das ist, was vor Augen tritt, dann ist ein Zustand, das was, vor Augen ist. Man stelle sich vor: Ein Akteur sitzt im Kino und schaut sich einen Film an. Er sieht Handlungsabläufe. Es geschieht ein Mord. Die Polizei eilt zum Tatort. Die Fahndung nach dem Mörder nimmt seinen Lauf. Es jagt ein Ereignis das andere. Plötzlich hält jemand die Filmrolle an und der Akteur erhält ein Standbild.

²⁶Ob jemand notwendig bestehende Sachverhalte als Charakteristika für Zustände erachtet, so dass jemand $2 + 2 = 4$ als Zustand zählt bzw. das Bestehen dieses Zustandes als immer währendes Ereignis erachtet, stelle ich frei. Ich persönlich würde Chisholm zu stimmen, vgl. [44], dass Ereignisse in irgendeiner Weise kontingent sein müssen. Ein Ereignis muss eintreten, demzufolge muss es einen Zeitpunkt geben, an dem das Ereignis nicht eingetreten war. Für Zustände würde ich demzufolge ähnliches postulieren. Teilt jemand diese Intuition nicht, kann er annehmen, dass das Zeitintervall, in dem das Ereignis eintritt, nicht beschränkt ist und Zustände über alle Zeiträume hinweg vorliegen können.

2. Zustände, Sachverhalte und Ereignisse

Er sieht die Szene, wie sie zu dem gegebenen Zeitpunkt ist. Er weiß, wer zu diesem Zeitpunkt wen liebt, wer welche Fakten weiß, wer sitzt, wer steht, wer welche Kleidung trägt etc.: Mit anderen Worten die Menge an Zuständen, die zu dem Zeitpunkt dabei vorliegen, werden ihm standbildhaft vor Augen geführt und nicht die Ereignisse. De Swart fasst dies wie folgt zusammen: „Events move the story line on, whereas states typically provide background informations“ [49, S.348].

Dabei ist es möglich, dass ein und derselbe Sachverhalt sowohl einem Ereignis als auch einem Zustand zuzuordnen ist. In [99, S.24] wird darauf eingegangen, wie ein „stative main clause can be turned into an event expression“. Die Bezeichnung des Sachverhaltes kann unterscheiden, ob ein Sachverhalt einem Ereignis oder einem Zustand zugeordnet wird, vgl. ‘der Sonnenschein’ und ‘dass die Sonne scheint’. Es ist jedoch nicht notwendigerweise der einzelne Sachverhalt, der die Unterscheidung beinhaltet. In dem Satz ‘when Pete came in, I knew that something was wrong’ beschreibt der Satz ein Ereignis, obwohl das Verb ein Zustandsverb ist [99]. Dennoch gibt es Sachverhalte bzw. Propositionen, die sich anhand ihrer Beschreibungen bzw. Bezeichnung eindeutig entweder in die Kategorie ‘charakterisieren Ereignisse’ oder in die Kategorie ‘charakterisieren Zustände’ einordnen lassen.

Moens und Steedman untersuchen verschiedene aspektuale, zeitliche Bestimmungen, die Hinweise auf Ereignisse bzw. Zustände geben können. Nach ihrer Meinung tritt im Falle von Ereignissen häufig ein sogenannter „Nucleus“ als Komplex-Ereignis auf. Ein solcher besteht aus einem Ziel-Ereignis, einem Ereignis, welches die Durchführung dessen beschreibt, und dem Folgezustand, der erreicht wird. ‘He reaches the top’ kann so ein Ereignis-Komplex sein. Es ist ein Ziel-Ereignis gegeben, den Gipfel zu besteigen, eine Durchführung des Auf-den-Gipfel-Steigens und schlussendlich der Folgezustand des Auf-dem-Gipfel-Seins. Letzteres bezeichnen sie ebenso wie ich nicht als Ereignis. Der Grund hierfür ist das Wort ‘Sein’. Ein ‘Sein’ deutet auf einen Zustand hin, wohingegen das Wort ‘Steigen’ eher eine Handlung und damit ein Ereignis bezeichnet. Der Frage, der nun nachgegangen wird, ist, welche Beschreibung eines Ereignisses dieses als Handlung kennzeichnet.

In diesem Kapitel wurde festgehalten, dass eingetretene Ereignisse und Zustände im aktuellen Geschichtsverlauf konkrete Entitäten sind, die durch Sachverhalte charakterisiert werden. Situationen, Zeitpunkte in Geschichtsverläufen, werden durch Mengen von Sachverhalten bzw. deren Sachverhalteigenschaften bestimmt. Sachverhalte, die in einer Situation in einem Geschichtsverlauf bestehen, sind durch Sätze ausdrückbar. Somit liefert der bestehende Sachverhalt in einer Situation über diesen Satz eine Beschreibung der Situation. Wenn dieser Sachverhalt ein Ereignis bzw. Zustand cha-

Charakterisiert, liefert er einen Teil einer Beschreibung des Ereignisses bzw. des Zustandes. Wenn Ereignisse und Zustände betrachtet werden sollen, stellt sich die Frage, wie kann ich anhand der ausgedrückten Sachverhalte erkennen, ob ein Zustand bzw. ein Ereignis charakterisiert wird. Kann vielleicht umgekehrt festgehalten werden, wie ein konkretes Ereignis bzw. ein konkreter Zustand beschrieben werden sollte?

Dies ist interessant, da Handlungen Ereignisse und propositionalen Einstellungen, wie Überzeugungen und Wünschen, Zustände zugrunde liegen, siehe Kapitel 3. Wenn dies der Fall ist, ist es möglich, an einem ausgedrückten Sachverhalt nicht nur abzulesen, dass es sich nicht nur um ein Ereignis, sondern um die Handlung eines Akteurs handelt.²⁷ Für einen Stit-Satz 'der Akteur α sorgt dafür, dass φ ' ist offensichtlich, dass der Satz ein Ereignis benötigt, um wahr zu sein, das dem Akteur als Handlung zugeschrieben wird und ein charakterisierender Sachverhalt besteht, der durch φ gegeben ist. Gibt es bestimmte Beschreibungen, die bereits auf eine Handlung eines Akteurs schließen lassen? Oder gibt es Beschreibungen, die von vornherein eine Charakterisierung als Ereignis ausschließen? Kann anhand der Beschreibung bzw. anhand eines Sachverhaltes möglicherweise eine Einteilung von Ereignissen getroffen werden, die berücksichtigt, wann etwas eine Handlung ist, bzw. wie dieses Ereignis möglicherweise zu anderen Ereignissen bzw. Zuständen in Beziehung steht.

Im nächsten Kapitel werde ich daher Bezeichnungen bzw. Beschreibungen von Ereignissen und Zuständen untersuchen. Im ersten Teil plädiere ich dafür, dass ein Ereignis einer Handlung zugrunde liegt und dass sich das Ereignis aufgrund von Stit-Paraphrasierungen der Bezeichner bzw. der Beschreibungen des Ereignisses als Handlung zu erkennen gibt. Im zweiten Abschnitt werde ich auf die Klassifikation eingehen, die Vendler für Verben eingeführt hat, um der Frage nachzugehen, ob bei Sachverhalten bzw. Propositionen die Beschreibung ausreicht, um etwas als ein Ereignis oder einen Zustand zu charakterisieren. Im Anschluss werde ich diskutieren, ob die daraus resultierende Klassifikation von Ereignissen Aufschluss geben kann, ob es eine Klasse von Ereignissen gibt, die möglicherweise der Klasse aller Handlungen zugrunde liegt, bzw. ob es eine Klasse von Ereignissen gibt, die eine Beschreibung als Handlung ausschließt.

Im dritten Teil werde ich mentale Zustände und mentale Ereignissen näher betrachten. Bei den mentalen Ereignissen werden Unterschiede zwischen mentalen Akten und mentalen Vorgängen festgehalten werden. Dies ist nötig, im Hinblick auf die Model-

²⁷Bei Zuständen ist es wesentlich unstrittiger, ob ein Akteur in einem bestimmten Zustand ist oder nicht, ob eine Überzeugung vorliegt oder nicht. Die Beschreibung 'der Akteur hat die Überzeugung, dass φ ' ist eine Beschreibung des Überzeugungszustandes. Bei Handlungen ist dies weniger offensichtlich.

2. Zustände, Sachverhalte und Ereignisse

lierung von Interaktionen zwischen den mentalen Zuständen, wie Überzeugungen und Wünsche, und den mentalen Ereignissen, wie dem Bilden/Aufgeben/etc. von mentalen Zuständen. Welche Interaktionen sich daraus für die Stit-Operatoren und den Operatoren für die mentalen Zustände ergeben, wird ebenfalls in Abschnitt 3.3 festgehalten.

3 Handlungen oder die Frage „Actions as Events“

Im letzten Kapitel wurde definiert, was Zustände, Ereignisse und Sachverhalte sind. In diesem Kapitel werde nun ich auf die Fragen zurückkommen, die zu Beginn des Kapitels 2 aufgekomen sind. Welche Analyse eines Satzes ist geeignet, um einen Satz als Handlungssatz zu kennzeichnen. Dabei wird auf die Begriffe des Zustandes und Ereignisses zurückgegriffen werden. In der Diskussion bei der Analyse von Handlungssätzen stelle ich zwei Ansätze gegenüber, auf der einen Seite Davidsons Ansatz ‘Actions-as-Events’ auf der anderen Seite die Stit-Paraphrasierung, wie sie Belnap et al. vorgeschlagen wurde.

Ich werde dafür argumentieren, dass einer Handlung zwar ein Ereignis zugrunde liegt, bzw. spreche ich dann abkürzend davon, dass ein Ereignis eine Handlung ist. Doch ist nicht das Ereignis ausschlaggebend, ob eine Handlung vorliegt, sondern die Bezeichnungen des Ereignisses bzw. die Propositionen, die das Ereignis charakterisieren, liefern einen Hinweis darauf, dass es sich um eine Handlung handelt. An dieser Stelle werde ich die Stit-Paraphrasierung einsetzen, um zu zeigen, dass es sich bei dem Ereignis, um eine Handlung handelt. Nach Davidson wird ein Ereignis als Handlung ausgemacht, wenn es eine Beschreibung des Ereignisses gibt, die eine bestimmte Bedingung erfüllt, [48, Essay 3]. Ich werde diese Ansicht abwandeln. Eine Stit-Paraphrase drückt unbestritten einen Handlungssatz aus, so dass im Falle von Ereignisbezeichnungen ebenfalls eine Stit-Paraphrase hinreichend ist, um eine Handlung zu identifizieren, ohne dass es nötig ist, dem Akteur in irgendeiner Weise eine Absichtlichkeit zu zuschreiben.

In Abschnitt 3.2 wird diskutiert, welche Ereignisse bzw. Ereignisbeschreibungen einen Hinweis darauf geben, ob eine Handlung vorliegt und ob eine Stit-Paraphrase eine äquivalente Beschreibung des Ereignisses ist. An dieser Stelle gehe ich auf Vendlers Klassifikation von Verben ein. Bevor untersucht werden wird, welche logische Form mit welchen Wahrheitsbedingungen einem Handlungssatz am nächsten kommt, stelle ich mentalen Zuständen mentale Ereignisse gegenüber. Es wird von Bedeutung

3. Handlungen oder die Frage „Actions as Events“

sein, welcher Satz eine Handlung beschreibt und somit eine Zuschreibung einer Handlung zu einem Akteur ermöglicht, und welche Art von Sätzen eine Zuschreibung von Wünschen und Überzeugungen eines Akteurs ermöglichen. Dafür scheint es notwendig, der Frage nachzugehen, inwieweit Handlungen, die Ereignisse sind, und Überzeugungen, Intentionen und Wünsche, die als mentale Zustände angesehen werden, in Interaktionen treten können und in welchen sie stehen. Es wird diskutiert werden, wie es möglich ist, diese Interaktionen zwischen den Handlungen und den mentalen Zuständen eines Akteurs zu beschreiben. Somit ist eine Auseinandersetzung mit den folgenden Fragen unausweichlich: Was ist ein mentaler Zustand? Was ist ein mentaler Akt? Gibt es mentale Zustände, die der Akteur voluntativ hervorgebracht hat, so dass wir die Bildung des Zustandes als mentalen Akt werten können? Ist bereits anhand der Beschreibung, eine Einordnung des mentalen Vorkommnisses als Akt oder Zustand möglich? Ich werde Antworten auf diese Fragen in Abschnitt 3.3 geben.

3.1 Handlung und das darunter liegende Ereignis

Eine Handlung ist eine besondere Form eines Ereignisses. Niemand würde bestreiten, dass beim Durchführen oder Vollenden einer Handlung ein einziger Zeitpunkt als Beschreibung nicht ausreicht. Es wird ein Zeitintervall betrachtet, in dem die Handlung vollzogen wird. Mir ist bewusst, dass Belnap et al. die Auffassung, dass es sich bei Handlungen um Ereignisse handelt, nicht vertreten wollen [19, S. 78f]. In [17, 19] wird eine Handlung als eine Folge einer Entscheidung des Akteurs zwischen verschiedenen Geschichtsverläufen beschrieben. Die Handlung wird in einem Moment beschrieben. Handlungen als ein Ereignis, dem eine bestimmte Beschreibung zukommt, wie bei Davidson [48, Essay 3], ist mit einer Beschreibung in der Stit-Theorie nicht ohne weiteres in Einklang zu bringen [19, S.79]. Eine Entscheidung in einem Moment zu treffen, ist ebenfalls ein Ereignis, auch wenn es ein mentales ist. Es schließt den Ansatz, der eine Handlung als ein spezielles Ereignis auffasst, nicht von vornherein aus, sondern ist notwendig, da ebenso wie Handlungen Entscheidungen nicht punktuell sind. Es bleibt zu erklären, wie die Beschreibung einer Handlung durch einen instantanen, modallogischen Operator in einem Moment erfolgen kann, wenn die Handlung einen Zeitraum einnimmt. Meine Antwort ist: Ob eine Handlung stattgefunden hat, kann an dem durch die Handlung erreichten Folgezustand festgemacht werden. Natürlich ereignet sich eine Handlung innerhalb eines Zeitintervalles. Es gibt einen Zeitpunkt vor Beginn der Handlung und einen Zeitpunkt, ab dem die Handlung als abgeschlossen erachtet werden kann. Aber um eine Handlung von einem Versuch einer Handlung zu

3.1. Handlung und das darunter liegende Ereignis

unterscheiden, ist die erfolgreiche Durchführung vonnöten. Der Versuch einer Handlung liegt ebenfalls ein Ereignis zugrunde, das möglicherweise sogar eine Handlung ist, je nachdem ob die den Versuch charakterisierenden Sachverhalte und Propositionen Stit-Paraphrasen zulassen. Dieses ist aber verschieden von dem der erfolgreichen Handlung. Der Satz 'er erklimmt den Gipfel' drückt eine abgeschlossene Handlung aus, sobald er den Gipfel erreicht hat. Der Satz 'er versucht, den Gipfel zu erreichen' drückt noch keine abgeschlossene Handlung aus. Weitere Beschreibungen des Versuchs 'er ist los gelaufen, um den Gipfel zu erreichen' können den Versuch als Handlung kennzeichnen. Der Satz 'er hat den Gipfel erreicht' beschreibt jedoch kein Ereignis, sondern einen Zustand. Jedoch gibt die Paraphrase 'er sorgt dafür, dass er den Gipfel erreicht' dieselbe Handlung wieder, die der Satz 'er erklimmt den Gipfel' beschreibt, der zudem die erfolgreiche Durchführung der Handlung beinhaltet. Im Beispiel 'er hat den Gipfel erreicht' könnte man argumentieren, dass Erreichen bereits ein Handlungsverb ist und kein Zustand beschrieben wird. Der Satz 'er sorgt dafür, dass er den Gipfel erreicht hat' ist hingegen unsinnig.

Wenn der Satz 'Alexander ist auf dem Thron' durch einen Stit-Satz ausgedrückt wird, gibt es zwei Möglichkeiten entweder 'Alexander sorgt dafür, dass er in Bälde auf dem Thron ist'. Dies drückt nicht das erfolgreiche Durchführen seiner Handlung aus. Oder das erfolgreiche Durchführen der Handlung wird durch den erreichten Zustand beschrieben, 'Alexander hat dafür gesorgt, dass er auf dem Thron ist'. Diese Paraphrase drückt zwar aus, dass ein bestimmter Zustand vorliegt, aber auch, dass dieser Zustand durch ein Ereignis hervorgebracht wurde. Moens und Steedman bezeichnen ihn als „consequent state“ (Folgezustand) eines Ereignisses [99]. Eine Handlung ist dabei das Ereignis, welches den Zustand erfolgreich hervorgebracht hat. Dennoch ist sie nicht selbst der Zustand. Allerdings, um zu entscheiden, ob eine erfolgreiche Handlung stattfand, genügt die Betrachtung des hervorgebrachten Zustandes in einem Moment. D.h. die Zuschreibung einer Handlung eines Akteurs mit dem Ergebnis, dass dieser oder jener Zustand erreicht wurde, ist punktuell möglich, obwohl die Handlung als ein Ereignis über ein Zeitintervall individuiert wird.

Als Problem könnte erachtet werden, dass nicht Sätze Ereignisse beschreiben bzw. bezeichnen, sondern dass es andere Formen der Ereignisbeschreibung gibt, auf die eine Stit-Paraphrasierung nicht sinnvoll anwendbar ist. Der Name eines Ereignisses ist meist durch ein Substantiv/Eigennamen bzw. die Nominalisierung von Verben oder eine Nebensatzkonstruktion mit 'dass' gegeben.¹ Meines Erachtens ist es am ein-

¹Für einen Überblick, welche Ausdrücke Ereignisse im Deutschen bezeichnen können, siehe z.B. [120, S.7].

3. Handlungen oder die Frage „Actions as Events“

fachsten, ein Ereignis als solches zu benennen, indem das Wort ‘Ereignis’ vorweg geschickt wird. Wenn der ‘Tod Caesars’ ein Ereignis bezeichnet, so benennt ‘das Ereignis, bei dem der Tod Caesars eintritt’ dasselbe Ereignis. Ich möchte keine Ereignisunterscheidung aufgrund ihrer Fein- bzw. Grobkörnigkeit der Bezeichnung vornehmen. Der Ausdruck ‘der Tod Caesars’ bezeichnet bei festgelegtem Geschichtsverlauf dasselbe Ereignis wie die Nominalphrasen ‘der überraschende Tod Caesars’ bzw. ‘der gewaltsame Tod Caesars’, falls der Tod Caesars gewaltsam war. Einer Betrachtungsweise, die Ereignisse als ebenso feinkörnig wie Sachverhalte behandelt, gehe ich nicht nach, sondern stimme Davidsons Ansatz der Grobkörnigkeit eines Ereignisses zu.

Belnap's Bemerkung in [15], dass in der Stit-Theorie, die sich mit Handlungssätzen befasst, eine Unterscheidung zwischen dem Unterlassen und dem Nicht-Vollziehen von Handlungen besser darstellbar ist als im Falle der Handlung-als-Ereignis Betrachtung, werde ich mittels einer ‘Belnap-artigen’ Stit-Paraphrasierung eines Ereignisses entkräften. Nimmt man eine Benennung des Ereignisses vor, so kann diese durch eine Stit-Paraphrasierung mit festem Akteur immer eine Handlung ausdrücken. Ähnlich zu Davidsons Existenzforderung eines Ereignisses, das als absichtlich vom Akteur hervorgebracht betrachtet wird, wenn es als Handlung aufgefasst wird, gibt es für ein Subjekt eine Stit-Paraphrase für jede Benennung des Ereignisses, die ebenfalls eine Benennung des Ereignisses ist. Im aktuellen Geschichtsverlauf wären für die Ereignisbezeichnung ‘der Tod Caesars’ z.B. ‘das Ereignis, bei dem eintritt, dass Brutus für den Tod Caesars sorgt’ bzw. ‘das Ereignis, bei dem eintritt, dass Brutus dafür sorgt, dass der Tod Caesars eintritt’ Stit-Paraphrasen, die dasselbe Ereignis bezeichnen, oder aber auch ‘das Ereignis, bei dem eintritt, dass Brutus für die Ermordung Caesars sorgt’.

Stit-Paraphrase eines Ereignisses Ein Ereignis kann genau dann als eine Handlung eines Akteurs α bezeichnet werden, wenn für jede Bezeichnung ‘Q’ des Ereignisses gilt, dass ‘das Ereignis, bei dem α sorgt dafür, dass Q eintritt’ eine geeignete Paraphrase für ‘Q’ und somit eine Bezeichnung des Ereignisses ist.

Da eine Stit-Paraphrasierung, um ein Ereignis als Handlung zu kennzeichnen, keinerlei Kenntnisse über den hervorgebrachten Sachverhalt oder Absichtlichkeit des Akteurs voraussetzt, beschreiben die Sätze bzw. charakterisieren die zugehörigen Sachverhalte ‘dass der Akteur zum Abendstern fliegt’ und ‘dass der Akteur zum Morgenstern fliegt’ dasselbe Ereignis, obwohl dem Akteur in keiner Form zugänglich oder transparent sein muss, dass der Abendstern der Morgenstern ist.

Es stellt sich nun die Frage, wann etwas eine Bezeichnung für ein Ereignis ist bzw. wann es keine Bezeichnung ist. Die Unterscheidungen im Englischen zwischen per-

3.1. Handlung und das darunter liegende Ereignis

fect nominal und imperfect nominal, z.B. von Bennett [21] und Vendler [152], kann im Deutschen am ehesten durch eine Infinitiv- bzw. Nebensatzkonstruktion wiedergegeben werden. Vendler und Bennett unterscheiden zwischen ‘his pushing of the rock’ und ‘pushing the rock’. Wenn beide Gerundien ins Deutsche übersetzt werden, wird die Assoziation, die Vendler erwähnt, deutlich. Während perfect nominals zumeist mit Nominalphrasen der Form ‘sein Stoßen des Steines’ übersetzt werden, können die Verwendungen des imperfect nominals in englischen Sätzen im Deutschen durch einen mit durch ‘indem’ (o.ä. Konjunktionen ‘dass’, ‘dadurch’) eingeleiteten Nebensatz wiedergegeben werden, ‘indem er den Stein stoß’. Da letztere im Deutschen durch vollständige Nebensätze wiedergegeben werden und die ausgedrückte Proposition daher eher einem Sachverhalt korrespondiert, der ein (möglicherweise anderes) Ereignis näher charakterisiert, anstelle dass er tatsächlich eine Ereignisbezeichnung zu sein scheint. So eignen sich perfect nominals im Deutschen eher als Ereignisbezeichnungen denn imperfect nominals. D.h., während bei einem perfect nominal bereits der Ereignis-Charakter des dahinterliegenden Sachverhaltes offenliegt, ist im Falle des imperfect nominals unklar, ob der korrespondierende Sachverhalt ein Zustand oder ein Ereignis charakterisiert.² Für perfect nominals ist jedoch eine Stit-Paraphrase der Bezeichnung des Ereignisses, wie eben erwähnt, möglich.

Zu einem Ereignis wird eine gewisse Zeitspanne zum ‘sich ereignen’ benötigt und eine Menge an Sachverhalten, die dieses Ereignis charakterisieren und in dieser Zeitspanne, wenn sich ein konkretes Ereignis ereignet, bestehen (in einer gewissen zeitlichen Abfolge). Anhand der das Ereignis charakterisierenden Sachverhalte ergeben sich die Bezeichnungen des Ereignisses. Es kommt vor, dass Bezeichnungen sich auf in das Ereignis involvierte Objekte oder Sachverhalte beziehen, z.B. ‘der Tod Caesars’. Ebenso können aus Beschreibungen des Ereignisses selbst wieder Bezeichnungen abgeleitet werden, z.B. indem die Beschreibung das Ereignis bezeichnet, ‘das Ereignis, auf das die Beschreibung *p* . . . zutrifft’.

Auch wenn sich Ereignisbezeichnungen aus sprachlichen Ausdrücken der das Ereignis charakterisierenden, bestehenden Sachverhalte bzw. den ihnen korrespondie-

²Bennett beruft sich auf Vendlers Unterscheidung „perfect nominals refer to events and imperfect ones to facts“ [21, S.130]. Für ihn gilt: „Events are facts of a kind: many facts are not events (e.g. ones that are not about zones), but every event is a fact. [. . .] We think of facts as ‘events’ when we are naming them with perfect nominals and as ‘facts’ when we are naming or expressing them with imperfect nominals. But they are the very same items, however named“. Seine These lautet: „An imperfect nominal names the fact it expresses; a perfect nominal names a fact that includes the fact it expresses“ [21, S.129]. Somit kennzeichnen perfect nominals, dass der ausgedrückte Sachverhalt („Fact“) ein Ereignis *ist*.

3. Handlungen oder die Frage „Actions as Events“

renden Propositionen ergeben, sind Ereignisse keine Sachverhalte und auch keine bestehenden oder Tatsachen. Sie sollten nicht gleichgesetzt werden, wenn ein Sachverhalt das Ereignis vollständig charakterisiert und der Sachverhalt in der Situation besteht. Ein Sachverhalt ist ein notwendigerweise existierendes, abstraktes Objekt, ein Ereignis hingegen ist konkret und seine Existenz auf den jeweiligen Geschichtsverlauf beschränkt. Ein weiterer Unterschied ist die Feinkörnigkeit. Für das Argument der Feinkörnigkeit von Sachverhalten gegenüber der Grobkörnigkeit von Ereignissen hat Davidson mehrere Beispiele ins Feld geführt.³

Daher werde ich Ereignisse nicht mit einer Klasse von Sachverhalten identifizieren, wie z.B. [41], bzw. als eine bestimmte Art von Tatsachen ansehen, wie z.B. [21]. Einer Beschreibung eines Sachverhaltes ist es nicht notwendigerweise immanent, ob ein Ereignis oder ein Zustand gemeint ist, wie im Beispiel ‘die Sonne scheint’. Wenn der Zeitpunkt fixiert ist, meint sie den Zustand, dass die Sonne scheint. Ebenso kann der Satz und damit die ausgedrückte Proposition und der damit korrespondierende Sachverhalt ein Ereignis beschreiben bzw. sein, nämlich das Ereignis des Sonnenscheins. Voraussetzung dafür ist jedoch die implizite Annahme zweier Zeitpunkte, die den zeitlichen Rahmen dieses Ereignisses kennzeichnen, von wann bis wann das Ereignis stattfand. Dies ist ebenfalls bei Zuständen möglich. Es kann gefragt werden, von wann bis wann bestand dieser Zustand? Dann ist aber das Vorliegen des Zustandes über diesen bestimmten Zeitraum ein Ereignis und kein Zustand.

Eine exakte sprachliche Unterscheidung zwischen Ausdrücken, die ein Ereignis beschreiben bzw. bezeichnen, ist nicht klar zu umreißen. Es ist immer auch eine Frage des Kontextes bzw. der Intuition, ob es sich um die Bezeichnung eines Ereignisses handelt. Im Folgenden möchte ich auf einige sprachliche Unterscheidungen eingehen, die es ermöglichen Unterschiede festzuhalten. Sie sind möglicherweise hinreichend

³Zu den Beispielen zählen die bereits in Abschnitt 2.3 erwähnten, unterschiedlichen Sachverhalte zu ‘Meyer climbed Kibo’ und ‘Meyer climbed the highest mountain in Africa’. Weitere Beispiele von Chisholm sind, dass die in den Sätzen ausgedrückten Sachverhalte ‘the author of Waverley being knighted’ vs. ‘the author of Marmion being knighted’ verschieden sind [43, S.114]. Davidson geht weiter und behauptet, nicht nur dass die Ersetzung coreferentieller Terme in einer Ereignisbeschreibung nicht zu einem anderen Ereignis führt, sondern auch dass es möglich ist, das Tätigkeitswort resp. die Tätigkeit selbst in einer Beschreibung zu ersetzen, so dass diese immer noch dasselbe Ereignis bezeichnet [48, Essay 8]: ‘das Abstechen von Caesar durch Brutus’ ist laut Davidson dasselbe Ereignis wie ‘das Töten Caesars durch Brutus’. Auch wenn die Sachverhalte, dass Brutus Caesar abstach und dass Brutus Caesar tötet, ein und dasselbe Ereignis charakterisieren, sind sie nicht identisch. Die Ereignisse ‘das Abstechen von Caesar durch Brutus’ und ‘das Töten von Caesar durch Brutus’ sind nach Davidson identisch, denn sie haben dieselben Ursachen und Wirkungen.

3.1. Handlung und das darunter liegende Ereignis

aber nicht notwendig. Zuvor werde ich kurz auf mentale Zustände zu sprechen kommen, die in Abschnitt 3.3 ausgiebiger betrachtet werden, und an einem Beispiel den Unterschied zwischen der Überzeugung als mentalem Zustand und dem Ereignis des Habens einer Überzeugung verdeutlichen.

Nach meiner Intuition sind das Zustandekommen einer Überzeugung, das Auftreten eines Wunsches, das Formen einer Intention Ereignisse. Des Weiteren sind es das Aufgeben einer Überzeugung, eines Wunsches oder das Revidieren einer Intention. Allgemein gesprochen ist eine Änderung des mentalen Zustandes bzw. das Festhalten an einem mentalen Zustand ein Ereignis. Am Beispiel der Überzeugung lässt sich dies darstellen. Liegt eine Überzeugung, dass p , im Moment m vor, ist dies ein Zustand. Das Vorliegen der Überzeugung über einen bestimmten Zeitraum hingegen, kann als ein Ereignis betrachtet werden. Die sprachliche Unterscheidung von ‘eine Überzeugung liegt vor’ und ‘das Vorliegen der Überzeugung’ kann einen Hinweis darauf geben, ob ein Ereignis oder ein Zustand beschrieben wird. Eine Nominalphrase als Bezeichnung deutet fast immer auf ein Ereignis hin.

Die Bildung der Überzeugung kann als ein Ereignis betrachtet werden. Es gibt einen Moment m_0 , zu dem der Zustand noch nicht vorlag, und einen Moment m_1 , in dem die Überzeugung (der mentale Zustand) gebildet vorlag. Es fand somit eine Zustandsänderung bei diesem Ereignis zwischen m_0 und m_1 statt. Dem Ereignis kann das Attribut dynamisch zugeschrieben werden. Nun liegt diese Überzeugung, dass p , im Subjekt eine geraume Zeit, z.B. innerhalb des Intervalls $[m_1, m_2]$, vor. Dann wäre das Vorliegen der Überzeugung vom Moment m_1 bis zum Moment m_2 ein Ereignis, dem das Attribut ‘statisch’ zuschreibbar wäre, da im Intervall (m_1, m_2) selbst keine Änderung eintritt, vgl. zu dem Begriff ‘Zustand’ S.34.⁴ Die Überzeugung zu dem oder dem Zeitpunkt $m_3 \in [m_1, m_2]$ ist jedoch ein Zustand. Das Vorliegen der Überzeugung im Intervall $[m_1, m_2]$ ist auch ein Ereignis, wenn die Überzeugung über das Zeitintervall hinaus besteht. Offensichtlich ist, dass bei der Beschreibung, dass ein Akteur eine bestimmte Überzeugung hat, die Frage nach dem Ereignis des Zustandekommens keine Rolle spielt. Ebenso wenig spielt es eine Rolle, wie lange diese Überzeugung in dem Akteur bereits vorliegt. Nur der Zustand zu dem Zeitpunkt der Beschreibung ist von Bedeutung.

⁴Es ist nötig, hier das offene Intervall zu betrachten, da in m_1 die Überzeugung möglicherweise noch nicht vorlag und es für die Dynamik nicht entscheidend ist, da die Längen der Intervalle (m_1, m_2) und $[m_1, m_2]$ nach üblichem Maß identisch sind.

3.2 Klassifikation von Ereignissen

Ein Ereignis kann hinreichend als Handlung eines Akteurs identifiziert werden, wenn jede Beschreibung des Ereignisses dem Akteur die Handlung zuschreibt. Dies ist sicherlich ein ungenügendes Kriterium, um zu entscheiden, ob das Bestehen eines Sachverhaltes durch die Handlung eines Akteurs hervorgebracht wurde. Dass bestimmte Verben und damit gebildete Sätze eher als Beschreibung für Zustände als für Ereignisse dienen, ist eine weit verbreitete Ansicht. Die Unterscheidung von Zuständen und Ereignissen und die weitere Unterteilung von Ereignissen in Handlungen und Nicht-Handlungen kann auf einer Analyse der Bezeichnungen bzw. Beschreibungen basieren, der das Ereignis bzw. den Zustand charakterisierenden Sachverhalte.

Warum ist es möglich, dass es Beschreibungen geben kann, die hinreichend sind, um einem Akteur eine Handlung zuzuschreiben? Ereignisse sind meist schwer abgrenzbar. Was ein bestimmtes Ereignis ausmacht, welcher Sachverhalt charakteristisch ist und welcher nicht, ist nicht möglich, vollständig zu beschreiben, vgl. Abschnitt 2.3 bzw. [21]. Dennoch ist es möglich, Ereignisse zu erfassen. So ist für den Mord an Caesar der durch den Satz ‘Brutus ermordet Caesar’ ausgedrückte Sachverhalt ausreichend, um zu beschreiben, dass es ein Ereignis ‘Mord an Caesar’ gegeben hat, auch wenn es mehr Sachverhalte sind, die dieses Ereignis charakterisieren.

Jedoch ist in diesem Fall die Beschreibung des Ereignisses durch den Satz ‘Brutus ermordet Caesar’ hinreichend, um zu sagen, dass auch alle anderen Beschreibungen des Ereignisses nach Stit-Paraphrasierung mit Brutus als Akteur immer noch dasselbe Ereignis beschreiben.

In diesem Abschnitt beschäftige ich mich der Frage, welche Beschreibungen die Klasse bzw. eine bestimmte Klasse von Ereignissen darstellen. Wir werden zudem feststellen, dass es Klassen von Ereignissen gibt, deren Beschreibungen eine Handlung ausschließen. Schwieriger ist die Frage nach einer Beschreibung *der* Klasse von Ereignissen, denen Handlungen zugrunde liegen. Diese Frage werde ich nicht beantworten können und pflichte Belnap in seiner Bestürzung über die Abwesenheit eines solchen linguistischen Testes bei [19, S.3].

Dennoch ist eine Betrachtung jeder Beschreibung eines Ereignisses nicht notwendig, um sicherzustellen, dass ein Ereignis eine Handlung eines Akteurs darstellt. Für die Untersuchung von Beschreibungen betrachte ich Verben. Bereits Davidson argumentierte, dass das Verb, dasjenige sei, was ausschlaggebend dafür ist, ob ein Satz eine Handlung beschreibt oder nicht. Es wird sich zeigen, dass nicht jedes Verb eindeutig eine Handlung bzw. ein Ereignis oder einen Zustand beschreibt, aber dass es

3.2. Klassifikation von Ereignissen

Verben gibt, die die eine oder andere Einteilung konsequent ausschließen.

Verben werden nach semantischen Kriterien in Handlungs-, Vorgangs- und Zustandsverbklassen unterteilt, vgl. hierfür [77, S.36ff] . ‘Akzeptieren’, ‘töten’ und ‘aufstehen’ sind eindeutige Beispiele für sogenannte *Handlungsverben*. Es wird eine mentale Aktivität bzw. eine Handlung beschrieben, die aktiv durch ein Subjekt ausgeführt wird. ‘Blühen’ und ‘verdauen’ sind Beispiele für *Vorgangsverben*. Im Unterschied zum Handlungsverb, wird ein Geschehen bzw. Vorgang beschrieben, das am oder dem Objekt geschieht, auf das das grammatikalische Subjekt des Satzes referiert. Dieses ist bei einem Vorgang passiv. Bei Vorgängen ebenso wie bei Handlungen wird meist eine Zeitspanne beschrieben, in der sich der Vorgang ereignet. D.h. Vorgangsverben geben einen Hinweis darauf, dass der korrespondierende Sachverhalt ein Ereignis charakterisiert.

‘Glauben’, ‘lieben’ oder ‘stehen’ sind Beispiele für *Zustandsverben*. Durch Zustandsverben wird beschrieben, dass etwas der Fall ist. Es beinhaltet keine Änderung. Zustandsverben beschreiben im Allgemeinen Zustände, solange der Satz sich auf einen festen Zeitpunkt bezieht und nicht auf die Zeitspanne, in der der Zustand vorliegt. Der Satz ‘er steht an der Tür’ beschreibt ohne weiteren Zusammenhang einen Zustand. Der Satz ‘er steht seit halb drei an der Tür’ beschreibt sowohl ein Ereignis, das Vorliegen eines Zustandes, aber auch den Zustand selbst.⁵ Der Zustand, dass er an der Tür steht, wird durch das Ereignis des Vorliegens des Zustandes impliziert.

Ein Zustandsverb kann somit je nach Kontext ebenfalls einen Vorgang oder eine Handlung beschreiben, nämlich dann wenn das Zustandekommen/Aufrechterhalten/Vorliegen/ ... des Zustandes ausgedrückt wird. Handlungsverben bzw. Vorgangsverben implizieren jedoch die Beschreibung eines Ereignisses. Die Frage, ob ein Zustand passiv oder aktiv ist, d.h. ob das Objekt (Akteur), auf das das grammatikalische Subjekt des Satzes referiert, in einem Zustand aktiv oder passiv ist, stellt sich nicht.⁶ Wenn

⁵So zu finden bei Parson [103, S.188]: ‘stand’ im Sinne von ‘irgendwo gelegen sein’ drückt einen Zustand aus, wohingegen ‘stand’ auch als *process verb* fungieren kann und dann ausdrückt, dass eine Person etwas tut, nämlich seit halb drei vor der Tür stehen. Dieses Ereignis kann eine Handlung oder ein Vorgang sein.

⁶In [77, S.36ff] sind diese Aussagen wiederzufinden. Es gibt Verben, die je nach Kontext einen Zustand oder einen Vorgang bzw. eine Handlung beschreiben. Es ist ebenfalls möglich, dass ein und dasselbe Verb je nach Kontext sowohl eine Handlung als auch einen Vorgang wiedergeben kann. Ein Beispiel für ein solches Verb ist ‘schwimmen’: ‘Das Holz schwimmt im Wasser’, (Vorgang). ‘Michael schwimmt im Wasser’, (Handlung). Unpersönliche Passivkonstruktionen sind jedoch nur mit Handlungsverben, wenn sie als solche gebraucht werden, und nicht mit Vorgangsverben möglich. ‘Er steht vom Tisch auf’. ‘Es wird vom Tisch aufgestanden’. ‘Es wird im Wasser geschwommen’. Im Falle des Holzes

3. Handlungen oder die Frage „Actions as Events“

es eine Antwort auf die Frage gibt, ob der mit einem Zustandsverb beschriebene Sachverhalt ein passives oder aktives Geschehen ist, impliziert der Zusammenhang bereits die Betrachtung eines Ereignisses, z.B. das Anhalten des Zustandes, und nicht den Zustand selbst.

So finden sich bei Zustandsverben, wenn diese ins Passiv gesetzt werden, folgende Phänomene wieder. Entweder der Satz wird sinnlos. Bspw. ‘Hans steht an der Tür’ würde, wenn wir nur das Verb ins Passiv setzen, zu ‘Hans wird an der Tür gestanden’. Dies ist unsinnig. Die andere Möglichkeit ist nur noch eine unpersönliche Passivkonstruktion ‘es wird von Hans an der Tür gestanden’. Diese drückt jedoch keinen Vorgang an einem Subjekt aus, sondern ist entweder die Beschreibung eines Zustandes oder die Beschreibung des Vorliegen eines Zustandes. Das Vorliegen des Zustandes wird dabei als aktiv durch den Akteur hervorgebracht beschrieben.

Es tritt ein Phänomen auf, das bei einigen Zustandsverben zu beobachten ist. Durch das Setzen des Zustandsverbs ins Passiv wechselt es die Kategorie in Vorgangs- bzw. Handlungsverb. Bspw. ‘Hans glaubt, dass p ’ wird zu ‘Hans wird geglaubt, dass p ’. Während ‘Hans glaubt, dass p ’ einen Zustand beschreibt, drückt ‘Hans wird geglaubt, dass p ’ möglicherweise keinen Zustand aus, sondern beschreibt ein Ereignis, bei dem Hans äußert, dass p , und jemand die Überzeugung bildet. Allerdings ist dies nicht bei jeder Passivkonstruktion von ‘glauben’ der Fall. Die unpersönliche Passivkonstruktion ‘es wird geglaubt, dass p ’ kann ebenfalls sowohl den Zustand beschreiben, dass jemand glaubt, dass p , oder das Ereignis, bei dem jemand die Überzeugung bildet, dass p . Im Falle von Zustandsverben sind unpersönliche Passivkonstruktionen möglich aber nicht aussagekräftig, ob ein Zustand vorliegt oder nicht. Somit ist in diesem Zusammenhang nicht eindeutig, ob eine Handlung eines Akteurs vorliegt, bei dem der Akteur aktiv ist, oder ob ein Vorgang vorliegt, bei dem der Akteur passiv ist.⁷ Wenn von dem Ereignis der Überzeugung gesprochen wird, ist meist das Auftreten der Überzeugung, die Bildung der Überzeugung etc. gemeint. Die Verben ‘bilden’, ‘auftreten’ etc. fallen in die Klasse Handlungs- bzw. Vorgangsverb und sind charakteristisch für Ereignisse.

Mir ist klar, dass (mentale) Zustände oft als passiv beschrieben werden. Cohen [45] formuliert, dass Glauben ein passiver Zustand ist und erläutert, wie passive Zustände andere Zustände produzieren, in dem Sinne, dass ein Zustand die Grundlage für einen

hätte die Aussage, ‘es wird im Wasser geschwommen’ (Handlung), keinen Sinn.

⁷‘Aktiv’ und ‘passiv’ sind Attribute eines Akteurs, wenn die Frage im Raum steht, ob der Akteur eine Handlung vollzogen hat oder ob es nur ein Vorgang am Akteur ist. Auf der anderen Seite kommen die Attribute der Beugung von Verben zu, z.B. können diese ins Passiv gesetzt werden.

3.2. Klassifikation von Ereignissen

anderen Zustand bildet. Dies ist nicht adäquat. Selbst wenn er nicht den Zustand, sondern das Vorliegen des Zustandes meint, kann sich meines Erachtens, ob etwas passiv oder aktiv ist, nur auf ein Ereignis beziehen, das den Zustand hervorgebracht hat, wobei ein Akteur in irgendeiner Weise involviert ist. Cohen meint zweifelsohne, dass ein Zustand, der nicht durch eine Handlung aktiv hervorgebracht wurde, passiv ist. Sollte bei der Überzeugungsbildung der Akteur aufgrund anderer Überzeugungen eine Überzeugung bilden, ohne dabei selbst aktiv zu werden, heißt das nicht, dass die Überzeugung passiv ist, sondern dass die Überzeugungsbildung möglicherweise in dem Akteur passiv vorgeht (Vorgang).

Aber da ein Zustand nicht notwendigerweise immer durch denselben Vorgang bzw. dieselbe Handlung, sondern möglicherweise einmal aktiv einmal passiv erzeugt wird, scheint mir eine generelle Einteilung in passive und aktive Zustände nicht sinnvoll. Die Zuordnung passiv/aktiv kann sich in diesem Zusammenhang somit immer nur auf das Ereignis beziehen, das den Zustand hervorgebracht hat. 'Der Akteur α hat blond gefärbte Haare' ist eine Zustandsbeschreibung. Der Akteur α ist dabei weder passiv noch aktiv. Lässt das Subjekt α die Haare beim Frisör färben oder färbt α sie sich sie selbst, so ist α im ersteren Fall passiv im letzteren aktiv. Ein Beispiel für ein passives mentales Ereignis ist die Überzeugungsbildung. Bis auf wenige Ausnahmen wird eine Überzeugung passiv durch das Subjekt gebildet. Etwas geschieht dem Subjekt. Im übertragenen Sinne verdaut das Subjekt die Evidenzen und erreicht dadurch den Zustand des Glaubens oder eben nicht.⁸ Die Bildung der Überzeugung beschreibt somit einen Vorgang. Glauben ist jedoch kein Vorgangs- oder Handlungsverb, sondern ein Zustandsverb.

Eine andere auf semantischer Ebene begründete Unterscheidung von Verben trifft Vendler [151], die zu ähnlichen Ergebnissen führt. Aufgrund der unterschiedlichen Auszeichnung von Zeitpunkten die Verben im Einzelnen implizieren, unterscheidet er zwischen *Activities*, *Accomplishments*, *Achievements* und *States*. Verben der Kategorien *Activity* (Aktivität) und *Accomplishment* (Durchführung) besitzen sinnvolle Verlaufsformen, d.h. sie beschreiben Vorkommnisse, die implizieren, dass eine bestimmte Zeitspanne benötigt wird, um ihren Verlauf bzw. Umsetzung zu vollenden. Vendler spricht nur im Falle der *Achievements* (Errungenschaften) und *States* (Vorliegen von Zuständen)⁹ von Zeitintervallen und sonst von sich erstreckenden Zeitpunkten.

⁸Es sei an die Redewendung 'er muss erstmal verdauen, dass p ' erinnert.

⁹Vendler benutzt den Begriff 'state' als etwas das über einem Zeitintervall hinweg besteht. Dies unterscheidet sich von meiner Auffassung eines Zustandes, so dass ich 'state' nicht mit Zustand, sondern mit Vorliegen eines ausgezeichneten Zustandstyps gleichsetze, vgl. Abschnitt 2.2 S.54.

3. Handlungen oder die Frage „Actions as Events“

Dies scheint mir unplausibel. Ein Augenblick (instant) erstreckt sich nicht. Ein sich erstreckender Moment ist ein Zeitintervall, daher spreche ich auch im Falle der Activities und Accomplishments von Intervallen. Durchführungen und Aktivitäten benötigen demnach ein Zeitintervall, in dem sie stattfinden, aber während bei Aktivitäten zu jedem Zeitpunkt des Intervalls die Aktivität dieselbe ist, gibt es bei Durchführungen einen Zeitpunkt im Intervall, der ausgezeichnet ist und die Durchführung abschließt.

Für Aktivitäten kann gefragt werden: Wie lange hat man sie betrieben? Es ist hingegen unsinnig zu fragen: Wie lange hat man für die Aktivität gebraucht? Bei Durchführungen ist es umgekehrt, siehe [151].

Angenommen im Intervall $[t_1, t_2]$ rennt der Akteur α , so tut er zu jedem Zeitpunkt dasselbe. Es handelt sich um eine Aktivität. Baut α ein Haus oder malt α einen Kreis in $[t_1, t_2]$, so gilt für alle $t \in [t_1, t_2]$, dass α einen Kreis malt, aber in $t = t_1$ und $t = t_2$ unterscheiden sich die Bedingungen und jeder Zeitpunkt der Tätigkeit kann einzigartig beschrieben werden. Es hat keinen Sinn zu fragen, wie lange er einen Kreis malt oder wie lange er ein Haus baut, sondern man fragt, wie lange er dafür gebraucht hat, den Kreis zu malen bzw. das Haus zu bauen. Demzufolge wäre das Malen eines Kreises eine Durchführung. Zielloses Kritzeln mit einem Stift hingegen ist demnach eine Aktivität, siehe [151].

Von den Verben, die eine Verlaufsform haben, unterscheidet Vendler, die die keine besitzen, und teilt sie in Achievements (Errungenschaften) und States (Vorliegen von Zuständen) ein. Keine Verlaufsform zu haben, deutet daraufhin, dass sie keine Dauer beinhalten, sondern punktuelle Vorkommnisse beschreiben. Bei Errungenschaften betrachtet Vendler wiederum ein Zeitintervall und zu irgendeinem Zeitpunkt tritt ein ausgezeichneter Moment ein, an dem etwas erreicht ist. In $[t_1, t_2]$ gewann er das Rennen, bedeutet, dass es einen Zeitpunkt $t \in [t_1, t_2]$ gibt, zu dem er das Rennen gewonnen hat. Bei ‘State’ liegt das, was durch das Verb beschrieben wird, in jedem Zeitpunkt $t \in [t_1, t_2]$ vor. Jemand liebt jemanden in $[t_1, t_2]$, dann liebt er ihn auch punktuell in jedem $t \in [t_1, t_2]$, vgl. S. 34. Vendler bemerkt, dass es sich bei ‘State’, dem Vorliegen eines Zustandes, nicht um Handlungen handeln kann. Im Falle der Errungenschaften kann es zwar, es muss aber nicht eine Handlung sein [151].

In Tabelle 3.1 ordne ich Vendlers Unterscheidung in Handlungs- und Vorgangskategorien ein und stelle fest, dass es Verben gibt, die eine Verlaufsform besitzen, jedoch nicht immer eine Handlung beschreiben, so dass entweder die Einordnung als Activities und Accomplishments falsch oder die Bezeichnung irreführend ist.¹⁰

¹⁰Vendler selbst diskutiert am Beispiel ‘seeing’, in welche Kategorien ein und dasselbe Wort aufgrund unterschiedlichen Gebrauchs fallen kann, und dass sich dies nicht auf einen fehlerhaften Gebrauch in

3.2. Klassifikation von Ereignissen

	Handlung	Vorgang
im Verlauf	Activities, Handlungen, die andauern, 'Hans läuft', Accomplishments, die Handlungen beschreiben, 'Hans läuft den Berg hoch'	Activities/Accomplishments ohne Akteur, 'das Wasser läuft', 'der Eimer läuft voll Wasser', 'the time is going by'
punktuell Ablauf impli- zit, jedoch nicht beschrieben	Achievements als vollzogene Handlung, 'er erreicht den Gipfel', States nach vollzogener Handlung, 'er steht am Tisch'	Achievements ohne Handlung, 'der Eimer ist vollgelaufen', States ohne Handlung, 'er glaubt weiterhin, dass p'

Tabelle 3.1: Vendlers Unterscheidungen in Bezug auf Handlung und Vorgang

Die Kategorien Activity oder auch Accomplishment, die einen Akteur beinhalten, können aufgrund der Verlaufsform eine Handlung des Akteurs beschreiben. Die Verlaufsform drückt aus: der Akteur ist gerade dabei, etwas zu tun. Den Ausdruck 'der Akteur ist gerade dabei zu ...', kann man selbst nicht mit einem Vorgangsverb ergänzen, wie z.B. 'ertragen', 'erdulden', 'verdauen' etc., ohne dass eine gewisse Form von aktiver Beteiligung des Akteurs ausgedrückt wird. Demzufolge geben Activity / Accomplishment-Verben mit involviertem Akteur in der Verlaufsform zumeist Handlungen wieder.

Jedoch ist nicht allein durch das Verb bestimmt, in welche Vendlersche Kategorie bzw. Klasse das durch den jeweiligen Satz beschriebene Ereignis fällt. Im Deutschen gibt es Verben, die je nach Kontext in mehrere Klassen fallen können. Ein Beispiel ist 'laufen'. 'Das Subjekt läuft', (Activity). 'Das Subjekt läuft den Berg hinauf', (Accomplishment). 'Das Subjekt läuft seine Bestzeit', (Achievement). Nun liegt dies nicht am Deutschen. Ein Verb im Englischen mit einer Verlaufsform ist 'drive'. Jedoch gibt es Kontexte, in denen der Gebrauch der Verlaufsform falsch wäre und diese demnach Zustände oder Erungenschaften beschreiben. 'He is driving a car', (Activity). 'He

dem jeweiligen Zusammenhang zurückführen lässt, [151].

3. Handlungen oder die Frage „Actions as Events“

is driving home’, (Accomplishment). ‘He drives into a tree’, (Achievement), seltsam hingegen wäre, ‘he is driving into a tree’. ‘The log drives on the river’, (State). Ebenfalls eigenartig klinge ‘the log is driving on the river’ (korrekt wäre ‘the log is floating on the river’, wobei ‘is floating’ keine Verlaufsform ist, sondern die Zuschreibung der Eigenschaft ‘floating’ also eine Zustandsbeschreibung ist).

Vendler hat keine Unterscheidung von Verben, sondern von Ereignissen aufgestellt. Jedoch treten Korrelationen zwischen den vier Klassen und bestimmten Verben und ihren Bedeutungen auf. So fällt ein Ereignis, was durch den Gebrauch des Verbs ‘believe’ charakterisiert wird, fast immer in Vendlers Kategorie ‘state’. Dabei ist ‘state’ in meiner Terminologie ein Ereignis, dass ein Zustand über einen bestimmten Zeitraum besteht. Ein Ereignis, das durch das Wort ‘laufen’ gekennzeichnet wird, fällt zumeist in die Kategorie Aktivitäten, jedoch eben auch in andere Kategorien. Die meisten Verben, die keine Verlaufsform im Englischen haben und sich daher für die Kategorien Aktivität und Durchführung nicht eignen, werden im Deutschen meist als Zustandsverben wiedergegeben, z.B. ‘know’, ‘want’, ‘believe’, etc.

Wiederum gibt es Verben, wie ‘drive’, die sehr wohl eine Verlaufsform haben, jedoch in bestimmten Kontexten Bedeutungen annehmen, die keine Verlaufsform zulassen. So scheint es möglich zu sagen, dass der Kontext darüber entscheidet, in welche Kategorie das gebrauchte Verb einzuordnen ist. In [99, S.17] schreiben Moens und Steedman: „However, it cannot be stressed too often that these aspectual profiles are properties of sentences *used in a context*: sense-meanings of sentences or verbs in isolation are usually compatible with several (or even all possible) Vendlerian profiles“. Sie betonen zudem, dass linguistische Mittel wie Änderungen der Zeit, der Aspekte und der adverbialen Bestimmungen die Einteilung eines Satzes in nahezu jede Klasse ermöglichen. Auch wenn bestimmte Verben einige Kategorien vollständig ausschließen, ist die Einteilung des beschriebenen Ereignis nicht immer nur an dem Verb festzumachen, sondern an dem Kontext, in dem der Satz einen bestehenden Sachverhalt wiedergibt. Es ist nötig, weitere in der Situation bestehende Sachverhalte zu betrachten. Dies ist wenig überraschend, da Ereignisse nicht mit bestehenden Sachverhalten identifiziert werden, ist es somit häufig bzw. gewöhnlich nicht möglich, aufgrund eines beschriebenen Sachverhaltes zu schließen, in welche Vendlersche Klasse das Ereignis fällt, das dieser bestehende Sachverhalt charakterisiert. Möglicherweise sind es sogar mehrere Ereignisse, für die der Sachverhalt charakteristisch ist.

Davidson schlägt vor [48, Essay 6 S.121]: „there are two ways we can imply that a person acted as an agent: we may use a verb that implies it directly, or we may use a verb that is non-committal, and add that the act was intentional.“ Davidson erkennt,

3.2. Klassifikation von Ereignissen

dass es nicht möglich ist, anhand des Verbs allein festzustellen, ob es sich bei einem durch das Verb beschriebenen Ereignis um eine Handlung handelt. Das Hinzufügen ‘it was intentional for x that p ’ zu einem Satz soll kennzeichnen, dass es sich um einen Handlungssatz handelt. Das Hinzufügen von Prädikaten der Form ‘the agent x brings it about that’ hält Davidson jedoch für unplausibel, da diese zwar eine Handlungsbeschreibung garantieren, jedoch nur hinreichend und nicht notwendig sind.

Demzufolge ist ein Test, ob ein Verb eine Verlaufsform im Englischen besitzt, keine hinreichende Bedingung, um eine Handlung durch das Objekt, auf das das grammatische Subjekt des Satzes referiert, zu untermauern, selbst wenn man einschränkt, dass es eine Person sein soll. Es ist jedoch hinreichend dafür, dass etwas keine Handlung ist, wenn die Verlaufsform in der verwendeten Zeitform den Satz unsinnig erscheinen lässt. So beschreibt ‘he knows, that p ’ keine Handlung, denn ‘he is knowing, that p ’ ist unsinnig. Der Satz ‘He reaches the top’, in dem Moment wo er den Gipfel erreicht hat, beschreibt keine vorliegende Handlung mehr, sondern nur eine bestimmte Aktivität bzw. Durchführung, die bereits vollendet wurde, und die Erreichung eines bestimmten Zustandes durch eine nicht explizit erwähnte Aktivität. Der Satz ‘he is reaching the top’ klingt ungrammatisch, vgl. [151]. Das schließt natürlich nicht aus, dass eine Handlung des Subjektes zum Erreichen des Gipfels geführt hat und somit dieses Erreichen aktiv durch das Subjekt herbeigeführt wurde.

Betrachtet man die Stit-Paraphrase des Satzes ‘he reaches the top’ aus Kapitel 1, erkennt man, dass es sich bei ‘he reaches the top’ nicht um eine Handlungsbeschreibung handeln muss. ‘He reaches the top’ (er erreicht den Gipfel, in dem Moment hat er ihn bereits erreicht) und ‘he sees to it that he reaches the top’ (er sorgt dafür, den Gipfel zu erreichen und er wird ihn erreichen) sind nicht äquivalent, da sie über zwei verschiedene Situationen bzw. Zeitintervalle Aussagen treffen. Somit kann letzteres keine Stit-Paraphrase nach Belnap darstellen, da eine Paraphrasierung nicht den Bewertungszeitpunkt ändern sollte, wenn er bereits auf dem Gipfel ist, sorgt er nicht mehr dafür, den Gipfel zu erreichen. Demzufolge würde ersterer Satz keine Handlung beschreiben, und auch nicht deutlich machen, dass in der Vergangenheit eine Handlung mit diesem Ergebnis vollzogen wurde.¹¹

Dies trifft nicht das intuitive Verständnis, das allgemein hin jemand von einem Erreichen des Gipfels hat. Die meisten würden ‘das Erreichen des Gipfels durch α ’ als Handlung von α bezeichnen. So ist es im Falle der Nominalphrase auch unstrittig, dass

¹¹Die Unterscheidung, wann eine Errungenschaft des Akteurs passiv oder aktiv ist, siehe Tab. 3.1, richtet sich dann nach dem Ereignis, das diese Errungenschaft hervorbrachte bzw. ihr voran ging, und somit ist die Errungenschaft selbst nach Belnaps Stit-Paraphrasierung keine Handlung mehr.

3. Handlungen oder die Frage „Actions as Events“

‘das Ereignis, bei dem α dafür sorgt, dass das Erreichen des Gipfels durch α eintritt’ eine Paraphrase von ‘das Erreichen des Gipfels durch α ’ ist, da sie dasselbe Ereignis bezeichnen. Auch die Ereignisbeschreibung, ‘dasjenige Ereignis, bei dem α den Gipfel erreicht’ bezeichnet mit der Stit-Paraphrase dasselbe Ereignis ‘dasjenige Ereignis, bei dem α dafür sorgt, dass er den Gipfel erreicht’, und meines Erachtens ist dies für jede Bezeichnung dieses Ereignisses möglich, vgl. Stit-Paraphrase eines Ereignisses S. 68. Der Test aus Kapitel 1 auf S.9 mithilfe der Stit-Paraphrasierung besagt aber, dass im Falle eines Achievement-Verbs der Satz keine Handlung ausdrückt. Denn eine Stit-Paraphrase ist für Achievement- und State-Verben selten äquivalent zur ursprünglichen Beschreibung des Ereignisses, aufgrund der Tatsache, dass sich das Verb ‘sorgen dafür’ meist in die Kategorie Accomplishment einordnet. Ist es tatsächlich der Fall, dass es sich bei Achievement und State in keinem Fall um Handlungen handeln kann? Vendler selbst behauptet, dass ein Achievement-Verb durchaus eine Handlung bezeichnen kann [151]. Im State-Fall verneint er es.

Nimmt man, das von mir beschriebene Verständnis von Zustand und Ereignis, könnte man die Unterscheidungen wie folgt wiedergeben. Ein Activity-Ereignis ist ein Ereignis, das keinen ausgezeichneten Zustand als Ziel hat, und über einen Zeitraum andauert. Ein Accomplishment-Ereignis ist ein Ereignis, das einen ausgezeichneten Zustand als Ziel hat und ebenfalls über einen Zeitraum andauert. Er läuft den Berg hoch, sobald er im Zustand ‘oben auf dem Berg seiend’ ist, ist die Durchführung des Berg-hoch-Laufens beendet. Ein Achievement-Ereignis ist ein Ereignis, das durch das Erreichens bzw. Hervorbringens eines ausgezeichneten Zustand gekennzeichnet ist. Ein State-Ereignis ist ein Ereignis, das sich durch das Fortbestehen eines Zustandes auszeichnet. State und Achievement werden von Vendler über einem Zeitintervall betrachtet, so dass der Gebrauch der Bezeichnung ‘State’ und ‘Achievement-Ereignis’ nur folgerichtig ist. Sie beinhalten selbst jedoch keine andauernde Tätigkeit.

Im Falle der Achievement- und State-Ereignisse sollte nicht der Satz, der die Errungenschaft (den ausgezeichneten Zustand nach dem Ereignis) oder den vorliegenden Zustand beschreibt, mit der Stit-Paraphrasierung untersucht werden. Dem in der Errungenschaft erreichten bzw. dem im State-Fall vorliegenden Zustand ging ein Ereignis voraus, das Zustandebringen des jeweiligen Zustandes. Dieses Ereignis des Eintretens des Zustandes sollte mittels einer Stit-Paraphrasierung auf eine Handlungsbeschreibung hin untersucht werden. Im Falle des Achievement-Ereignisses ist es das Ereignis selbst, das sich durch den errungenen Zustand als solches auszeichnet. Das Ereignis des Erreichens des Zustandes enthält im Achievement-Fall implizit diesen Vorgang bzw. diese Handlung, die den Zustand zustande bringt. Im Falle des State-

3.2. Klassifikation von Ereignissen

Ereignisses stellte Vendler dar, dass dieses keine Handlungen sein kann. Dennoch kann das Hervorbringen des State-Zustandes, welches wiederum ein Ereignis ist, unter Umständen eine Handlung darstellen. Obwohl States Ereignisse sind, die durch eine Handlung aktiv hervorgebracht wurden, beschreiben sie selbst keine Handlung.

Stellt sich heraus, dass das untersuchte Achievement-Ereignis durch Stit-Paraphrasen der Bezeichnungen äquivalent beschrieben werden kann, ist es im Achievement-Fall möglich, dass es eine Handlung ist. Vendler schreibt, dass Errungenschaften u.U. Handlungen sind. Meines Erachtens sind sie es genau dann, wenn aus dem Kontext hervorgeht, dass der erreichte ausgezeichnete Zustand durch eine Handlung erreicht wurde, siehe auch [151]. Wenn man also untersuchen möchte, ob ein Ereignis, das in die Kategorie Achievement fällt,¹² eine Handlung ist, muss man dasjenige Ereignis, das dazu führte, dass der Zustand erreicht wurde, mit der Stit-Paraphrasierung untersuchen.

Oft ist es sogar der Fall, dass wir mit dem Achievement-Verb ein Ereignis benennen, damit jedoch nicht dieses Ereignis beschreiben wollen, sondern dasjenige, welches zu dieser Errungenschaft führte, sogar dann, wenn es noch nicht feststeht, ob der zu erringende Zustand tatsächlich erreicht wird. Ryle schreibt, „it has been all the easier to overlook these differences, since we very often borrow the achievement verbs to signify the performance of the corresponding task activities, where the hopes of success are good. A runner may be described as winning his race from the start, despite the fact that he may not win it in the end; and a doctor may boast that he is curing his patient’s pneumonia, when his treatment does not in fact result in the anticipated recovery. ‘Hear’ is sometimes used as a synonym of ‘listen’ and ‘mend’ as a synonym of ‘try to mend’“ [117, S.149f]. Dies gibt aber wiederum einen Hinweis darauf, dass es Handlungen gibt, die erst als durchgeführt erachtet werden können, wenn es einen Zustand gibt, der Ziel dieser Handlung war.

Zusammengefasst gibt es folgende Möglichkeiten, eine Handlung anhand eines Satzes oder der Ereignisbezeichnungen eines Ereignisses zu erkennen: Falls es sich um einen Satz handelt, der ein Ereignis charakterisiert, dass in die Kategorien State oder Achievement fällt, entscheiden die Beschreibungen des Ereignisses, das Zustandebringen des jeweiligen ausgezeichneten Zustandes, darüber, ob eine Handlung vorliegt. Für die Kategorien Accomplishment/Activity-Ereignis genügt es, das Ereignis selbst zu testen. Anhand der Stit-Paraphrase des Satzes wird die Handlung deutlich. Im Falle

¹²Im Falle von mentalen Akten [Handlungen] sind es meist Achievement- Ereignisse, bei denen der ausgezeichnete Zustand ein mentaler Zustand ist, z.B. die Akzeptanz der Wahrheit einer Proposition bzw. das Formen einer Intention.

3. Handlungen oder die Frage „Actions as Events“

eines Achievement-Satzes ‘he reaches the top’ beschreibt er eine Handlung, falls es eine Beschreibung des Ereignisses (Satzes) gibt, das die Durchführung oder Aktivität wiedergibt, die dazu führte, dass er den Gipfel erreicht, und für diese Beschreibung gilt, dass sie mit der Stit-Normalform des Satzes ‘he sees to it that he reaches the top’ äquivalent ist. Das Ereignis ‘sein Erreichen des Gipfels’ bezeichnet eine Handlung, wenn es für alle Bezeichnungen des Ereignisses gilt, dass sie eine Paraphrase zu ‘das Ereignis, bei dem er dafür gesorgt hat, den Gipfel zu erreichen’ sind. Die Frage, die hierbei offen zu bleiben scheint, ist die nach der erfolgreichen Durchführung. Wie Ryle in seinen Beispielen deutlich macht, verwenden wir Verben, die eine Errungenschaft beschreiben, obwohl es noch nicht feststeht, dass wir den ausgezeichneten Zustand auch tatsächlich erreichen. Jedoch fällt das Verb „sorgen für“ in die Klasse der Durchführungen, so dass generell angenommen wird, wofür gesorgt wurde, das ist der Fall.

Auch wenn eine Handlung ein Ereignis ist und in einem Zeitintervall stattfindet, ist es nicht der Zeitraum, durch den bestimmt ist, ob die Handlung abgeschlossen wird. Vielmehr ist eine Handlung eine Durchführung im Vendlerschen Sinne oder ein Ereignis, das nach [99] als ein Nukleus bezeichnet wird: Ein Ereignis, welches sich aus drei Teilen einer „culmination“, einem „preparatory process“ und einem „consequent state“ zusammensetzt. Die Entscheidung, etwas zu tun, wäre das erste Teilereignis „culmination“, der Prozess, der dem Abschluss der Durchführung vorausgeht, das zweite Teilereignis, aber erst das Hervorbringen des Zustandes „consequent state“ nach der abgeschlossenen Handlung entscheidet, ob es bei einem Versuch geblieben ist oder zu einer Vollendung der Handlung kam. Demzufolge ist es möglich, eine Stit-Paraphrasierung zu benutzen, um einen Handlungssatz oder eine Ereignisbeschreibung als Handlungsbeschreibung zu identifizieren, selbst wenn es sich um ein State- oder Achievement-Ereignis handelt. Gleichzeitig rechtfertigt dies, dass zu *einem* Zeitpunkt in einem festgelegten Geschichtsverlauf bewertet wird, ob eine Handlung stattgefunden hat, wie im Falle des deliberativen Stit-Operators, auch wenn das Ereignis der Handlung selbst einen Zeitraum einnimmt. Somit ist die Frage, ob etwas eine Handlung beschreibt, erst zu beantworten, wenn die Handlung durchgeführt ist, d.h. der „consequent state“ erreicht ist.

Dies ist auch ein Argument, warum es semantisch nicht sinnvoll ist, den Moment, in dem sich der Akteur entscheidet, die Handlung auszuführen, als Moment der Bewertung heranzuziehen, wie im Falle der Wahrheitsbedingungen des *astit*-Operators, Definition 3 siehe S. 18. Eine Entscheidung, die dazu geführt hat, den oder den Zustand herbeizuführen, kann aufgrund der äußeren Umstände bei der Durchführung in ande-

3.3. Mentale Zustände und mentale Ereignisse

ren Handlungen resultieren. Erst die Sachverhalte, die nach dem Vollziehen der Handlung bestehen, und der damit erreichte Zustand charakterisieren die Handlung. Wenn der Arzt sich entschieden hat, seinen Patienten zu heilen, ist die Behandlung noch nicht notwendigerweise erfolgreich. Unabhängig vom Behandlungserfolg hat sich der Arzt aber für einen Geschichtsverlauf entschieden und zwar für den, in dem er ihn behandelt (Durchführung). Man könnte einwenden, der Arzt hat sich nur entschieden, den Patienten zu behandeln, und das erfolgreiche Behandeln als eine andere Handlung angesehen. Das halte ich jedoch im Allgemeinen für unplausibel. Dies wäre höchstens der Fall, wenn es für den Arzt eine andere Entscheidung gibt, Handlungen zu vollziehen, die eine erfolgreiche Behandlung verhindern. Er hätte sich dann entschieden, den Patienten zu behandeln, aber nicht zu heilen. Jedoch würde ich dann nicht mehr von einer Heilbehandlung sprechen. Die Entscheidung des Arztes für die Behandlung des Patienten ist in den meisten Fällen dieselbe Entscheidung wie diejenige, den Patienten mit der Behandlung zu heilen.

3.3 Mentale Zustände und mentale Ereignisse

Im folgenden Abschnitt werden mentale Vorkommnisse, wie das Formen, Bilden, Haben und Revidieren von Überzeugungen, Wünschen und Intentionen daraufhin untersucht, in welchen Beziehungen sie zu Begriffen wie Ereignis, Handlung und Zustand stehen bzw. inwieweit sie selbst welche sind. Dazu möchte ich die in Kapitel 2 und 3 dargelegten Begriffe auf mentale Ereignisse und mentale Zustände übertragen.

Unter mentalen Ereignissen verstehe ich das Formen, Beibehalten und Aufgeben von mentalen Zuständen oder generell Aktivitäten, die auf das Mentale beschränkt sind. Diese mentalen Ereignisse sind wiederum in *mentale Akte* und *mentale Vorgänge* analog zu den Begriffen 'Handlung' und 'Vorgang' sowie mentales Vorkommnis einzuteilen. Ein mentaler Vorgang ist ein mentales Ereignis, das in dem Akteur vorgeht, ohne die Möglichkeit aktiv einzugreifen. Im Folgenden spreche ich dann zumeist von Subjekten statt von Akteuren. Dahingegen ist ein mentaler Akt ein mentales Ereignis, das der Akteur in irgendeiner Weise aktiv mitgestaltet oder über das er in irgendeiner Weise Kontrolle hat.

Das Bilden einer Überzeugung (Achievement-Ereignis) fällt in die Kategorie mentaler Vorgang. Das Fassen einer Intention (Achievement) oder das Akzeptieren einer Aussage (Accomplishment) als wahr zähle ich zu den mentalen Akten. Unter Akte fallen sowohl Accomplishments, Achievements als auch Activities im Vendlerschen Sinn. Das Rechnen im Kopf wäre ein Beispiel für eine mentale Activity. Das Be-

3. Handlungen oder die Frage „Actions as Events“

rechnen von 79 + 139 wäre ein Accomplishment-Ereignis. Wenn die Unterscheidung zwischen Accomplishment, Achievement und Activity in der nachfolgenden Diskussion unerheblich ist, spreche ich im Allgemeinen von mentalen Akten und mentalen Vorgängen, wobei die Unterscheidung darauf abzielt, ob der mentale Zustand aktiv (d.h. in irgendeiner Weise unter der Kontrolle des Akteurs) hervorgebracht wird bzw. passiv im Subjekt gebildet wird.

Falls es unerheblich für das mentale Ereignis ist, ob der Zustand aktiv oder passiv hervorgebracht, beibehalten oder aufgegeben wird, spreche ich von einem mentalen Vorkommnis. Unter ein mentales Vorkommnis fällt z.B. ein State-Ereignis der Art, dass in einem Subjekt in einem bestimmten Zeitraum eine Überzeugung vorliegt. Dies unterscheidet sich natürlich vom mentalen Zustand der Überzeugung des Subjektes zu einem bestimmten Zeitpunkt.¹³

Um zu prüfen, ob im mentalen Fall die Stit-Normalform der Beschreibung eines Sachverhaltes eine Handlung bzw. einen mentalen Akt beschreibt, betrachte man die Beschreibung der Ereignisse: Subjekt α akzeptiert, dass p , und Subjekt α fasst die Intention, dass p . Die Variationen des Kontextes sind bei mentalen Ereignissen und auch Zuständen weniger möglich als z.B. im Falle des Schwimmens. Wenn es bei, ‘es schwimmt im Wasser’, noch unklar sein könnte, ob es sich um ein Kind oder Holz handelt, ist es im Falle ‘es glaubt, dass p ’ unstrittig, dass ‘es’ ein Subjekt bezeichnet. Es ist immer ein Subjekt, das dasjenige ist, welches den Akt oder Vorgang erlebt. Daher ist die Unterscheidung von Vorgang und Akt abhängig von den mentalen Gegebenheiten des Subjektes. Betrachten wir die Unterscheidung zwischen Accomplishment/Activity und Achievement/State. Fällt ein mentales Ereignis in die erste Kategorie, kann es sich nur um einen mentalen Akt handeln, da das Subjekt die Aktivität bzw. die Durchführung ausführt. ‘Das Subjekt α fasst die Intention’ impliziert, dass α gerade dabei ist, eine Intention bzw. die Intention, dass p , zu fassen. Demzufolge fällt das Beispiel als Durchführung in die Kategorie Accomplishment/Activity. Das Fassen einer Intention wäre somit ein mentaler Akt.

Im Vergleich dazu ist dies im Achievement/State-Fall komplizierter. Die Aussage ‘ α

¹³Es gibt Autoren, die Intentionen als mentale Zustände ansehen [25]. Dies wird, z.B. von Thompson, bestritten: Absichten sind Akte [142]. Daher spreche ich vom Fassen der Intention, um deutlich zu machen, dass ich das Ereignis des Auftretens oder Zustandekommens der Intention meine, das ich ebenfalls zu den mentalen Akten zähle. Wenn eine Intention gefasst ist und im Subjekt vorliegt, d.h. zu jedem Zeitpunkt des State-Ereignisses ‘Vorliegen der Intention, dass p ’, schließe ich mich Bratman an und betrachte eine Intention oder die Absicht, dass p , als einen mentalen Zustand. Das Ereignis, eine bestimmte Intention über einen gewissen Zeitraum zu haben, fällt dann unter die Kategorie State.

3.3. Mentale Zustände und mentale Ereignisse

beabsichtigt, dass p fällt in die Kategorie Achievement/State. Notwendig dafür, dass das Subjekt beabsichtigt, dass p , ist jedoch ein Ereignis, nämlich dass das Subjekt die Absicht formt, dass p . Betrachtet man die Stit-Paraphrasen für diese Ausdrücke, dann drückt 'das Subjekt sorgt dafür, dass es p beabsichtigt' dasselbe aus wie 'das Subjekt formt die Absicht, dass p '. In beiden Fällen liegt im Subjekt die Absicht, dass p nach dem beschriebenen Ereignis vor. Würde man die Stit-Paraphrase des letzten Satzes bilden 'das Subjekt sorgt dafür, dass es die Absicht, dass p , formt' erhalten wir wiederum einen äquivalenten Ausdruck. Demzufolge ist die Stit-Paraphrase des Satzes, der das Vorliegen des mentalen Zustandes beschreibt, äquivalent zu einer Beschreibung des Bildens des mentalen Zustandes und beschreibt somit dasselbe Ereignis. Aus dieser Betrachtung ergibt sich zudem, dass das Fassen einer Intention ein mentaler Akt ist.

Dass das Bilden eines mentalen Zustandes ein mentaler Akt ist, kann nur für mentale Zustände gelten, deren Bildung bzw. Zustandekommen an Entscheidungen bzw. unter der Kontrolle des Akteurs liegt. Mentale Akte liegen nur in den Fällen vor, in denen allein ein Dafür-Sorgen, dass der mentale Zustand gebildet wird, den mentalen Zustand tatsächlich zur Folge hätte. Sollte im Subjekt ein mentaler Zustand vorliegen, betrachten wir mit der Stit-Paraphrase der Beschreibung des Zustandes das Ereignis, das den Zustand hervorgebracht hat, und vernachlässigen, dass möglicherweise ein anderes Ereignis, für das der Akteur gesorgt hat, das Ereignis 'Bildung des Zustandes' zur Folge hatte. Dies wäre eine Form der indirekten Kontrolle, auf die ich später eingehen werde, vgl. Abschnitt 4.1. Wenn es sich bei der Bildung eines Zustandes um einen mentalen Vorgang handelt, sollte die Stit-Paraphrase der Zustandsbeschreibung somit kein Äquivalent zur Beschreibung des Bildens des Zustandes sein.

Im Falle der Überzeugung beschreibt ' α glaubt, dass p ', dass ein Zustand vorliegt. Ein Ausdruck der Form 'ich Sorge dafür, dass ich glaube, dass p ' impliziert jedoch nicht, dass ich ab sofort anfangen zu glauben, dass p . Als ein Beispiel zur Verdeutlichung des Vorganges bzw. Aktes und der verschiedenen Formen des Dafür-Sorgens in beiden Fällen stelle man sich folgende Situation vor. Sie sitzt in einem stickigen Saal, die Luft ist schlecht. Das Fenster ist geschlossen. Es wäre gut, ein Fenster zu öffnen. Ohne weiteres kann sie dafür sorgen zu beabsichtigen, dass sie das Fenster öffnet, d.h. dass sie die Absicht formt, das Fenster zu öffnen. Das Dafür-Sorgen, dass sie beabsichtigt, führt dazu, dass sie die Absicht bildet. D.h. sie beabsichtigt, dass Fenster zu öffnen. Im Fall der Überzeugungsbildung kann sie nicht ohne Weiteres die Überzeugung formen, dass sie das Fenster öffnet. Sie kann die Überzeugungsbildung indirekt beeinflussen, indem sie das Fenster öffnet. Nachgerade kann sie die Überzeugung bilden, dass sie das Fenster öffnet. Sie hat dabei aber keine direkte Kon-

3. Handlungen oder die Frage „Actions as Events“

trolle über die Überzeugungsbildung ausgeübt, sondern indirekte Kontrolle über das Dafür-Sorgen, dass sie das Fenster öffnet. Das Dafür-Sorgen, dass sie glaubt, dass p , resultiert aus dem Dafür-Sorgen, dass p . Dass das Dafür-Sorgen, dass sie die Überzeugung bildet, dass sie das Fenster öffnet, tatsächlich bei geschlossenem Fenster ohne ihr weiteres Zutun dazu führt, dass sie die Überzeugung bildet, dass sie das Fenster öffnet, scheint unsinnig.

Nun könnte jemand argumentieren: Sie kann die Überzeugung bilden, dass sie das Fenster öffnen wird. D.h. sie kann dafür sorgen, dass sie glaubt, dass sie das Fenster öffnen wird. Dies ist wiederum eine Form der indirekten Überzeugungsbildung. Was ist die Evidenz dafür zu glauben, dass sie das Fenster öffnen wird? Die einzig mögliche Antwort müsste lauten, dass sie die Absicht hat, das Fenster zu öffnen. Somit bezieht sich die Kontrolle, die sie ausgeübt hat, nicht auf die Überzeugungsbildung, sondern auf die Absicht, das Fenster zu öffnen. Das Ereignis ‘das Formen der Absicht, das Fenster zu öffnen’ hat den Zustand ‘die Absicht haben, das Fenster zu öffnen’ zur Folge. Aus dieser Absicht resultiert die Überzeugungsbildung, dass sie das Fenster öffnen wird. Dafür zu sorgen zu glauben, dass sie das Fenster öffnen wird, setzt in dieser beschriebenen Situation das Ereignis voraus, die Absicht zu formen, das Fenster in einem zukünftigen Moment zu öffnen. Dieses Ereignis wäre der mentale Akt und nicht die Überzeugungsbildung, die sich als mentaler Vorgang möglicherweise an den Akt anschließt.

Die Stit-Paraphrase einer Beschreibung einer Überzeugungsbildung ist somit in den meisten Fällen nicht äquivalent zu der Überzeugungsbildung. Die Stit-Paraphrase einer Überzeugungsbildung beschreibt meist ein Ereignis, aus dem die Überzeugungsbildung resultiert. Jedoch ist dies nicht notwendigerweise in beiden Fällen dasselbe Ereignis. Intuitiv versteht man unter ‘er sorgt dafür, dass er glaubt, dass p ’, dass das Subjekt Handlungen ausführt, die darauf hinauslaufen, dass er Evidenzen sammelt, die für p sprechen, so dass ‘er glaubt, dass p ’ wahr wird. Der Satz drückt nicht aus, dass der Folgezustand (consequent state) dieses ‘Dafür-Sorgens’ tatsächlich die Überzeugung, dass p , ist, wie es im Falle der Absicht wäre. Die gesammelten Evidenzen könnten statt für p gegen p sprechen, so dass aus seinem Dafür-Sorgen, dass er glaubt, dass p , folgt, dass er nicht glaubt, dass p . Die Stit-Paraphrase des Zustandes ist somit nicht semantisch äquivalent zur Beschreibung der Bildung des mentalen Zustandes. Im Gegensatz dazu ist der Satz ‘er sorgt dafür, dass er beabsichtigt, dass p ’ äquivalent zu ‘er beabsichtigt, dass p ’.

Demzufolge kann der Zustand der Überzeugung nicht durch einen mentalen Akt allein hervorgebracht worden sein. Somit ist das Bilden des Überzeugungszustandes

3.3. Mentale Zustände und mentale Ereignisse

kein mentaler Akt und kann nicht unter die Kategorie Accomplishment bzw. Activity fallen. Es handelt sich bei der Bildung der Überzeugung um einen Vorgang am Subjekt. Da jedes mentale Accomplishment/Activity-Ereignis einen Akt des Subjektes darstellt, fällt die Überzeugungsbildung nicht in die Kategorie Accomplishment/Activity, sondern Achievement/State. Im Fall der Überzeugungsbildung tritt eine Zustandsänderung ein. Somit ist die Kategorie State ausgeschlossen. Eine Überzeugungsbildung ist daher ein Achievement-Ereignis, für das im Normalfall keine Beschreibung eines Accomplishment-Ereignisses existiert, die das Erreichen des ausgezeichneten Zustandes als Handlung darstellt.

Eine ähnliche Behauptung findet sich bei Ryle. Er unterscheidet zwischen Task-Verben (oder auch try verbs) und Achievement-Verben („got it verbs“) und ordnet die kognitiven Verben unter letztere ein [117, S.151f]. Er hält den Unterschied wie folgt fest: „One big difference between the logical force of a task verb and that of a corresponding achievement verb is that in applying an achievement verb we are asserting that some state of affairs obtains over and above that which consists in the performance, if any, of the subservient task activity“ [117, S. 150]. Ryle lehnt es ab, dass kognitive oder perzeptuelle Verben Durchführungen (performance) oder Aktivitäten sind. „They are not tryings, but things got by trying or by luck“ [117, S. 152].

Nun könnte man Überzeugungen nicht nur bilden, sondern fassen. ‘Fassen’ scheint eher in die Kategorien, die Akte bzw. Handlungen kennzeichnen, zu fallen. Dass ‘ α die Überzeugung fasst, dass p ’, klingt ungewöhnlich bzw. schwingt im Fassen der Überzeugung eine gewisse Gleichgültigkeit gegenüber den Evidenzen mit. Wenn er die Überzeugung gefasst hat, dass p , kann kommen, was wolle. Er hält daran fest. Dies ist jedoch keine Form eines perzeptuellen oder kognitiven Prozesses, sondern drückt eher eine Akzeptanz, dass p , bzw. ein Beharren auf p , ohne dass notwendigerweise ein Zustand, in dem er glaubt, dass p , tatsächlich vorliegt. Wie ich im Weiteren ausführen werde, sind ‘das Akzeptieren, dass p wahr ist’ und ‘das Glauben, dass p ’ zwei verschiedene Ereignisse bzw. Zustände.

Im Falle der Akzeptanz liegt eine Besonderheit vor. ‘Das Subjekt akzeptiert, dass p ’ bezeichnet sowohl das Vorliegen des mentalen Zustandes als auch das Zustandekommen dessen (mentaler Akt). Handelt es sich nun in Vendlers Terminologie um ein State oder ein Accomplishment-Ereignis? Die Antwort lautet weder noch. Die Aussage ‘das Subjekt akzeptiert, dass p als wahr’ beschreibt die Entscheidung des Subjektes, eine bestimmte propositionale Einstellung zu p anzunehmen, und gleichzeitig drückt es aus, diese auch beizubehalten. Der Sachverhalt charakterisiert zwei Teilereignisse, die den Kategorien Accomplishment bzw. State zugeordnet werden können. Das gesam-

3. Handlungen oder die Frage „Actions as Events“

te Ereignis fällt weder in die Kategorie Accomplishment noch State, sondern ist ein Achievement-Ereignis, da ein ausgezeichneter Zustand erreicht wird.

Vendlers Tests bestätigt dies, vgl. [151]. ‘Das Subjekt ist gerade dabei zu akzeptieren, dass p ’ ist ungewohnt und drückt keinen Verlauf also keine Activity oder Accomplishment aus. Es ist eher möglich zu fragen ‘seit wann akzeptiert er, dass p wahr ist?’ (Achievement) als zu fragen ‘wie lange dauerte es zu akzeptieren, dass p ?’ (Accomplishment) bzw ‘für wie lange akzeptiert er, dass p ?’ (State). Zu fragen ‘wie lange brauchte er zu akzeptieren, dass p ?’ und ‘wie lange brauchte er, bis er endlich akzeptierte, dass p ?’ ist nicht unmöglich. Allerdings wird nicht nach dem Ereignis der Akzeptanz gefragt, sondern nach der Dauer eines Ereignisses, welches dem Ereignis ‘er akzeptiert, dass p ’ vorausging. ‘Wie lange hat es gedauert, bis er (endlich) akzeptiert hat, dass p ?’ wäre die Frage nach dem Gesamtereignis, bestehend aus dem der Akzeptanz vorausgehenden Ereignis und der Akzeptanz selbst. Das Gesamtereignis ist ein Accomplishment-Ereignis, was sich anhand der Frage erkennen lässt. Die Frage ‘wie lange hat er akzeptiert, dass p ?’ wird meist verwendet, um zu fragen, wie lange hielt der durch die Akzeptanz erzielte Zustand an, und fragt eher nach dem State-Ereignis, welches nach der Akzeptanz eintritt. So kann nach Teilereignisse gefragt werden. Die Frage, die am offensichtlichsten erscheint, ist erstere ‘seit wann akzeptiert er, dass p ?’ . Daher ist das Ereignis Akzeptanz, dass p wahr ist, eher in die Kategorie Achievement als in State oder Accomplishment einzuordnen. Eine Stit-Paraphrasierung des Satzes ‘das Subjekt sorgt dafür, dass es akzeptiert, dass p ’ macht deutlich, dass ‘das Subjekt akzeptiert, dass p ’ einen mentalen Akt statt eines mentalen Vorganges beschreibt.

Die Unterscheidung zwischen Akt und Vorgang treffe ich aufgrund der Kontrolle, die wir über die einen mentalen Ereignisse und über die anderen möglicherweise nicht haben. ‘Möglicherweise nicht’ soll implizieren, dass etwas, das als mentaler Vorgang angesehen wird, unter anderen Umständen ein Akt sein kann. Überzeugungen und Wünsche bilden und ändern sich aufgrund dessen, wie unsere Umwelt auf uns einwirkt, sich präsentiert bzw. welche Bedürfnisse wir aufgrund von Trieben, Emotionen, etc. haben.¹⁴ Sie liegen im Falle der Überzeugung fast nie unter unserer direkten in-

¹⁴Hier ist die von Searle benannte, bereits bei Anscombe beschriebene Direction of fit (Passrichtung) gemeint, jedoch in beide Richtungen Geist-Welt und Welt-Geist. Überzeugungen adaptieren sich an der Welt und sind Geist-Welt gerichtet. Das Ziel einer Überzeugung nach Searle ist, dass der Überzeugungsgehalt wahr ist, sozusagen sich der Welt anpasst. Bei Wünschen ist es umgekehrt. Das Ziel eines Wunsches ist es, die Welt seinem Wunschgehalt anzupassen. Wenn die Welt unseren Wünschen angepasst ist, so können sich Wünsche ändern bzw. sich neue bilden. Auf die Passrichtung gehe ich in Abschnitt 4.1 ein.

3.3. Mentale Zustände und mentale Ereignisse

haltsgerichteten Kontrolle.¹⁵ Daher stellen Überzeugungsbildungen zumeist mentale Vorgänge dar. Im Fall der Wunschbildung ist dies nicht so eindeutig. Wir haben Kontrolle darüber, was wir uns wünschen, üben diese bei der Bildung jedoch nicht bzw. nur selten aus.

Unter Kontrolle und damit als einen möglichen mentalen Akt benenne ich das Annehmen/Ablegen/Bilden/Formen/Fassen/Beibehalten etc. von Einstellungen. Kontrolle wird dann ausgeübt, wenn es an meiner Entscheidung liegt, dass diese Einstellung angenommen/abgelegt/ etc. wird oder nicht. Grob gesagt liegt im mentalen Fall eine Handlung oder ein mentaler Akt vor, wenn ich entscheiden kann, ob ich eine Einstellung annehme/ablege etc. oder nicht, und ich die Fähigkeit habe, diese Entscheidung umzusetzen.

Es stellt sich die Frage, wie ich entscheiden kann, ob ich bei jedem Fassen z.B. einer Intention mich entscheiden muss und ob ich diese Entscheidung bewusst treffen. Letzteres ist sicherlich nicht der Fall. Ich möchte am Beispiel der Überzeugung versus Akzeptanz der Wahrheit einer Proposition die Unterscheidung zwischen einer Entscheidung, die einen mentalen Akt darstellt, und des Zustandekommens eines Zustandes ohne Akt deutlich machen.

Wie oben erwähnt, liegt im Falle der Akzeptanz eine Besonderheit vor. Wenn ein Subjekt akzeptiert, dass p , ist das eine propositionale Einstellung. Ob ein Subjekt etwas akzeptiert oder nicht, liegt in der Entscheidungsgewalt des Subjektes. Gleichzeitig mit der Entscheidung für die Akzeptanz liegt im Subjekt der mentale Zustand vor, dass das Subjekt eine Proposition als wahr akzeptiert hat. So schreibt Cohen:

„Acceptance is thus a mental act (as what was called ‘judgement’ often used to be), or a pattern, system or policy of mental action.“ [45, S.368]

Es gibt propositionale Einstellungen, die in der Entscheidungsgewalt des Subjektes liegen und somit einen mentalen Akt kennzeichnen, jedoch gleichzeitig als das erfolgreiche Ausüben des Aktes beschrieben wird, welcher zu einem mentalen Zustand des Subjektes führt. Sie stellen somit ein Achievement-Ereignis dar, nach dessen Erreichen des ausgezeichneten Zustandes sich ein State-Ereignis anschließt. Im Gegensatz zur Durchführung liegt im Falle des Achievement-Ereignisses jedoch keine Dreiteiligkeit in Form des Nukleus von Moens und Steedman [99] vor. Im Falle der Akzep-

¹⁵Auf eine Unterscheidung von doxastischen Kontrollarten, die Nottelmann [102] trifft, gehe ich in Abschnitt 4.1 ein. Inhaltsgerichtete doxastische Kontrolle liegt vor, vgl. [102, S.90], wenn wir für eine bestimmte Proposition p (Inhalt) bewirken können, voluntativ eine doxastische Einstellung anzunehmen, d.h. zu glauben, dass p , abzulehnen zu glauben, dass p , (disbelief) oder uns willentlich enthalten.

3. Handlungen oder die Frage „Actions as Events“

tanz spiegelt das Ereignis ‘dass α akzeptiert, dass p ’ sowohl einen Zustand (‘ α hat akzeptiert, dass p ’) als auch ein Ereignis, dass das Subjekt p akzeptiert, wider.¹⁶ Im Gegensatz zur Überzeugung ist jedoch das Ereignis, das dieser Sachverhalt charakterisiert, ein mentaler Akt. Einer Akzeptanz geht voraus, dass die Proposition, die es zu akzeptieren gilt, dem Subjekt gewahr wird. Es trifft aufgrund der ihm vorliegenden Evidenzen, wobei ich als Evidenzen auch Wünsche und andere Einstellungen, Emotionen, Schmerzempfinden, etc. zähle, die ich unter den mentalen Gesamtzustand des Subjektes subsumiere, aktiv und unter seiner Kontrolle die Entscheidung, den mentalen Zustand ‘akzeptiert zu haben, dass p ’ eine geraume Zeit anzunehmen.

Cohen zeigt in [45, 46]¹⁷ an mehreren Beispielen, dass zwischen der Akzeptanz einer Proposition p als wahr und dem Überzeugtsein davon, dass p , ein Unterschied liegt. Er spricht in beiden Fällen nur von den mentalen Zuständen, die häufig zusammen auftreten. Meines Erachtens wird diese Unterscheidung vor allem dadurch deutlich, wenn man die Ereignisse betrachtet, die zu dem jeweiligen Zustand führen. Interessant sind die Fälle, in denen jemand akzeptiert, dass p , aber nicht glaubt, dass p , und Fälle, in denen jemand glaubt aber nicht akzeptiert, dass p .

Im ersten Fall stelle man sich einen Richter vor, der aufgrund der entlastenden Beweise, die der Anwalt des Angeklagten präsentiert, akzeptiert, dass dieser unschuldig ist. Aber seit Anbeginn der Verhandlung glaubt der Richter, dass er schuldig ist, z.B. weil der Angeklagte ihm unsympathisch ist oder vielleicht jemandem ähnlich sieht, der schwere Verbrechen begangen hat, oder der Angeklagte bereits einschlägig vorbestraft ist. Ob die akzeptierte oder geglaubte Proposition handlungswirksam ist, sei dahingestellt, dies hängt von weiteren Einstellungen des Richters ab.¹⁸ Als Fälle von Überzeugt sein, dass p , aber nicht akzeptieren, dass p , benennt Cohen in [46, S.17] z.B. Überzeugungen von Wissenschaftlern, dass bestimmte Naturgesetze aufgrund von Annahmen gelten, die sie aufgestellt haben, die jedoch nicht als wahr akzeptiert werden, solange man sie nicht bewiesen hat. Ein anderes Beispiel soll zeigen, dass es manchmal notwendig ist, die Wahrheit einer Proposition zu akzeptieren, selbst wenn man keine Überzeugung aufgrund seiner Evidenzen bilden kann. Folgendes Beispiel stammt von Alston [4, S.126] und zeigt, dass Überzeugung nicht notwendig für

¹⁶Vendler beschreibt, dass es für die Kategorie Achievement charakteristisch ist, dass es keinen Bedeutungsunterschied zwischen der Benutzung des Perfektes anstatt des Präsens des Verbs gibt. Die Proposition ‘ α akzeptiert, dass p ’ impliziert ‘ α hat akzeptiert, dass p ’ bereits bzw. in unmittelbarer Zukunft wird es das, vgl. [151, S.147].

¹⁷Cohen ist nicht der einzige in dieser Diskussion. Unterschiede zwischen diesen beiden Einstellungen werden u.a. auch von Audi, Bratman, Nottelmann, Tuomela in [9, 26, 102, 143] u.v.m. herausgestellt.

¹⁸Das Beispiel lehnt sich an das Beispiel von Cohen [45, S.370] an.

3.3. Mentale Zustände und mentale Ereignisse

Akzeptanz ist.

Ein Befehlshaber soll entscheiden, ob er die feindliche Armee angreift oder nicht. Er würde seine Entscheidung aufgrund der Stärke und Anzahl des Gegners treffen, doch er hat anhand der ihm zur Verfügung stehenden Evidenzen keinerlei Überzeugung gebildet, ob er schwächer oder stärker ist. Entscheidet er sich für den Angriff, hat er angenommen, dass der Gegner schwächer ist. Würde er nicht als wahr akzeptieren, dass er schwächer ist, hätte er nicht angegriffen. Analoges gilt, wenn er nicht angreift. Entscheidet er sich gegen den Angriff, hat er angenommen, dass der Gegner nicht schwächer ist. Dennoch muss er eine Entscheidung fällen angreifen oder nicht angreifen. Er wird eine der beiden Proposition akzeptieren, ohne von einer der beiden überzeugt zu sein.

Obwohl Cohen [46] mit Beispielen, wie dem des Richters, zeigt, dass Überzeugung und Akzeptanz unterschiedliche Einstellungen sind, stellt er fest, dass es normal ist zu glauben, was wir für wahr als gegeben annehmen, und zu akzeptieren, wovon wir überzeugt sind. Auf der anderen Seite resümiert er:

„More generally we can say that, though a person who accepts nothing that he believes is intellectually self-paralysed, a person who habitually accepts everything that he believes is recklessly uncritical.“ [46, S.19]

Dass eine Überzeugung, dass p , einen Grund zu akzeptieren, dass p wahr ist, darstellt, scheint Cohen natürlich. Alles unkritisch zu hinterfragen, was man glaubt und einfach als wahr zu akzeptieren, scheint ihm nicht vernünftig.

Umgekehrt scheint es unplausibel zu sein, dass die Akzeptanz, dass p , einen Grund für die Überzeugung, dass p , darstellt. Die Akzeptanz, als Grund für das Glauben der Proposition anzugeben, scheint irrational. Vielmehr würde man davon ausgehen, dass das Überzeugtsein von einer Proposition dazu führt, die Wahrheit dieser zu akzeptieren. Indem Cohen Fälle nennt, in denen dieser Grund nicht hinreichend ist, sich für die Akzeptanz von p zu entscheiden, ist es möglich zu glauben, dass p , ohne zu akzeptieren, dass p wahr ist. Auf der anderen Seite kann die Akzeptanz, dass p , ursächlich dazu beitragen, dass das Subjekt überzeugt ist, dass p , [46, S.18]. Der einmalige oder wiederholte Akt der Akzeptanz kann dazu führen, dass der Vorgang der Überzeugungsbildung, dass p , abläuft, ohne dass sich der Akteur dafür entschieden hat. Cohen lässt die Möglichkeit offen, dass, wenn jemand immer wieder eine Proposition als wahr annimmt und sie akzeptiert, es ursächlich dafür sein kann, dass er sie glaubt. Somit kann allein die Akzeptanz zum Bilden einer Überzeugung führen, ohne

3. Handlungen oder die Frage „Actions as Events“

dass das Subjekt weitere Gründe hat. Diese Überzeugungsbildung stellt jedoch keinen mentalen Akt dar, da der mentale Akt, der die Überzeugung verursacht und bei dem das Subjekt direkte inhaltsgerichtete Kontrolle ausübt, die wiederholte Akzeptanz und nicht die Überzeugungsbildung ist.

Dies ist ein Grund zwischen mentalen Akten, mentalen Vorgängen und mentalen Zuständen zu unterscheiden. Zwei mentale Zustände liegen vor oder liegen jeweils nicht vor. Eine Zustandsänderung kann auf zweierlei Arten bewerkstelligt werden, einmal passiv im Subjekt oder aktiv von dem Subjekt selbst initiiert. Während im Falle, dass das Subjekt akzeptiert hat, dass p , Cohen davon spricht, dass durch die Akzeptanz das Subjekt (mentaler Akt) ohne sein aktives Dazutun das Vorliegen des Zustandes der Überzeugung erreicht (mentaler Vorgang), kann nach Cohen die Überzeugung als Grund dazubetragen (Vorliegen des mentalen Zustandes), dass das Subjekt sich entscheidet zu akzeptieren, dass p (mentaler Akt). Beide mentalen Vorkommnisse haben mentale Zustände zur Folge. Der Unterschied muss nicht im gebildeten Zustand liegen, sondern liegt bereits in der Art und Weise begründet, wie der Zustand zustande kam, entweder initiiert durch einen Akt oder durch einen mentalen Vorgang.

Im Falle einer Intention liegt die Sache etwas anders. Der mit ‘ α intendiert, dass p ’ beschriebene Sachverhalt kann ebenfalls ein Ereignis aber auch einen mentalen Zustand charakterisieren, wobei das Ereignis beschrieben werden kann durch: Der Zustand, der durch den Sachverhalt ‘das Haben der Intention’ charakterisiert wird, liegt eine gewisse Zeit vor (State). Angenommen man stimme zu, dass Intentionen unter der Entscheidungsgewalt des Subjektes liegen. Dann wird durch den mit ‘ α intendiert, dass p ’ beschriebenen Sachverhalt ebenfalls ein Ereignis charakterisiert, dass das Subjekt die Entscheidung, diese Intention zu haben, getroffen hat. Auch wenn ‘intendieren’ kein reines Achievement-Verb, wie z.B. akzeptieren, ist, beinhaltet das Bestehen eines Sachverhaltes ‘ α intendiert, dass p ’, dass eine Entscheidung, die Einstellung anzunehmen, stattgefunden hat (Accomplishment), die Einstellung angenommen wurde (Achievement) und dies zum Eintreten eines mentalen Zustandes führte, der dann fortbesteht (State).

Auf sprachlicher Ebene kann ein Unterschied zwischen den Teilereignissen gemacht werden, indem vom Haben einer Intention und dem Fassen einer Intention gesprochen wird, d.h. zwischen dem Vorliegen des mentalen Zustandes und dem Durchführen des mentalen Aktes. Die eben genannten drei Teilereignisse könnte man dann auf zwei Arten zu einem Ereignis zusammenfassen: Entweder man legt die Betonung auf den Sachverhalt, dass der mentale Zustand eine Weile besteht. Dann wäre es ein State-Ereignis. Oder man legt die Betonung auf das Fassen der Intention, das wie-

3.3. Mentale Zustände und mentale Ereignisse

Einstellung ¹⁹	Zustände	Ereignisse	
		Vorgang	Akt
Überzeugung	Haben	Formen, Ablegen, Beibehalten	—
Wunsch	Haben	Formen, Annehmen, Ablegen, Beibehalten können sowohl Vorgang als auch Akt sein.	
Absicht	Haben	Beibehalten, Ablegen	Formen, Ablegen, Beibehalten
Akzeptanz	Subjekt hat akzeptiert	—	Subjekt akzeptiert

Tabelle 3.2: Einstellungen als mentale Zustände und Ereignisse

derum alle drei als Teilereignis enthält. Dann ist es ein Achievement-Ereignis. Die Verwirrung darüber, ob Intentionen Handlungen oder Zustände sind, vgl. [142], resultiert meines Erachtens aus der Missachtung des Unterschiedes, eine Intention zu haben bzw. eine Intention zu fassen. Aufgrund dessen, dass ‘ α intendiert, dass p ’ ausdrückt, dass ein Akt vollzogen wurde, nämlich der der Bildung der Intention, und gleichzeitig angibt, dass im Subjekt α der mentale Zustand, dass ‘ α intendiert, dass p ’ vorliegt, wird im weiteren Verlauf in Bezug auf Intentionen sowohl von Akten als auch von Zuständen gesprochen. Je nach Zusammenhang ist das Subjekt in einem bestimmten Zustand oder aber das Zustandekommen des Zustandes gemeint.

Weitere propositionale Einstellungen, die einen mentalen Akt voraussetzen, welche das Auftreten eines mentalen Zustandes nach sich ziehen, sind z.B. Urteilen und Ablehnen. Es gibt jedoch Ereignisse, die mentale Akte darstellen, ohne gleichzeitig das Eintreten eines mentalen Zustandes zur Folge zu haben, z.B. Rechnen und Folgern. Wenn mentale Akte Ereignisse sind, über die wir Kontrolle in der Form haben, dass wir über ihr Eintreten entscheiden, impliziert das, dass diese Form der Entscheidung als Basishandlung verstanden wird. Jeder Handlung liegt eine Entscheidung zur Handlung zugrunde. Eine Entscheidung oder ‘ α entscheidet, p zu tun’ verstehe ich, wie im

¹⁹Es kann sowohl mentale Akte als auch mentale Zustände geben, die keine propositionalen Einstellungen sind, z.B. eine bestimmte Wahrnehmungserfahrung bzw. Emotionen.

3. Handlungen oder die Frage „Actions as Events“

Folgenden erläutert. Einer Entscheidung muss keine bewusste Deliberation vorausgehen. Eine Entscheidung trifft das Subjekt anhand der bereits vorhandenen mentalen Zustände, die zum Zeitpunkt der Entscheidung im Subjekt vorliegen. Die Zustände können aus vorangegangenen mentalen Akten resultieren. Es ist jedoch nicht notwendigerweise der Fall, dass das Subjekt über seine Wünsche, Motive, Absichten reflektiert oder anderweitige Kenntnisse hat, um seine Entscheidung zu treffen.

Eine Änderung eines vorliegenden mentalen Zustandes ist ein mentales Ereignis. Das Vorliegen eines mentalen Zustandes ist ein mentales Ereignis. Jedoch ist der Zustand unabhängig davon, ob es ein Zustand durch Akzeptanz oder Überzeugung ist, kein Ereignis. Auch auf mentaler Ebene verhalten sich Ereignis und Zustand analog. Eine Nichtänderung eines Zustandes kann ebenfalls ein mentales Ereignis sein. Dann ist aber die Nichtänderung und nicht der Zustand selbst das Ereignis. Ein mentaler Akt ist ein solches aktives mentales Ereignis. Das Vorliegen einer Überzeugung ist ein mentales Ereignis, jedoch kein mentaler Akt. Das Bewusstwerden, diese Überzeugung zu haben, oder der Vorgang, dass die Überzeugung gebildet wird, sind ebenfalls mentale Ereignisse. Ob ein mentales Ereignis tatsächlich eine Handlung, oder besser formuliert ein mentaler Akt ist, bzw. die Frage, welche Bedingungen vorliegen müssen, so dass ein mentales Ereignis ein Akt ist, sind nicht eindeutig geklärt.²⁰

Es ist in der Literatur oft diskutiert worden, ob es mentale Handlungen gibt und inwiefern mentale Aktivitäten oder Phänomene überhaupt Handlungen bzw. Akte eines Akteurs sein können. Was auffällt, ist, dass dabei oft nicht zwischen Zuständen und Ereignissen (Zustandekommen des mentalen Zustandes) unterschieden wird. Ich finde diese Unterscheidung notwendig, um Argumente zu widerlegen, die sich gegen die Auffassung richten, dass mentale Phänomene als Handlungen gelten. Ryle argumentiert, dass wir mentale Akte nicht mit Bedacht ausführen können. Denn diese Forderung hätte einen infiniten Regress zur Folge.

„That someone has been carefully watching a bird on the lawn does not entail that he has also been metaphorically ‘watching’ his watching; [...] Doing something with heed does not consist in coupling an executive performance with a piece of theorising, investigating, scrutinising or ‘cognising’: or else doing anything with heed would involve doing an infinite number of things with heed.“ [117, S.137]

Proust [112] versucht sich an einer positiven Definition, dessen was eine mentaler Akt

²⁰Strawson hat einen Vorschlag gemacht, wann etwas ein „mental act“ ist. Er wurde jedoch von Buckareff kritisiert, dass sein Ansatz zu restriktiv ist, [36, 138].

3.3. Mentale Zustände und mentale Ereignisse

im Gegensatz zu einer mentalen Operation (einem mentalen Phänomen) ist, um Ryles infiniten Regress und dem von ihr benannten ähnlichem Problem des Zuganges zu entgehen. Als Beispiel für das Zugangsproblem nimmt sie den Begriff „attention“. Um auf ein Ereignis aufmerksam zu werden, egal ob es ein mentales oder physisches Ereignis ist, bedarf es des Ereignisses. Im Falle des Aufmerksam-Werdens bedeutet es, dass, falls das Subjekt Kontrolle darüber ausüben will und es als ein Akt des Subjektes zählt, dass es aufmerksam wird, es erst einmal darauf aufmerksam werden muss. Das Ereignis 'aufmerksam zu sein' bedarf bereits des Ereignisses 'aufmerksam zu sein'.

„The particular form it takes in attention is the following: how can I set myself to perceive some external or internal property, if I cannot yet represent it and thus appreciate how relevant to me it will be? What distinguishes this puzzle from Ryle's infinite regress, is that in the puzzle, the content of attention seems to be a precondition for accessing it, whereas in Ryle's objection, the act supposes that you have already performed it (as an act token) (access puzzle) or that some previous act of the same type was performed (Ryle's puzzle).“ [112, S.113].

Wie lassen sich die beiden Probleme lösen? Als erstes hält Proust fest, dass man den Verstand eines Subjektes nicht als sein Werkzeug ansehen kann, das er benutzt und wieder zur Seite legt, sondern es ist das Subjekt selbst, welches denkt. Ein Denker kontrolliert seine mentalen Vorgänge nicht hin und wieder, wie er gern möchte, sondern für Proust ist Folgendes der Fall:

„[A] mental action is one that exercises control on antecedently uncontrolled mental processing in order to modify the output of that mental processing.“ [112, S.113]

In gewisser Weise trifft sie die Unterscheidung, die ich oben angeführt habe. Sie trennt zwischen den Prozessen (mentalen Vorgang bzw. mentalen Akt) und dem Output (mentalen Zustand) eines mentalen Prozesses. In meiner Bezeichnung stellt ein mentaler Prozess ein mentales Ereignis dar. Nach dem Eintreten des Ereignisses liegt ein Zustand vor, den man mit Prousts Output des Prozesses gleichsetzen kann. Die meisten Prozesse bei Proust laufen unbewusst ab. Ein Output verursacht den nächsten Output. Nach Proust sind die meisten mentalen Operationen unbewusst, während ihrer Meinung nach mentale Akte bewusst ablaufen [112, S.114].²¹ Proust betont, dass im Falle von mentalen Akten das Subjekt auf die mentalen Operationen einwirken kann,

²¹Proust beschreibt „mental operations“ analog zu dem von mir verwendeten Begriffes des mentalen Vor-

3. Handlungen oder die Frage „Actions as Events“

indem es bestimmte Intentionen hat, etwas zustande zu bringen. Aber genausowenig wie wir beim Fangen eines Balles wissen, welche Muskeln wir wie anspannen, wissen wir, wie wir den Ball fangen, und können uns mit ein bißchen Training auf unser Können verlassen. Dieselbe Form von Kontrolle bei solch physischen Handlungen postuliert sie für mentale Akte. Es erfordert Metakognition zu wissen, welche Operationen uns helfen, welche Ziele zu erreichen. Nicht erforderlich ist dabei zu wissen, wie die Operationen genau ablaufen.

Proust schlägt fünf hinreichende und notwendige Bedingungen dafür vor, wann ein Subjekt einen mentalen Akt A vollzieht, so dass X die mentale Eigenschaft F zur Folge hat [112, S.121]. Übersetzt in meine Terminologie lauten sie wie folgt:

- (1) zum Zeitpunkt t geschehen Vorgänge O im Subjekt α , die zum Zustand M führen;
- (2) α hat in t eine Proeinstellung zum Zustand F ;
- (3) α weiß, wie es den Zustand F erreicht, wenn die Vorgänge O im Zustand M vorliegen;
- (4+5) α 's Proeinstellung und Know How verursachen unter regulären Umständen ohne Devianz, dass α den Zustand F erreicht.

Auf Probleme und Vorteile dieser Charakterisierung möchte ich kurz eingehen. Dafür werde ich Kriterien nennen, die eine Handlung (mental und nicht-mental) erfüllt. Die Kriterien, ob etwas eine Handlung bzw. ein mentaler Akt ist oder nicht, sollten sich für mentale und nicht-mentale Ereignisse nicht unterscheiden. Auch wenn es Unterschiede zwischen mentalen Akten und nicht-mental Handlungen gibt, sollten beide unter einer Form der voluntativen Kontrolle des Akteurs und auf ihn zurückführbar sein. Generell wird davon ausgegangen, dass eine Handlung eine Einwirkung des Subjekts

ganges. Vier Eigenschaften sind notwendig, die auf die mentalen Vorkommnisse zutreffen, die ich als Vorgänge bezeichne. Eine Operation transformiert einen Input in einen Output und dies in einer determinierten Art und Weise, jedoch ohne dass es nötig ist, dass bestimmte Motive oder Ziele für den Output vorliegen müssen. Sie sind nicht auf den Output gerichtet, sondern sie werden durch ihren Input und anderen vorliegenden, nicht unter der zu dem Zeitpunkt ausübaren Kontrolle stehenden Vorgängen und Zuständen, wie z.B. Wahrnehmungen, Überzeugungen, Wünsche, verursacht. Sie können nicht deliberativ ablaufen und sind dem Subjekt (möglicherweise) nicht bewusst, vgl. [112, S.114]. Da ich über die Darstellung des mentalen Gehaltes nicht an dieser Stelle diskutieren möchte, lasse ich die Einlassung Proust beiseite, dass Operationen bzw. mentale Zustände einen Gehalt in Form einer Eigenschaftszuschreibung haben „M having N“ und bezeichne einfach den mentalen Zustand (Output der Operation), der einen bestimmten, wie auch immer darstellbaren Gehalt hat.

3.3. Mentale Zustände und mentale Ereignisse

auf das Bestehen eines bestimmten Sachverhaltes beinhaltet. Dabei ist Voraussetzung, dass das Subjekt in irgendeiner Form Kontrolle über die Handlung hat und sie dem Subjekt als Urheber zuschreibbar ist. Das Subjekt hat die Handlung aktiv mit verursacht, so dass das Subjekt das Ereignis nicht nur über sich geschehen lässt (sondern es sich dafür entschied).²²

Das Subjekt befindet sich zu einem gegebenen Zeitpunkt in einem mentalen Gesamtzustand. Es liegen mentale Zustände vor, die propositionale Einstellungen darstellen, wie z.B. Volitionen, Überzeugungen, bereits bestehende Ziele etc., aber auch Zustände mit nicht notwendigerweise propositionalem Gehalt, wie Emotionen, Wahrnehmungen, etc. In diesem Zustand treten z.B. durch konfligierende Wünsche und Überzeugungen oder Wünsche, die mit Überzeugungen und Emotionen konfligieren, Spannungen auf und führen zu mentalen Ereignissen. Ein solches mentales Ereignis kann z.B. das Bilden einer Intention sein (mentaler Akt), um diese Konflikte abzubauen, und führt zu einer Änderung des mentalen Gesamtzustandes des Subjektes. Es sind auch andere mentale Ereignisse denkbar, die keinerlei Kontrolle des Subjektes voraussetzen, wie das Bilden bzw. Revidieren von Überzeugungen, die den Zustand des Subjektes entspannen. Weder diese mentalen Ereignisse noch die Zustände müssen dem Subjekt bewusst, zugänglich oder transparent sein. Es ist möglich, aber nicht notwendigerweise der Fall, dass es diese Vorgänge und Zustände wahrnimmt oder Überzeugungen darüber ausbildet. Im Gegenteil zeigt doch Ryles Argument, dass ein Bewusstwerden ebenfalls ein mentaler Vorgang ist. Dass dem Subjekt ein mentaler Zustand bewusst ist, ist selbst wieder ein mentaler Zustand, der jedoch nur durch einen mentalen Vorgang zustande gekommen ist. Demzufolge muss nicht jedem mentalen Akt inhärent sein, dass das Subjekt ihn bewusst erlebt.

Anscombe hat eine Eigenschaft beschrieben, die bei ihr für eine absichtliche Handlung hinreichend ist, nämlich „ohne Beobachtung zu wissen“, die Handlung vollzogen zu haben [7]. Meines Erachtens ist diese Eigenschaft in gewisser Weise für mentale Akte gegeben. Das Subjekt weiß nicht notwendigerweise, dass es *p* intendiert. Es muss ihm nicht bewusst sein. Richtet es jedoch seine Aufmerksamkeit auf seine Ein-

²²Mit dem Begriff des Urhebers lehne ich mich an Pauens erste Minimalforderung des Urheberprinzips an, die eine freie Handlung erfordert und beschreibt, dass die Handlung einem Akteur zuschreibbar sein muss [104]. Allerdings verzichte ich auf die zweite Forderung Pauens, das Autonomieprinzip, dass es in gewisserweise unabhängig von den äußeren Umständen sein muss. Dies ist auch nicht nötig, da ich keine Kriterien für Handlungsfreiheit erörtern möchte. Diese Forderung der Urheberschaft ist dennoch schwächer als die von Davidson verlangte Beschreibung, dass es intentional für den Akteur ist [48, Essay 3], die Handlung durchzuführen, bzw. auch schwächer, als die Aussage, dass Wünsche und Überzeugungen Ursache unserer Handlungen sind [48, Essay 1].

3. Handlungen oder die Frage „Actions as Events“

stellungen zur Proposition p , wird ihm ohne Weiteres gewahr, dass es das Fassen der Absicht (mentaler Akt) mit dem Folgezustand das Haben der Absicht (mentaler Zustand) vollführt hat. Mentale Akte sind dem Subjekt ohne weitere Vorgänge, als dass das Subjekt seine Aufmerksamkeit auf den Folgezustand des Aktes lenkt, wissbar bzw. transparent. Dies gilt für viele mentale Zustände nicht, die Folgezustand eines mentalen Vorganges sind. Eine Überzeugung ist dem Subjekt nicht immer transparent, selbst wenn es seine Aufmerksamkeit in Selbstreflexion auf die doxastische Einstellung richtet, die es hat, ist sie ihm nicht transparent. Manche Überzeugungen entdeckt das Subjekt erst bei sich anhand der Handlungen, die es vollzieht. Dies unterscheidet eine Überzeugung wiederum von der Akzeptanz einer Proposition als wahr.

Allerdings ist zumeist nicht der mentale Akt zugänglich bzw. transparent, sondern der Folgezustand des mentalen Aktes, da der Akt bereits stattgefunden hat. Dies unterscheidet mentale Akte nicht von Handlungen. Oft werden physische, nach außen sichtbare Handlungen uns anhand des Zustandes gewahr, den die Handlung herbeiführte. Die Handlung des Armhebens wird uns gewahr, wenn der Arm gehoben ist. Wenn ich dafür Sorge, dass p , erkenne ich mein Sorge-Tragen-Dafür daran, dass danach p wahr ist. Ein Beispiel ist Autofahren. Niemand wird während des Autofahrens alle seine Handlungen, die ablaufen, bewusst erleben. Er erkennt, dass er den Fuss auf das Gaspedal bzw. Bremspedal setzt, daran, dass er beschleunigt bzw. bremst.

Es ist somit nicht nötig, darin möchte ich Proust in Punkt (2) widersprechen, dass der Ausgangszustand einer Handlung vom Subjekt tatsächlich in einer Proeinstellung des Subjektes vor der Handlung wiederzufinden ist. Eine Proeinstellung bei Proust ist eine Einstellung zu einem Ziel und einer Überzeugung, dass es das richtige ist, etwas zu tun, um das Ziel zu erreichen. Die Proeinstellung ist eine motivationale und konative Komponente des mentalen Aktes [112, S.121]. Vor Beginn der Handlung ist ein mentaler Vorgang vonnöten, der die Proeinstellung zur Folge hätte. Dies ist meines Erachtens unnötig. Für nicht-mentale Handlungen scheint dies offensichtlich.

Prousts Konzept von Proeinstellung ist etwas speziell gefasst und erinnert an eine Handlungsmotivation [129] bzw. einen primären Grund zu handeln [48, Essay 1]. Da der Begriff *Proeinstellung* von mehrere Autoren verschieden gebraucht wird, möchte ich kurz auf diesen Begriff eingehen. Ich werde ‘Proeinstellung’ anders als Proust fassen jedoch so, dass „meine“ Proeinstellung notwendig ist, um eine Proeinstellung in Prousts Sinn zu haben. Eine Proeinstellung ist eine Einstellung eines Akteurs zu p (Proposition, Sachverhalt, mentaler Repräsentation u.ä.) in der Passrichtung Welt-Geist, die positiv motivational zu p und konativ ist. Typische Beispiele für Proeinstellungen sind Wünsche, Hoffnungen und das Ziel zu haben, dass p . ‘Beseitigen wollen,

3.3. Mentale Zustände und mentale Ereignisse

dass p' wäre keine Proeinstellung, da sie zwar möglicherweise die richtige Passrichtung hat und konativ ist, jedoch nicht positiv motivational zu p ist.²³ Ob sie die richtige Passrichtung hat, ist diskutabel, da, falls das Beseitigen-Wollen von p impliziert, dass der Akteur eine kognitive Einstellung zu p (bspw. Erkennen, dass p) hat, würde diese Einstellung auch die andere Passrichtung einnehmen. Ob es möglich ist, dass mentale Zustände beide Passrichtungen haben können, so dass bestimmte Formen von Überzeugungen ebenfalls Proeinstellungen sein können, z.B. 'zu glauben, dass es gut ist, eine bestimmte Handlung auszuführen' werde ich in Abschnitt 4.1 diskutieren.

Jemand, der Fahrerflucht begeht, handelt möglicherweise, ohne eine Proeinstellung zu haben, Fahrerflucht zu begehen. Er wird sich seiner Tat gewahr, als er seine Aufmerksamkeit darauf richtet, die Proeinstellung zur Fahrerflucht jedoch fehlt. Seine Motivation für die Entscheidung zu fliehen war Angst und nicht der Wunsch zu fliehen.²⁴ So würde ich auch zwischen absichtlichen und unabsichtlichen Handlungen unterscheiden. Daher scheint es, dass Handlungen i.A. keiner Wünsche (Proeinstellungen) bedürfen, und somit ist diese Annahme für mentale Akte ebenfalls nicht notwendig. Dass Wünsche und Überzeugungen, wie Davidson meint [48, Essay 1], u.U. Handlungen verursachen können, gebe ich zu, aber sie sind nicht notwendigerweise Auslöser oder primäre Ursache. Auch für mentale Akte kann es möglich sein, dass weder eine Proeinstellung zu dem Akt bzw. noch zu dem Folgezustand des Aktes vorliegt bzw. dass Proeinstellungen sowohl zu der Handlung als auch zum Unterlassen der Handlung vorliegen.

Nehmen wir den Richter, der aufgrund der Beweise die Unschuld des Angeklagten annimmt. Er muss keine Proeinstellung zu der Handlung der Akzeptanz oder dem bei ihm daraus folgenden Zustand, dass er die Unschuld akzeptiert hat, besitzen. Er kann seinem Verhalten ablehnend gegenüber stehen. Er kann auch die Unschuld des Angeklagten ablehnen. Seiner Handlung zu akzeptieren, d.h. zu bewirken, dass er in den Zustand der Akzeptanz kommt, ruht keine Proeinstellung zum Endzustand inne. Dennoch gibt es Wünsche, die als Grund für diese Entscheidung dienen, z.B. der Wunsch, seinem Amt gerecht zu werden.

Punkte (3-5) der Charakterisierung von mentalen Akten, wie Proust sie vorschlägt, können wie folgt zusammengefasst werden. Das Vermögen bzw. die Fähigkeiten, die das Subjekt hat (Know How), zusammen mit der Proeinstellung bewirken auf kau-

²³Die Unterscheidung von positiv konativ und negativ konativ zu treffen geht auf einen Hinweis von Heinrich Wansing zurück.

²⁴Auf die Motivation zu einer Handlung und dem Vorliegen von Wunsch und Absicht gehe ich in Kapitel 4 detaillierter ein.

3. Handlungen oder die Frage „Actions as Events“

sal nicht-deviante Weise, dass der erwünschte Zustand eintritt. Dies zählt für Proust bereits als Handlung. Dies scheint mir jedoch zu wenig. Die Kontrolle, die die Akteure über das Know-How und die Proeinstellung haben, ist zu schwach, um eine Handlung zu postulieren. Über Proeinstellungen, z.B. in Form von Wünschen, hat ein Akteur nicht notwendigerweise Kontrolle bzw. hat er sie i.A. nicht ausgeübt. Wäre das Erlangen der Proeinstellung ein Teil des Aktes, würde wiederum Ryles infinites Regress Argument greifen, also muss die Möglichkeit der Kontrolle im Punkt (3) liegen. Proust formuliert allerdings, dass das Subjekt das Know How nur haben muss und dieses Know How etwas verursacht.

Eine Fähigkeit mit einer Proeinstellung allein führt noch nicht zur Handlung. Ich habe das Know How von diesem Stuhl aufzustehen, möchte es, um ins Kino zu gehen, tue es aber dennoch nicht, weil ich u.U. andere Proeinstellungen habe, die mit dem erstgenannten Wunsch konfigieren. Das Know How, wie ich aufstehe, ist in diesem Fall ursächlich dafür, dass ich auch weiß, was ich tun muss, um sitzen zu bleiben. Wenn nun beide Proeinstellungen ‘es ist richtig aufzustehen’ und ‘es ist richtig sitzen zu bleiben’ im Subjekt auftreten, müssten beide Ereignisse eintreten. Somit ist die Fähigkeit zum Sitzenbleiben bzw. zum Aufstehen zusammengenommen mit der Proeinstellung nicht ausreichend, um eine Handlung, Aufstehen oder Sitzenbleiben, tatsächlich zu verursachen. Eine Fähigkeit und eine Proeinstellung zu haben, führt nicht ohne Weiteres zum Ausüben einer Handlung.

Der mentale Gesamtzustand enthält möglicherweise sowohl Proeinstellungen als auch Kontraeinstellungen zu einer Proposition p (Kontraeinstellung im Sinne, dass das Subjekt eine Einstellung zu p hat, die Welt-Geist gerichtet ist und die negativ motivational zu p und konativ ist, beispielsweise zu wünschen, dass p nicht der Fall ist). Sicherlich ist das Besitzen der Fähigkeit ursächlich relevant, wenn das Subjekt sie ausübt, aber ebenso ursächlich relevant, wenn das Subjekt es unterlässt, sie auszuüben. Im Falle einer Handlung scheint es mir, sogar notwendig zu sein, dass ein Subjekt nicht nur die Fähigkeit haben sollte, die Handlung auszuführen, sondern auch die Fähigkeit, die Handlung nicht auszuführen. Somit würden Proeinstellung und Fähigkeit zum Handeln als auch Kontraeinstellung und dieselbe Fähigkeit zum Unterlassen der Handlung führen. Das Subjekt kann jedoch nicht beides tun. Der Gesamtzustand müsste also auf der einen Seite mehr pro als kontra sein. Wäre das Abwägen ein Vorgang ohne Einwirken des Subjektes, würde dem Subjekt jede inhaltsbezogene Kontrolle bei der Handlung abgesprochen werden. Also muss es eine Entscheidung des Subjektes sein, ob der mentale Gesamtzustand dafür oder dagegen spricht und damit ein Ausüben der Fähigkeit erfolgt oder nicht.

3.3. Mentale Zustände und mentale Ereignisse

Die Ausübung der Fähigkeit zur Durchführung der Handlung ermöglicht es nicht nur, dem Akteur eine gewisse Form von Kontrolle auf der einen Seite zu zusprechen, sondern auf der anderen Seite ist es zudem möglich, damit eine Negativ-Bedingung für die Zuschreibbarkeit der Handlung aufzustellen. Hat ein Akteur nicht die Fähigkeit, die er zum Durchführen der Handlung benötigt, ist sie ihm auch nicht zuschreibbar.

Die Fähigkeit zu haben, impliziert dabei nicht die Möglichkeit, die Handlung auszuüben. Wenn jemand an den Stuhl gefesselt ist, verliert er nicht die Fähigkeit zu gehen. Es ist ihm lediglich in dem Moment nicht möglich, sie auszuüben. Die Fähigkeit ist notwendig, um eine Handlung erfolgreich auszuführen, aber nicht hinreichend. Wenn das Subjekt sich für etwas entschieden hat, ist es nicht sofort realisiert. Allein aus der Fähigkeit resultiert bei einer Handlung noch nicht die Möglichkeit.

Auf mentale Akte lässt sich dies meiner Ansicht nach übertragen. Die Fähigkeit 'eine Rechenaufgabe zu lösen' impliziert auch die Fähigkeit, dies zu unterlassen. Oft wünscht ein Subjekt, die Rechenaufgabe zu lösen, u.U. ist es dazu auch in der Lage und vollzieht, die Rechenaufgabe zu lösen, dennoch nicht. Nicht nur weil das Subjekt die Fähigkeit zwar hat, sich aber dennoch hin und wieder verrechnen kann, sondern vor allem deswegen, weil die Fähigkeit, die Rechenaufgabe zu lösen, und die Proeinstellung nicht zum Bearbeiten der Aufgabe überhaupt führen.²⁵ Prousts Konzept scheint mir daher nicht hinreichend.²⁶

Am Anfang ihres Aufsatzes benennt Proust allgemein, wie man testen kann, ob ein Verb eine (mentale) Handlung beschreibt oder nicht, den sogenannten Try-Test. Wenn man versuchen kann, etwas zu tun, scheint das Tun eine Handlung zu sein, die man versuchen kann. Man kann versuchen zu laufen, aber man kann nicht versuchen, in

²⁵Argumente dafür, dass es für ein Ereignis nicht reicht, als Handlung gezählt zu werden, weil es aufgrund mentaler Zustände und Vorgänge im Subjekt verursacht wird, findet man in [36, S.84] und [138, S.228].

²⁶Fara in [57] unterscheidet verschiedene Möglichkeiten des Fehlschlagens der Ausübung der Fähigkeit, obwohl man die Fähigkeit hat. Die Fähigkeit zu haben und nicht die Möglichkeit zu haben, wäre am Beispiel des Rechnens, die Fähigkeit zum Rechnen zu haben, aber keine Rechenaufgabe vor sich zu haben. Ein zweites Fehlschlagen ist die Fähigkeit und die Möglichkeit zu haben, eine Rechenaufgabe zu lösen und vor sich zu haben, aber es gar nicht zu versuchen. Als drittes Fehlschlagen der Ausübung der Fähigkeit benennt er, die Fähigkeit zu haben, die Aufgabe vor sich zu haben und es zu versuchen, und dennoch fehlzuschlagen. Dies bezeichnet Fara als sogenannte „Masked Abilities“ (sein Penny-Stapel Beispiel bzw. Austins Golfer). Diese Art des Fehlschlagens der Rechenfähigkeit wäre nach Proust kein mentaler Akt, da das Know-How und eine Proeinstellung zwar vorliegen, aber nicht ursächlich zu dem gewollten Zustand führen. Es liegt keine Proeinstellung zum Verrechnen vor. Dennoch wäre das Rechnen selbst ein mentaler Akt. Davidson stimmt zu, dass Verrechnen nicht absichtlich sein kann [48, Essay 3 S.45], aber es eine andere Beschreibung des Ereignisses des Verrechnens geben kann, die dieses Verrechnen als mentalen Akt klassifiziert.

3. Handlungen oder die Frage „Actions as Events“

Ohnmacht zu fallen [112, S. 108]. Ich stimme ihr zu. Man kann versuchen, etwas als wahr zu akzeptieren, aber man kann nicht versuchen, von etwas überzeugt zu sein, jedenfalls nicht direkt und inhaltsgerichtet.²⁷ Sie sagt deutlich, dass dies ein guter Test ist, ob eine durch ein Verb beschriebene Tätigkeit als eine Handlung aufgefasst werden kann. Sie sagt jedoch auch, dass dieser Test keine Garantie ist, dass im Einzelfall tatsächlich eine Handlung vorliegt. Man kann versuchen, sich ein Ziel zu setzen, eine Rechenaufgabe zu lösen oder auf eine Äußerung hin zuzuhören. Man kann nicht versuchen, einen Wunsch zu haben, sich unabsichtlich zu verrechnen oder den lauten Knall nebenan nicht zu hören. Ein Try-Test kann daher nur ein Negativ-Test sein. Falls man etwas nicht versuchen kann, kann es sich nicht um eine Handlung handeln. Allerdings ist der Test ungeeignet, um von einer Ereignisbeschreibung festzustellen, dass es eine Handlung ist.

Buckareff schlägt vor, dass Intentionen zusammen mit dem Try-Test einen Hinweis auf das Vorliegen einer Handlung geben.

„What is more, the causation of the action by the intention must play a guiding role, with the agent controlling his behaviour via the intention, remaining responsive to feedback in the execution of the action.“ [36, S. 85]

Das Problem, das ich sehe, wenn eine Intention vorliegt und ein Ereignis mit verursachen muss, um es als Handlung zu klassifizieren, ist, dass das Fassen einer Intention selbst ein mentaler Akt ist und dies Ryles infiniten Regress Argument bzw. Proust Zugangsproblem zum Opfer fällt. Wenn Davidson sagt:²⁸

„Action does require that what the agent does is intentional under some description, and this in turn requires, I think, that what the agent does is known to him under some description“ [48, Essay 3 S.50].

ergeben sich mehrere Fragen. Was meint er damit „intentional under some description“ zu sein? Ist das Subjekt in einem Zustand, in dem es die Intention bereits hat,

²⁷Ich kann mir wünschen, p nicht zu glauben, und solange versuchen, mir den Wunsch einzureden, dass ich es tatsächlich glaube, aber die Fähigkeit, die ich ausübe, wäre mich selbst zu belügen. Dies ist keine direkte Kontrolle über Überzeugungsbildung, sondern indirekte, indem ich Evidenz für p schaffe.

²⁸In dem Aufsatz *Agency* argumentiert Davidson dafür, dass Handlungen in irgendeiner Weise dem Subjekt als absichtlich zugeschrieben werden, und dafür, dass jede Handlung sich auf elementare Handlungen zurückführen lässt, wobei die elementaren Handlungen reine „bodily movements“ sind. Unter der Annahme, dass sich mentale und nicht-mentale Handlungsbedingungen nicht unterscheiden, habe ich seine Bedingungen aufgegriffen und möchte an ihnen testen, ob sie dazu dienen können, um mentale Ereignisse als Handlungen zu klassifizieren.

3.3. Mentale Zustände und mentale Ereignisse

und ist es sich dessen bewusst? Eine Intention hat man gebildet oder nicht und man ist sich dessen bewusst oder nicht.²⁹ Oder meint Davidson ein Konditional der Form: Wenn *X* der Fall gewesen wäre, hätte das Subjekt ebenso gehandelt, nur dabei auch mit der Absicht so zu handeln. Die Frage ist dann, was ist *X*? Wenn ein Konditional denkbar wäre, in der das Subjekt genau das nicht getan hätte (sogar obwohl die Absicht vorlag), wäre der Vorgang in der aktuellen Situation dann keine Handlung? Wenn hingegen ‘absichtlich unter einer Beschreibung zu sein’ erfordert, dass die Intention in irgendeiner Form im Subjekt vorliegt, dann müssen wir dem Akteur eine Intention zuschreiben, wobei, diese zu formen, kein mentaler Akt aufgrund des Regressargumentes sein kann. Wenn aber eine Intention gebildet wird und dies kein mentaler Akt ist, sondern ein Ereignis ist, das dem Subjekt geschieht, ohne die Fähigkeit einzuwirken, dann ist wiederum unklar, worin in Davidsons Ansatz die Möglichkeit der Kontrolle besteht. Damit wäre nicht festgelegt, worin der Unterschied zwischen ‘das geschieht dem Subjekt nur’ und ‘das Subjekt ist selbst aktiv’ liegt.

Mein Vorschlag ist: Wir ersetzen ‘absichtlich’ durch ‘vom Subjekt entschieden’, ohne eine möglicherweise vorherige Entscheidung für eine Proeinstellung (Wunsch, Absicht) des Subjektes zu fordern. Dieses Verständnis von Entscheidung beinhaltet keine Deliberation des Akteurs, wie Entscheidung z.B. in der Entscheidungstheorie beinhaltet. Auf die Frage, ob denn eine Proeinstellung zu einem Sachverhalt des Ereignisses oder des Folgezustandes des Ereignisses vorliegt, würde ich antworten, dass dies möglich, aber nicht notwendig ist. Wenn ein Subjekt etwas entscheidet, muss es weder dem Ereignis noch dem aus der Entscheidung resultierendem Zustand positiv gegenüber stehen. Seine Entscheidung hängt sicherlich ursächlich an dem mentalen Gesamtzustand, in dem sich das Subjekt befindet. Ob seine Entscheidung hinreicht, um den Akt zu vollziehen, liegt an den Fähigkeiten, die das Subjekt hat, und an den äußeren Möglichkeiten. Da jedoch bei Handlungen ein Subjekt sich innerhalb seiner Fähigkeiten entscheidet, etwas zu tun bzw. zu bewirken, kennzeichnet die Entscheidung ein Ereignis als Akt bzw. Handlung, sobald die Entscheidung, wenn die äußeren Umstände dies zulassen, umgesetzt ist. Diese Art von Entscheidung ist einem Akteur inhärent, wenn er Handlungen vollzieht.

²⁹Searle kritisiert „unter einer Beschreibung“ wie folgt. Unter einer Beschreibung ist ein Feuerwehrauto rot und unter einer anderen Beschreibung ist das Feuerwehrauto nicht rot. Was sagt mir das über das Feuerwehrauto und die Eigenschaft ‘rot zu sein’. Searle meint die Frage ist, was macht einen Handlung absichtlich unter der einen und nicht-absichtlich unter der anderen Beschreibung. Das ist der Kern der Handlung. [123, S.123f]. Eine Beschreibung impliziert für Searle, dass es ein sprachliches Problem ist. Ich denke jedoch nicht, dass Davidson etwas Derartiges im Sinn hat.

3. Handlungen oder die Frage „Actions as Events“

Daher kann untersucht werden, ob eine Verbalphrase eine Handlung beschreibt, indem ein ‘entscheidet sich zu’ davor gestellt wird. Man kann sich entscheiden zu laufen, aber nicht dazu entscheiden, in Ohnmacht zu fallen. Man kann sich entscheiden, etwas zu akzeptieren, aber nicht von etwas überzeugt zu sein. Diesen Test könnte man in Analogie zu dem auf Ryle zurückgehenden ‘Try-Test’ ‘Decide-Test’ nennen. Man kann sich entscheiden, eine Rechenaufgabe zu bearbeiten, aber sich nicht unabsichtlich zu verrechnen, etc.

In Analogie zu Davidsons Formulierung, die Beschreibungen verwendet, würde ich formulieren: Ein Subjekt ist Akteur eines Ereignisses, wenn eine Beschreibung, dass das Subjekt sich entschieden hat, Akteur des Ereignisses zu sein, nicht den das Ereignis charakterisierenden Sachverhalt widerspricht. Dennoch ist ‘absichtlich’ bzw. ‘entscheidet sich zu’ problematischer als ‘sorgen dafür, dass’. Beide Ausdrücke scheinen zu implizieren, dass eine Einstellung zum Sachverhalt vorliegt. So kann eine Handlung absichtlich als auch unabsichtlich sein, je nachdem welcher das Ereignis charakterisierende Sachverhalt gerade betrachtet wird. Wenn daraufhin geschlossen wird, dass das Ereignis keine Handlung ist, weil der „falsche“ Sachverhalt betrachtet wurde bzw. die Beschreibung, die dazuführte, dass es sich um eine Handlung handelt, nicht berücksichtigt wurde, wird dem Akteur das Ereignis möglicherweise nicht als Handlung zugeschrieben.

Wenn jede Beschreibung, jeder das Ereignis charakterisierende bestehende Sachverhalt p , derart ist, dass er durch ‘der Akteur sorgte dafür, dass p besteht’ paraphrasiert werden kann, deutet das darauf, dass eine Handlung des Akteurs vorliegt, vgl. die Stit-Paraphrase für Ereignisse S.68. Ob und welche Sachverhalte das Ereignis außerdem charakterisieren, ist für die Zuschreibung des Akteurs zu dem Ereignis nicht wesentlich. Es steht fest: Es hat eine Handlung stattgefunden, der ein Ereignis zugrunde liegt, und der Akteur hat dafür gesorgt, dass die Sachverhalte bestehen. Es gibt nach eben geführten Diskussionen eine Beschreibung, dass es eine Entscheidung des Akteurs war, für einen Sachverhalt zu sorgen, der das Ereignis charakterisiert.

Im Weiteren nehme ich daher an, dass eine Handlung eines Subjektes α stattgefunden hat, wenn im Gesamtzustand G des Subjektes α ein Ereignis E in einem Moment abgeschlossen ist und zum Zustand F führte³⁰ und α sich für E entschied, auch wenn ein anderes Ereignis möglich gewesen wäre (Negativ-Bedingung des *dstit*-Operators). Um eine Handlung einem Akteur zuzuschreiben, hätte es einen möglichen Geschichtsverlauf zum Zeitpunkt der Entscheidung für die Handlung geben müssen, in der sich

³⁰Es ist hierbei durchaus möglich, dass F in G bereits vorlag, z.B. wenn es sich um das Aufrechterhalten einer Akzeptanz handelt.

3.3. Mentale Zustände und mentale Ereignisse

das Ereignis nicht vollzogen hätte. Dies soll verdeutlichen, dass es sich um eine Entscheidung des Akteurs handelt. Für die Entscheidung, für F mittels des Ereignisses E zu sorgen, ist der Gesamtzustand G ursächlich, wobei nicht notwendig ist, dass eine Proeinstellung für F in G vorlag. Die Fähigkeit lag, für das Ereignis E zu sorgen, in G vor. Die Bewertung, dass der Akteur α für das Ereignis E und damit für den Zustand F gesorgt hat, erfolgt in einem Moment eines Geschichtsverlaufs. Die Bewertung entspricht dabei der von Belnap für den deliberativen Stit-Operator vorgeschlagenen Wahrheitsbedingung. Ein Akteur sorgt dafür, dass p , genau dann, wenn der Akteur es in einem Moment in seiner Entscheidungsgewalt hat, dafür zu sorgen, dass p , wobei p einen Sachverhalt darstellt, der das Ereignis E bzw. den Zustand F charakterisiert. Der Akteur muss, wenn er für p sorgt, nicht beabsichtigen, dass p . Er kann dabei einen anderen beliebigen Sachverhalt beabsichtigen, der möglicherweise E bzw. F charakterisiert. Es kann ebenfalls sein, dass der Akteur keine Absicht zu irgendeinem F oder E charakterisierenden Sachverhalt hat, z.B. wenn der Akteur mit seiner Handlung etwas anderes erreicht, als er beabsichtigt hat. Er hat sich in diesem Fall dennoch für die Ausführung der Handlung entschieden, auch wenn er sich in dem Folgezustand der Handlung möglicherweise getäuscht hat. Auf die Beziehung zwischen Handlung und Intention gehe ich in Abschnitt 4.3 weiter ein.

Die Entscheidung ist die Entscheidung zur Ausübung der Fähigkeit, das Nicht-Eintreten von E und damit das Nicht-Auftreten von F zu verhindern, indem er die Geschichtsverläufe verhindert, in denen E nicht eintritt. Dann ist auch offensichtlich, warum die Negativ-Bedingung bei dem deliberativen Stit-Operator beinhaltet, das ein Nicht-Eintreten von E möglich gewesen wäre. Hierbei ergibt sich eine Besonderheit im Falle des mentalen Aktes im Gegensatz zu einer nicht-mentalenen Handlung. Wenn ein Akteur einen Sachverhalt nicht hervorbringen kann, so kann dies im Falle der Handlung an Situationen liegen, die verschieden von der aktuellen Situation sind. Dies ist möglich, z.B. wenn in allen aktuell möglichen Situationen das Ereignis eintritt und dem Akteur somit nicht zuschreibbar ist. Für einen mentalen Akt ist diese Negativ-Bedingung nicht notwendig, um als Akt zu gelten. Ob es sich um einen mentalen Akt handelt, ist daran festzumachen, ob der Akteur eine Art von Kontrolle über den jeweiligen Akt hat, z.B. ob das Bilden des mentalen Zustandes wie einer Intention unter seiner Kontrolle liegt. Wenn der Akteur diese Kontrolle beim Bilden des mentalen Zustandes ausübt, d.h. dass er sich entschieden hat, diese Kontrolle auszuführen, dann entspricht die Ausübung dieser Kontrolle der Ausübung der Kontrolle bzw. der Fähigkeit, diesen Zustand nicht zu bilden. Ein und dieselbe Situation, nämlich die aktuelle, beinhaltet, dass der Akteur und nichts anderes und niemand sonst dafür gesorgt hat,

3. Handlungen oder die Frage „Actions as Events“

genau diesen mentalen Akt zu vollführen bzw. genau diese Intention zu bilden. Es wäre somit im Falle des mentalen Aktes falsch die Beschreibung eines mentalen Aktes mittels eines Stit-Operators, der eine Negativ-Bedingung erfordert, vorzunehmen. Der *dstit*- und *astit*-Operator ist somit zur Beschreibung mentaler Akte ungeeignet. Ein Operator, der allein die Positiv-Bedingung des *dstit*-Operators in Betracht zieht, bietet sich an, um einen mentalen Akt wie das Fassen einer Intention zu beschreiben. Diesen Operator gibt es. Er wird nach Chellas als *cstit*-Operator bezeichnet. Ich werde in Abschnitt 4.3 auf die Motivation für diesen Operator bzw. in Abschnitt 6.1 bei der Definition der *bdi-stit* auf ihn zurückkommen.

Abschließend lässt sich Folgendes sagen. Auch wenn nicht vollständig geklärt werden kann, wann ein mentales Ereignis eine Handlung ist, ist es entscheidbar, wann ein Ereignis als eine Handlung beschrieben werden kann bzw. wie beschrieben wird, dass eine Handlung bzw. ein mentaler Akt des Akteurs vorliegt. Unabhängig davon, ob es sich um ein mentales oder nicht-mentales Ereignis handelt, ist die Stit-Paraphrasierung ein Test dafür, ob ein Ereignis eine Handlung ist. Demzufolge ist es im Falle von Handlungen immer möglich, statt des die Handlung beschreibenden Satzes die Stit-Paraphrase des Satzes zu verwenden. Da ein Satz einem Sachverhalt korrespondiert und dieser sowohl das Ereignis als auch den Zustand charakterisieren kann, ist es unnötig zu ‘der Akteur dafür sorgt, dass p ’ hinzuzufügen, ob der p korrespondierende Sachverhalt das Ereignis oder den Zustand charakterisiert. Entweder ist p bereits die Beschreibung eines Ereignisses, somit der Handlung. Oder es bezeichnet den durch die Handlung erzielten Folgezustand, der nach vollzogener Handlung vorliegt. Dann wird die Handlung durch eine Beschreibung des Folgezustandes charakterisiert.

Dies wird offensichtlich, wenn wir eine wahrheitsgemäße Beschreibung der Handlung angeben wollen. Ob die Beschreibung korrekt ist, zeigt sich am Folgezustand nach der Handlung und nicht während der Handlung. Angenommen jemand möchte jemanden töten, indem er ihn vergiftet und er vergiftet ihn tatsächlich. Bevor derjenige am Gift sterben kann, wird er von einem anderen Akteur erstochen. Das Ereignis der Vergiftung wäre ein anderes Ereignis als das der Tötung, was sich mittels der Stit-Paraphrasen mit den jeweiligen Akteuren nachweisen lässt. Nicht der mit der Handlung möglicherweise intendierte Folgezustand macht die Handlung aus, sondern der tatsächliche Folgezustand, den der jeweilige Akteur hervorbringt. Somit kann sich eine Handlung, da ihr ein Ereignis zugrunde liegt, über einen bestimmten Zeitraum hinweg erstrecken. Die Vergiftung ist vollzogen, sobald das Opfer das Gift getrunken hat. Die Handlung wurde durch den ersten Akteur mit Tötungsabsicht ausgeführt. Getötet und damit die Handlung der Tötung vollendet, ist jedoch dem zweiten Akteur zuzu-

3.3. Mentale Zustände und mentale Ereignisse

schreiben, sobald der Tod aufgrund des Stiches eingetreten ist. Der Sachverhalt ‘dass das Opfer tot ist’ charakterisiert den Folgezustand dieses Ereignisses. Die Beschreibung der Handlung kann erst in der Situation (Zeitpunkt-Geschichte-Paar) erfolgen, in der die Handlung, Tötung, vollzogen ist.

Für die in dieser Arbeit auszuarbeitende Logik soll untersucht werden, wann eine Handlung vorliegt und in welchen Interaktionen diese mit den Wünschen, Überzeugungen und Intentionen des Handelnden notwendigerweise tritt. Dabei konzentriere ich mich im Folgenden auf den Zustand nach Eintreten des Ereignisses, wenn das Eintreten dieses Zustandes einen das Ereignis der Handlung charakterisierenden Sachverhalt darstellt. So ist allgemein festzuhalten, dass ein das Ereignis der Handlung charakterisierender Sachverhalt ausreicht und dass das Ereignis, das zu diesem Zustand bzw. zum Bestehen des Sachverhaltes führt, eine Handlung ist, die durch eine Stit-Paraphrase und somit durch einen Stit-Operator beschrieben werden kann. Es wird im Weiteren nicht diskutiert, was jemand im Einzelnen getan hat und welche Sachverhalte während des Ereignisses außerdem bestehen, sondern ich beschreibe nur den Sachverhalt, dessen Bestehen das Subjekt durch seine Handlung(en) hervorruft. Die Formel $\alpha \text{ dstit} : \varphi$ wird so gelesen, dass α durch eine Handlung dafür gesorgt hat, dass φ wahr ist, wobei φ eine Proposition repräsentiert, der ein Sachverhalt korrespondiert, für den α bewirkt hat, dass er in dem Moment besteht. Die Frage nach dem Ausgangszustand vor dem Ereignis beantwortet sich durch den Verlauf der Geschichte bis zum Moment im Modell, in dem die Formel ausgewertet wird. Aufgrund der Annahme der Eindeutigkeit der Vergangenheit³¹ sind die möglichen Geschichtsverläufe bis zum Moment m nicht unterscheidbar. Der Zustand nach der Handlung ist die Situation im Modell, in der die Formel ausgewertet wird. In dieser Situation hat sich der Akteur bereits für eine Wahlzelle entschieden, siehe Abschnitt 1.2.

Es gibt Logiken, die in atomaren Elementen einer Logik zwischen Variablen für Ereignissen und für Propositionen unterscheiden, z.B. [47, 79, 92]. Dabei stellt sich die Frage nach der Unterscheidung zwischen Ereignis und Sachverhalt und die weitere Frage, inwieweit Ereignisse als Handlungen einem Akteur zugeschrieben werden. Da meine Argumentation darauf hinausläuft, einem Akteur ein Ereignis als Handlung anhand des das Ereignis bzw. die Handlung charakterisierenden Sachverhaltes zuzuschreiben, der nach Durchführung der Handlung besteht, macht eine Unterscheidung zwischen Variablen für Propositionen und Variablen für Ereignisse keinen Sinn. Sachverhalte und Ereignisse sind dennoch zwei verschiedene Arten von Entitäten.

³¹Siehe Diskussion um die Eindeutigkeit der Historie Abschnitt 1.3.

3. Handlungen oder die Frage „Actions as Events“

Wenn ein Akteur α in einer Situation (m, h) zum Zeitpunkt m im Geschichtsverlauf h dafür sorgt, dass p , ist dies genau dann der Fall, wenn ein Ereignis eingetreten ist, dass zum Zeitpunkt m im Bestehen des durch p beschriebenen Sachverhaltes kulminierte, dessen Akteur α war, so dass es sich bei dem Ereignis um eine Handlung des Akteurs α handelt.

Wenn im Weiteren in der *bdi-stit* Logik sowohl mentale Akte als auch nicht mentale Handlungen bzw. Zustände nach der Handlung, als vom Akteur bewirkt, beschrieben werden, sind interessante Fragen, denen im folgenden Kapitel nachgegangen wird, inwieweit zum Zeitpunkt der Handlung Überzeugungen, Wünsche und Intention den Gesamtzustand charakterisieren und welche für das Eintreten der erfolgreichen Handlung vorliegen müssen, die zum gewünschten beabsichtigten Zustand führt. Von Interesse sind solche Interaktionen, die sich notwendigerweise beim Auftreten einer Handlung und eines Zustandes zeigen. In die formale Sprache übertragen lautet die Frage, wie die Logik außer den Handlungsoperatoren auch Operatoren darstellt, die beschreiben, dass ein Akteur wünscht, glaubt und beabsichtigt, und inwieweit sich der *dstit*-Operator mit Operatoren der Form $\alpha des: (\alpha$ wünscht, dass), $\alpha bel: (\alpha$ glaubt, dass) und $\alpha int: (\alpha$ beabsichtigt/intendiert, dass) kombinieren lässt. Den Fragen, welche Wahrheitsbedingungen diesen Operatoren zugewiesen werden und ob es notwendige Beziehungen zwischen diesen Operatoren gibt, gehe ich im nächsten Kapitel auf den Grund, indem ich die Funktion der mentalen Zustände und die Kontrolle bei deren Bildung analysiere. Eine erste Fragestellung, die sich aus den bisherigen Überlegungen ergibt, lautet sicherlich: Ist die Bildung von Wünschen, Überzeugungen und Absichten eine Handlung oder nicht? Eine zweite lautet: Welche Beziehungen lassen sich in *bdi-stit* Logik zwischen diesen Operatoren formulieren?

Vorweg sei gesagt, auch wenn eine mentale propositionale Einstellung eine Handlung ist, dass sie in *bdi-stit* Logik immer durch eine Formel beschrieben wird, die durch eine Proposition ‘das Subjekt hat die mentale Einstellung, dass . . .’ ausgedrückt wird. Wir beschreiben in der *bdi-stit* Logik sozusagen nicht das Ereignis der Bildung des Zustandes, d.h. nicht den mentalen Akt selbst, sondern den mentalen Zustand des Akteurs nach dem Ausführen, der dem Akt als Folgezustand inhärent ist.

4 Überzeugungen, Wünsche und Absichten

In den vorangegangenen Kapiteln ist den Fragen nachgegangen worden: Was ist eine geeignete Beschreibung von Ereignissen und Zuständen mit Berücksichtigung von mentalen versus nicht-mentalenen Ereignissen und Zuständen? Dieses Kapitel ist den konkreten mentalen Zuständen gewidmet. Mentale Zustände, die bei den Handlungen eines Akteurs und der Bewertung der Rationalität von Handlungen eine Rolle spielen, sind Wünsche, Überzeugungen und Absichten. Für den jeweiligen Typ eines mentalen Zustandes wird auf die Kontrolle, die wir über den jeweiligen Zustand haben, die Passrichtung und die Beziehungen der mentalen Zustandstypen untereinander eingegangen. Des Weiteren wird das Bestehen notwendiger Beziehungen der Zustände zueinander aber auch bei der Ausführung von Handlungen untersucht. Dabei steht die Frage im Vordergrund, ob einer Handlung immer eine bestimmte Überzeugung oder ein bestimmter Wunsch zugrunde liegen muss. Ob das Vorhandensein eines solchen Zustandes notwendig ist, wird sowohl nach deskriptiven als auch nach normativen Gesichtspunkten untersucht. Die normativen Gesichtspunkte sind auf die Frage nach der Rationalität einer Handlung beim Haben eines Zustandes bzw. bei der Bildung eines mentalen Zustandes beschränkt. Wenn ich im Folgenden von „normativ“ spreche, beziehe ich mich lediglich auf die Bewertung als rational oder nicht rational.¹

Paradigmatische Beispiele für mentale Zustände sind Wünsche und Überzeugungen. Sie werden meistens als gegensätzlich in ihrer Passrichtung und von vielen als grundlegend für andere propositionalen Einstellungen angenommen, so dass z.B. In-

¹Normative Urteile, moralische Fragestellungen, ob eine Handlung moralisch gut ist, möchte ich außer Acht lassen. Mir ist klar, dass es diskussionswürdig ist, ob Rationalität überhaupt normativ ist [35]. Diese Diskussion würde jedoch den Rahmen der Arbeit übersteigen. Ich setze voraus, wenn F rational für den Akteur ist, dass er F sollte oder dass es so sein sollte, dass F , wobei F für 'eine bestimmte Einstellung haben' oder 'eine bestimmte Handlung vollziehen' steht. Demzufolge ist Rationalität in gewisser Weise normativ. Weiterhin unterscheide ich nicht zwischen theoretischer und praktischer Rationalität, sondern formuliere, dass, eine Überzeugung zu haben, rational ist bzw. dass ein Akteur rational ist, eine Intention zu fassen.

4. Überzeugungen, Wünsche und Absichten

tentionen von einigen Philosophen auf Wünsche und Überzeugungen zurückgeführt werden [8, 48]. Von diesem sogenannten Überzeugung-Wunsch-Modell von Intentionen möchte ich mich distanzieren, wobei ich mich auf Bratman [24, 25] berufen werde. Intentionen sind eine eigene Form eines mentalen Zustandes. Sie unterscheiden sich in ihrer Funktion von den mentalen Zuständen der Wünsche und Überzeugungen, so dass Intentionen sich nicht auf letztere zurückführen lassen.

Wenn nach der Motivation für Beziehungen zwischen mentalen Zuständen und Handlungen gefragt wird, wird als Handlungserklärung meist eine Kombination aus Wunsch und Mittel-Zweck-Überzeugung angegeben, die von manchen als notwendig² bzw. von anderen als hinreichend für eine Handlungsmotivation³ bzw. für eine Ursache der Handlung in Betracht gezogen wird.⁴ Viele legen sich fest, dass einer Handlung immer ein Wunsch vorausgehen muss und dass eine Handlung nicht nur aufgrund von Verstandesüberlegungen durchgeführt wird [101, 129]. Wiederum wird von einigen festgehalten, dass eine solche Kombination von Wunsch und Überzeugung nicht nur ein rationaler Grund für eine Handlung sein kann, sondern dass ein solcher Grund ursächlich für eine Handlung ist [48, Essay 1]. Ob es notwendig ist, dass eine Handlung durch einen Wunsch motiviert wird bzw. dass einer Handlung immer eine Überzeugung der Form zukommt, dass diese Handlung vollzogen werden sollte, um diesen Wunsch zu erfüllen, werde ich kurz thematisieren.

Zu Beginn dieses Abschnittes setze ich mich mit den mentalen Zuständen ‘Wunsch’ und ‘Überzeugung’ auseinander, welche Passrichtung diese Einstellungen haben bzw. welche Passrichtungen propositionale Einstellungen im Allgemeinen haben.⁵ Auch

²Für notwendig aber nicht hinreichend einer solchen Kombination spricht Nagel: „[T]he temptation to postulate a desire at the root of every motivation is similar to the temptation to postulate a belief behind every inference. [...] the reply in both cases is the same: that this is true in the trivial sense that a desire or belief is always present when reasons motivate or convince — but not that the desire or belief explains the motivation or conclusion, or provides a reason for it“ [101, S.30f]. Das Vorliegen eines Wunsches verursacht allein keine Motivation. So wie man ein Schlussprinzip braucht, um aus einer Überzeugung eine weitere Überzeugung zu ziehen, bspw. Modus Ponens, so braucht man laut Nagel nicht nur den Wunsch, sondern auch eine Operation, die die rationale Motivation vervollständigt [101, S.35].

³Smith [129] lässt offen, ob eine hinreichende Handlungserklärung als eine Ursache gelten muss oder nicht. Er vertritt die Meinung, dass ein Akteur eine Motivation für eine Handlung genau dann hat, wenn ein Wunsch und die entsprechende Mittel-Zweck-Überzeugung vorliegt, wie diese Handlung den Wunsch befriedigt.

⁴Davidson in [48, Essay 1] vertritt, dass solch primäre Gründe (Wunsch-Überzeugung) eine Ursache für eine Handlung sind. Aber nicht jede solche Wunsch-Überzeugung-Kombination ist ein primärer Grund.

⁵In der Literatur existieren verschiedene Theorien, was Wünsche sind bzw. durch was sie definiert wer-

wenn es sich bei Wünschen und Überzeugungen um mentale Zustände handelt, geben nicht alle propositionalen Einstellungen notwendigerweise Zustände wieder. Z.B. bezeichnet ‘urteilen, dass p ’ ein mentales Ereignis. Welcher Ansatz auch betrachtet wird, es gibt zwei Arten von Passrichtungen, die einer propositionalen Einstellung zukommen können. Entweder die Welt passt sich dem Geist an, oder der Geist passt sich der Welt an. Smith zufolge ist für eine Handlungsmotivation erforderlich, dass eine Einstellung in der Passrichtung Welt-Geist vorliegt. Einen motivierenden Grund zu haben, ist unter anderem auch ein Ziel zu haben [129, S.55]. Smith identifiziert diese Einstellung, ein Ziel zu haben, mit einem Wunsch, das Ziel erreichen zu wollen.

Um einer Einstellung eine Passrichtung zuzuweisen, ist es nicht nötig, dass es sich dabei um einen mentalen Zustand handelt. Eine Einstellung in der Passrichtung ‘Welt-Geist’ sollte im Falle einer Handlungsmotivation vorliegen. Jedoch ist nicht notwendigerweise ein bestimmter Wunsch vonnöten. Ein Grund dafür ist, dass nicht nur Wünsche, sondern auch bestimmte Überzeugungen bzw. andere mentale Zustände beide Passrichtungen aufweisen können. Wenn nicht die konkrete Überzeugung in dem Akteur in der aktuellen Situation betrachtet wird, sondern von Überzeugungen und Wünsche als Zustandstypen gesprochen wird, so ist es nur natürlich anzunehmen, dass diese beiden Zustandstypen eigentlich entgegengesetzte Passrichtungen haben. Dennoch ist es möglich, dass ein konkreter, in einem Akteur vorliegender Wunsch

den und wofür sie gut sind, z.B. als Motivation für Handlungen, als Überzeugungen, das das Gewünschte gut für den Wünschenden ist, etc. Wünsche sind graduell. Man kann sich etwas mehr oder weniger wünschen als etwas anderes. Die Möglichkeit zur Graduierung haben sie mit Überzeugungen gemein. Obwohl man sich sowohl Objekte als auch Sachverhalte wünschen kann: 1. Er wünscht Bier. 2. Er wünscht Bier zu trinken. Oder 3. er wünscht, dass er Bier trinkt. Jeder der drei Wünsche impliziert einen anderen. Es wird niemand bezweifeln, dass aus 3. 2. und aus 2. 1. folgt. Die Frage, ob 3. aus 1. folgt, verneinen einige Autoren [140]. Als Begründung benennt Thagard, dass es Lebewesen gibt, die Wünsche haben, ohne dass sie eine Proposition begreifen müssen, die ihrem Wunschgehalt korrespondiert. Eine ähnliche Kritik wird von Ben-Yami in [20] vorgebracht. Er meint, dass die meisten mentalen Zustände nicht generell als propositionale Einstellungen betrachtet werden können. Als Beispiele führt er mentale Einstellungen an, die in der Sprache nicht durch eine Einstellung zu einem Satz, sondern durch Infinitivkonstruktionen beschrieben werden. ‘Ich möchte schlafen’ kann nicht durch ein ‘ich möchte, dass ich schlafe’ paraphrasiert werden, ohne die Aussage zu verändern. So scheint es, dass nicht nur graduell verschiedene Formen des Wünschens, sondern auch verschiedene Formen des Gehaltes existieren, worauf sich der Wunsch richtet. Dass Wünsche keinen propositionalen Gehalt haben, impliziert nicht, dass Wünsche einem Akteur nicht durch Konstruktionen ‘der Akteur wünscht, dass’ zugeschrieben werden können. Auch wenn ‘ich möchte, dass ich schlafe’ in der ersten Person nicht äquivalent ist zu ‘ich möchte schlafen’, ist diese Äquivalenz in der dritten Person unkritisch. Ähnlich argumentiert Smith in [129].

4. Überzeugungen, Wünsche und Absichten

bzw. eine konkrete Überzeugung (Instanzen des jeweiligen Zustandstypes) beide Passrichtungen haben kann. Dies führt dazu, dass einer Handlungsmotivation weder eine bestimmte Überzeugung noch ein bestimmter Wunsch zugrunde liegen muss. Es ist möglich, dass eine Überzeugung bzw. ein Wunsch allein Motivation für eine Handlung sein kann.

Dieses Kapitel ist wie folgt gegliedert. Im ersten Abschnitt stehen die mentalen Zustände bzw. propositionalen Einstellungen ‘Überzeugung’ und ‘Wunsch’ zur Diskussion. Ich werde zeigen, dass weder das Belief-Desire Modell von Davidson noch das als Humesche Theorie von Handlungsmotivation bezeichnete Vorhaben von Smith notwendige Beziehungen zwischen Überzeugung und Handlung bzw. Wunsch und Handlung postuliert. Dennoch unterscheiden sich Wünsche und Überzeugungen anhand der Passrichtung, die ihnen im Allgemeinen zugeschrieben wird. Verschiedene Vorschläge werden diskutiert, wie die Passrichtung zu verstanden werden kann.

Zum Abschluss des Abschnittes 4.1 wird beschrieben, welche logischen Operatoren in welcher Semantik Überzeugungen am besten beschreiben und ob Axiomenschemata motiviert werden können, die für den Operator als allgemeingültig bzw. erfüllbar anzusehen sind.

In Abschnitt 4.2 wird aufgezeigt, dass es sich bei Absichten und Wünschen aufgrund ihrer Funktionen, die die jeweiligen Zustände übernehmen, um zwei verschiedene Typen mentaler Zustände handelt.⁶ Dabei werde ich nicht nur eine geeignete Beschreibung mittels logischer Operatoren motivieren, sondern auch darauf eingehen, welche Interaktionen zwischen diesen mentalen Zuständen bestehen und welche sinnvoll in einem logischen System wiederzugeben sind, das diese Zustände beschreibt.

Dabei ist ein wichtiger Unterschied zwischen Absichten und Wünschen die Funktion einer Absicht, die Handlungen eines rationalen Akteurs zu steuern, um bestimmte Ziele und Pläne umsetzen zu können [25]. Dies führt dazu, dass im Falle eines rationalen Akteurs enge Zusammenhänge zwischen seinen Handlungen und seinen Intentionen bestehen. Auf diese Zusammenhänge zwischen Handlungen und Intentionen eines einzelnen Akteurs und welche Auswirkungen dies auf eine Beschreibung in einem logischen System hat, werde ich abschließend in Abschnitt 4.3 anführen.

Dieses Kapitel stellt zusammengefasst die Motivation für die in den folgenden Kapiteln zu postulierenden *bdi-stit* Logik dar. Es werden sowohl Gründe angegeben, warum die Operatoren für die propositionalen Einstellungen in der jeweiligen Semantik interpretiert werden, als auch die Argumente zusammengetragen, warum verschie-

⁶Ich verwende ‘Absicht’ und ‘Intention’ synonym.

dene Interaktionsschemata zwischen den Operatoren angenommen werden.

4.1 Überzeugungen und Wünsche

Eine Logik von Handlungen und Überzeugungen beinhaltet Zusammenhänge und Relationen zwischen dem Ausführen von Handlungen und dem Vorliegen bestimmter mentaler Zustände. Es gibt zwei Arten, wie diese Zusammenhänge motiviert werden. Zum einen bestehen sie aufgrund deskriptiver Aspekte zwischen Handlungen und Überzeugungen, dass ein Akteur eine Handlung nur durchführt, wenn eine bestimmte Überzeugung vorliegt. Zum anderen kann der Zusammenhang aufgrund einer normativen Begründung der Beziehung zwischen Handlung und Überzeugung bestehen, z.B. dass es für einen Akteur nur rational sein kann, eine Handlung zu vollziehen, wenn in ihm eine bestimmte Überzeugung vorliegt. Welche und ob solche Zusammenhänge existieren, soll in diesem Abschnitt diskutiert werden.⁷

Für viele Philosophen sind Überzeugungen und Wünsche die grundlegenden mentalen Zustände. Nach dem Überzeugung-Wunsch-Modell für die Motivation von Handlungen, wie es auf Hume zurückgeht, ist eine bestimmte Form von Mittel-Zweck-Überzeugung und Wunsch nicht nur notwendig, sondern für Smith in [129] auch hinreichend für die Motivation einer Handlung:

(P1) Ein Subjekt hat genau dann einen motivierenden Grund (Motivation) zu handeln, wenn das Subjekt einen Wunsch hat und die Überzeugung, dass mit der Handlung der Wunsch befriedigt würde.⁸

Dieses Modell ist oft kritisiert worden. Es geht von einer grundlegenden Unterscheidung mentaler Zustände hinsichtlich ihrer Funktion aus. Für Non-Kognitivisten ist bzw. kann ein Wunsch keine Überzeugung sein. Eine vom Verstand diktierte Überzeugung kann allein keine Motivation zu einer Handlung darstellen. Jeder (absichtlichen) Handlung ist ein Wunsch vorausgegangen. Überzeugungen sind allein nicht hinreichend für Handlungen, Wünsche sind jedoch notwendig. Ich werde in diesem Abschnitt auf diese Annahmen kurz eingehen.

⁷Die Debatte um die Motivation und Ursache von Handlungen (Handlungserklärungen) und um die normative Bewertung von Handlungen (Rechtfertigung) besteht seit Anbeginn der Philosophie. Somit kann dieser kurze Abschnitt nicht als Übersicht oder Beitrag zu dieser Debatte verstanden werden, sondern es soll ein bestimmter Vorschlag daraufhin untersucht werden, ob er mögliche logische Interaktionsaxiome zwischen den Operatoren motivieren kann.

⁸Die Formulierung ist dem Prinzip (P1) aus [129, S.36] entlehnt. Smith argumentiert dabei mittels der Passrichtung eines mentalen Zustandes bzw. einer Einstellung für die Plausibilität dieser Behauptung.

4. Überzeugungen, Wünsche und Absichten

Unbestritten sind den Einstellungen ‘Wunsch’ und ‘Überzeugung’ gemein, dass es sich um mentale Zustände handelt. Während bei Überzeugungen wenige bestreiten, dass es sich bei dem Inhalt einer Überzeugung um eine Proposition handelt, ist dies im Falle des Wunsches nicht offensichtlich.⁹ Da es in einer Logik um die Zuschreibung eines Wunsches zu einem Akteur und somit um die Repräsentation des Habens des Wunsches des Akteurs, denn um die genaue Referenz des gewünschten Objektes geht, werde ich im Weiteren die mentalen Zustände ‘Wunsch’ und ‘Überzeugung’ als propositionale Einstellungen auffassen.¹⁰

Passrichtung – konative und kognitive Einstellungen

Ein Unterscheidungskriterium, dass von vielen für propositionale Einstellungen bzw. mentale Zustände angenommen wird, ist die sogenannte Passrichtung.¹¹ Die Bezeichnung ‘Passrichtung’ („direction of fit“) geht auf Searle [123, S.23] zurück.¹² Seither hat die genaue Formulierung der Passrichtung viele Philosophen beschäftigt. Ich

⁹Der Wunsch ‘nach Bier’ hat den Wunsch zum Objekt ‘Bier’ und nicht zu einer Proposition inne [140]. Entkräftet könnte dies werden, indem der Wunsch ‘nach Bier’ nur eine abkürzende Formulierung für den Wunsch ist ‘dass der Akteur Bier zur Verfügung hat’ etc. Aber auch für Überzeugungen wurde vereinzelt argumentiert, dass der mentale Zustand nicht immer als Einstellung zu einer Proposition verstanden werden muss [20].

¹⁰Eine ähnliche Formulierung findet sich bei Smith an. Er zieht in Betracht, dass Wünsche durchaus nicht propositionalen Gehalt haben, jedoch es immer möglich ist, den Gehalt des Wunsches propositional zu fassen, so dass jeder Wunsch einem Akteur über die Einstellung ‘der Akteur wünscht, dass p ’ zugeschrieben werden kann [129, S.50].

¹¹Ich werde dafür argumentieren, dass propositionale Einstellungen vom Typ her möglicherweise durch das Haben einer Passrichtung unterschieden werden können. Dennoch ist es möglich, dass eine konkrete Einstellung in Form eines konkreten mentalen Zustandes oder in eines konkreten mentalen Aktes beide Passrichtungen aufweisen kann. Propositionale Einstellungen, die nicht in Form von mentalen Zuständen realisiert sind, sondern z.B. in Form von mentalen Akten bzw. mentalen Vorgängen, können ebenfalls eine Passrichtung aufweisen. Daher unterscheide ich nicht die Passrichtung für mentale Zustände, sondern unterscheide Passrichtungen von propositionalen Einstellungen.

¹²Die Unterscheidung der Passrichtung bei Sprechakten findet sich in Austin [10]. Searle leitet in Analogie zur Unterscheidung direkter und assertiver Sprechakte eine Unterscheidung von intentionalen Zuständen her. Ein direkter Sprechakt ist in der Ausrichtung von Welt auf Wort und deutet an, dass das Wort auf die Welt einwirken soll und diese sich dann an dem Wort orientiert. Ein Beispiel wäre ein Befehl. Ein assertiver Sprechakt ist in der Wort-auf-Welt-Ausrichtung. Das Wort soll die Welt wiedergeben, z.B. in Form von Behauptungen. Diese Ausrichtung für Sprechakte zwischen Wort und Welt überträgt Searle auf mentale Zustände und die Ausrichtungen zwischen Geist und Welt. Die Idee der Passrichtung für Einstellungen findet sich bereits bei Anscombe [7]. Sie unterscheidet Einstellungen, die sich auf die Welt richten und welche die sich nach der Welt richten.

möchte auf Vor- und Nachteile zweier Formulierungen eingehen. Die eine wird normativ, die andere deskriptiv die Idee der Passrichtung aufgreifen. Ich werde begründen, dass es sich bei der Passrichtung nicht nur um eine Metapher handeln muss, solange man sie deskriptiv beschreibt.

Die Idee der Passrichtung besagt, dass es zwei Arten von propositionalen Einstellungen geben kann. Entweder die Einstellung ist darauf ausgerichtet, dass ihr propositionales Objekt die Welt dem Subjekt abbildet. Oder aber die Einstellung bildet eine Welt ab, in der die Proposition der Fall ist, die es hervorzubringen gilt, so dass das Subjekt als Träger der Einstellung den Gehalt der Einstellung auf die tatsächliche Welt als Muster überträgt.

Einige verstehen die Passrichtung rein metaphorisch andere normativ. Letztere interpretieren die Passrichtung als eine Eigenschaft, wie Zustände wie Überzeugungen und Wünsche an die Welt bzw. den Geist angepasst sein *sollten*. Somit ist eine Einstellung in der Passrichtung Geist-Welt als eine Einstellung zu verstehen, für welche die Eigenschaft, auf Wahrheit abzuzielen, als konstitutiv erachtet wird, d.h. als Fehler in der Einstellung, wenn die Wahrheit des Gehaltes nicht gegeben ist. Dies ist die sogenannte Korrektheitsnorm. Eine Überzeugung *soll* nur angenommen werden, wenn sie korrekt ist, d.h. der Gehalt der Überzeugung wahr ist.¹³ Im Falle der Welt-Geist Einstellung *soll* es der Fall sein, dass das Subjekt den Gehalt der Einstellung zu realisieren bzw. zu erfüllen bestrebt sein sollte. Ist der Gehalt nicht realisiert, wird formuliert, dass im Falle der Welt-Geist-Richtung der Fehler in der Welt liegt, im Gegensatz zu der Geist-Welt-Richtung, wo es ein Fehler in der Einstellung ist, da diese die Welt falsch abbildet.¹⁴ Platts formuliert seine Unterscheidung von Überzeugungen

¹³Es gibt Befürworter einer stärkeren These, dass die Korrektheitsnorm verlange, wenn eine Proposition begriffen bzw. betrachtet wird, dass jede wahre Proposition geglaubt werden sollte [160]. Diese stärkere Norm und die Korrektheitsnorm sind in [37] scharf kritisiert und abgelehnt worden. Nichtsdestotrotz wird an der Diskussion dieser Norm deutlich, dass es nicht die Korrektheit von Überzeugungen ist, die die Passrichtung bestimmt, ob sie nun konstitutiv für Überzeugungen ist oder nicht.

¹⁴Anscombe [7] spricht davon, dass Fehler begangen worden sind, z.B. vom Dedektiv, wenn seine Untersuchung nicht mit dem übereinstimmt, was der Mann kaufte. Aber es ist nicht die Einkaufsliste, die falsch ist, wenn nicht alle Sachen der Liste gekauft wurden, sondern der Fehler liegt in der Welt. Searle sagt: „die Idee der Ausrichtung ist die Idee von der Verantwortung fürs Passen“ [123, S.23]. „Die Überzeugung trägt sozusagen die Verantwortung dafür, zur Welt zu passen, und wenn Welt und Überzeugung nicht zueinander passen, dann bringe ich die Situation dadurch in Ordnung, daß ich die Überzeugung abändere. [...] [w]enn [...] meine Wünsche sich nicht erfüllen, dann kann ich die Situation nicht ebenfalls wiederum dadurch richtigstellen, daß ich einfach [...] den Wunsch abändere. In diesen Fällen ist es sozusagen die Schuld der Welt, daß sie nicht [...] zu dem Wunsch paßt.“ [123, S.24]

4. Überzeugungen, Wünsche und Absichten

und Wünschen wie folgt:¹⁵

„The distinction is in terms of the *direction of fit* of mental states to the world. Beliefs aim at being true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit the world, not vice versa. Desires aim at realization, and their realization is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing *in the desire*, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa.“ [108, S.256f]

Eine solche normative Formulierung hat ihre Tücken, wenn die Korrektheitsnorm einer Überzeugung nicht nur konstitutiv, sondern regulativ verstanden wird. Im Hinblick darauf, dass die Bildung und Verwerfung von Überzeugungen nicht immer unter unserer direkten Kontrolle liegt, sind Formulierungen von „beliefs should be changed to fit the world“ oder „false beliefs should be discarded“ sicherlich problematisch und bedürfen weiterer Klärung, wenn die Passrichtung eine regulative Norm begründen soll. Zudem sind es nicht die Zustände, die auf etwas abzielen. Der Akteur befindet sich in einem bestimmten Zustand, der eine bestimmte Funktion hat. Die Norm, die sich dann durch Platts Formulierung ableiten lässt, ist, dass ein Akteur darauf abzielen sollte, dass seine Überzeugungen wahr und seine Wünsche befriedigt sind.¹⁶

Wenn die Korrektheitsnorm nicht regulativ, sondern konstitutiv verstanden wird, tritt ebenfalls ein Problem auf. Die Formulierung eignet sich möglicherweise für Wünsche und Überzeugungen. Wenn jedoch die Passrichtung ein Unterscheidungskriterium für Wünsche und Überzeugungen sein soll, muss sie unabhängig von diesen beiden Einstellungen für jeden mentalen Zustand, der eine propositionale Einstellung ist, definiert sein. Es sei denn, dass es keine propositionale Einstellungen gäbe, die nicht auf Wunsch oder Überzeugung zurückführbar sind.¹⁷ Wenn die Passrichtung ‘normativ’ einen Unterschied zwischen Einstellungen festhalten soll, gibt es zwei Möglichkeiten, wie ‘normativ’ zu verstehen ist. Die erste Möglichkeit ist, dass die Einstellung

¹⁵Die Formulierung ist normativ gehalten, aber Platts selbst vertritt auf den darauf folgenden Seiten eher eine metaphorische Ansicht hinsichtlich der Passrichtung.

¹⁶Worauf ich im Weiteren noch eingehen werde, ist dabei die Kontrolle, die ein Akteur über seine Wünsche und Überzeugungen hat. Kontrolle ist vonnöten, wenn eine Korrektheitsnorm beim Haben einer Einstellung regulativ aufgefasst wird, vgl. hierfür das in Abschnitt 2 S.31 bereits eingeführt Prinzip „Sollen impliziert Können“.

¹⁷Searle verneint dies [123, S.51f].

4.1. Überzeugungen und Wünsche

selbst normative Eigenschaften aufweist [173]. Die andere Möglichkeit wäre, dass die Einstellung im Zusammenspiel mit anderen Einstellungen normativ ist. Letztere Möglichkeit schließt sich jedoch aus. Wenn die Passrichtung charakteristisch für eine Einstellung sein soll, ist es nicht möglich, dies an weiteren Einstellungen festzumachen, die im Subjekt vorliegen, die selbst wieder eine feste Passrichtung haben. Ob es Zustände ohne Passrichtung gibt, die für die Ausrichtung der Passrichtung anderer Einstellungen vorliegen müssen, wird von Copp und Sobel bezweifelt [130, S.52].

Im Falle der ersten Möglichkeit liefern Copp und Sobel folgendes Gegenargument [130, S.51]. Wenn anhand der Passrichtung eine Unterscheidung von Einstellungen getroffen wird und diese durch die Passrichtung angezeigt wird, inwieweit eine propositionale Einstellung entweder auf Wahrheit oder auf Erfüllung in der Welt gerichtet sein sollte, ist es nicht möglich, dass für die Definition der Passrichtung auf den Unterschied zwischen Wahrheit und Erfüllung referiert wird. Wenn dies möglich wäre, dann wäre der Unterschied nicht normativ, sondern deskriptiv und die 'normative' Formulierung der Passrichtung eine reine Metapher.

Dass es möglich ist, dem Zustand einer Überzeugung normative Eigenschaften zuzusprechen und die Passrichtung deskriptiv zu fassen, zeigt Velleman. Für ihn besteht der Unterschied in der Passrichtung in einer unterschiedlichen Relation zwischen Proposition, Welt und Subjekt. Er weist daraufhin, dass normative Formulierungen der Passrichtung den Fehler machen, nicht zwischen 'etwas als wahr zu erachten' und 'es erfolgreich als wahr zu erachten' auf der einen Seite und 'es als wahr zu machend' und 'es erfolgreich als wahr zu machend' zu differenzieren.¹⁸ Sie machen entweder die Einstellung oder die Welt für den Erfolg bzw. Misserfolg verantwortlich [149, S.112]. Auch wenn Velleman der Meinung ist, dass Überzeugungen eine essentielle normative Eigenschaft haben [149, S.245], ist es nicht die Passrichtung, die einer Überzeugung eine normative Eigenschaft als konstitutiv zuschreibt.

Es gibt kognitive Einstellungen nach Velleman z.B. 'akzeptieren als wahr, dass p ' und 'annehmen, dass p ', die dieselbe Passrichtung wie eine Überzeugung haben, für die es jedoch nicht konstitutiv ist, ob der Gehalt der Einstellung wahr sein sollte oder nicht. Was eine Überzeugung von einem Wunsch unterscheidet, unterscheidet auch diese Einstellungen von einem Wunsch [149, S.111]. Dennoch ist es disputabel, diese

¹⁸Anstelle von der Wahrheit der Proposition zu sprechen, kann die Verbindung zwischen dem der Proposition korrespondierenden Sachverhalt, der Welt und dem Subjekt ausfindig gemacht werden. Anstelle von 'wahr zu erachten' wird der Sachverhalt 'als bestehend zu erachten' und anstelle von 'als wahr zu machen' 'als hervorzubringen' begriffen. Ich werde im Weiteren fast ausschließlich, die Formulierung mit dem Sachverhalt verwenden, vgl. Abschnitte 4.2 und 4.3.

4. Überzeugungen, Wünsche und Absichten

kognitiven Einstellungen einer Korrektheitsnorm wie Überzeugungen zu unterwerfen. Für Überzeugungen ist Wahrheit nicht nur konstitutiv, sondern die Wahrheit ist das konstitutive Ziel einer Überzeugung nach Velleman [149, 114]. Aber ein konstitutives Ziel zu haben, ist nicht die Unterscheidung, auf die die Passrichtung abzielt. Eine Definition von Passrichtungen in Geist-Welt-Richtung, die Wahrheit nicht als konstitutives Ziel haben, ist die folgende von Velleman:

„The term „direction of fit“ refers to the different ways in which attitudes can relate propositions to the world. In cognitive attitudes, a proposition is grasped as patterned after the world; whereas in conative attitudes, a proposition is grasped as a pattern for the world to follow. The propositional object of desire is regarded not as fact — not, that is as *factum*, having been brought about — but rather as *faciendum*, to be brought about; it’s regarded not as true but as to be made true.“ [148, S.105]

Diese Formulierung enthält keine normativen Terme.¹⁹ Sie definiert eine kognitive bzw. konative Einstellung auch nicht unter zu Hilfenahme anderer Einstellungen, so dass sie dem Gegenargument der Zirkularität von Sobel und Copp entgeht.²⁰ Als paradigmatische kognitive Einstellung zählt die Überzeugung. Eine paradigmatische konative Einstellung ist ein Wunsch. Die Bezeichnung der jeweiligen Passrichtung von ‘kognitiv’ mit ‘Geist-Welt’ bzw. ‘konativ’ mit ‘Welt-Geist’ finde ich ähnlich wie Humberstone [85, S.60 Fußnote 2] unintuitiv und werde stattdessen wie Velleman von ‘ko-

¹⁹Dabei wird ‘faciendum’ als ‘was hervorgebracht werden wird’ bzw. ‘was hervorgebracht werden müsste’, wobei ‘müsste’ im Sinne von notwendig und nicht als obligatorisch verstanden wird.

²⁰Dass eine deskriptive Beschreibung der Passrichtung Zirkularitätsprobleme hat, wenn sie eine Unterscheidung mentaler Zustände liefern soll, jedoch mentale Zustände bereits in der Definition enthält, die sich aufgrund ihrer Funktion unterscheiden, haben [130] am Beispiel der von Smith gebrachten Formulierung [129] aufgezeigt. Smiths Formulierung wirft Probleme auf, z.B. „a belief that *p* is a state that tends to go of existence in the presence of a perception that *not p*“ [129, S.54]. Dass eine Wahrnehmung, dass nicht *p*, bereits die Passrichtung in gewisser Weise vorwegnimmt, macht die Formulierung zirkulär. Aber Sobel und Copp gestehen am Ende ihres Aufsatzes dem Funktionalismus zu, dass es eine deskriptive Unterscheidung von Überzeugungen und Wünsche geben kann, ohne Verwendung der Passrichtung bzw. dass es möglicherweise eine Definition von Passrichtung geben könnte, die Überzeugungen und Wünsche unterscheidet und den von ihnen aufgezeigten Problemen entgegen könnte [130]. Wenn man unterstellt, dass das Erfassen einer Proposition als wahr bzw. als wahrzumachen bereits eine propositionale Einstellung ist, unterliegt man bereits dem Zirkularitätsproblem. Um eine propositionale Einstellung zu bilden, muss der Akteur jedoch die Proposition irgendwie erfasst haben. Was durch die Passrichtung zum Erfassen der Proposition hinzukommt, ist, dass beim Erfassen der Proposition mit erfasst wird, ob diese ein Muster von der Welt bzw. ein Muster für die Welt ist.

4.1. Überzeugungen und Wünsche

gnitiven' und 'konativen' Einstellungen sprechen.²¹ 'Konativ' drückt dabei nicht eine positiv konative Einstellung zu der Proposition aus, sondern soll als positiv oder negativ zu p strebende, antriebshafte Einstellung verstanden werden.²² Die Einstellung 'verhindern wollen, dass p ' ist ebenfalls eine konative Einstellung. Die Frage ist, ob diese Darstellung der Passrichtung ebenfalls nur eine Metapher sein kann oder ob sie tatsächlich eine Charakterisierung bzw. konstitutive Unterscheidung für propositionalen Einstellungen angeben kann. Dies führt zu der Frage: Kann ein mentaler Zustand, der eine konative Einstellung darstellt, die Passrichtung haben, die sonst nur kognitive Einstellungen haben bzw. vice versa?²³

Ich übernehme Vellemans Definition von Passrichtungen. Eine kognitive Einstellung hat die Funktion, dem Subjekt die Welt zu repräsentieren. Sie hat eine Art abbildende Funktion, was in der Welt der Fall ist bzw. welche Sachverhalte bestehen. Eine konative Einstellung hat die Funktion, das Subjekt zum Hervorbringen eines Sachverhaltes anzutreiben. Die Idee der Passrichtung nach Anscombe und Searle war jedoch „passen“ so zu begreifen, dass es auch der Fall sein *sollte*, dass für kognitive Einstellungen bzw. für die Passrichtung Geist-Welt der Inhalt der Einstellung 'wahr' und im Falle der konativen Einstellungen für die Passrichtung Welt-Geist der Gehalt 'gut' ist. Bei Anscombe ist es sonst als Fehler in der Welt anzusehen, dass der gewünschte Sachverhalt nicht erfüllt ist.

Eine solche 'normativ' konstitutive Eigenschaft eines mentalen Zustandes ist nach Velleman jedoch nur für Überzeugungen zu vertreten. Ein Zustand einer Überzeugung ist so geartet, dass er auf Wahrheit gerichtet ist in dem Sinne, dass der Zustand darauf zielt, tatsächlich den Wahrheitswert der Proposition korrekt zu erfassen [149, S.252],²⁴ „because belief entails not only regarding a proposition as true but, in addition, doing so with the aim of getting the truth right“ [149, S.118].

Eine Norm wie die Korrektheitsnorm, die Velleman für Überzeugungen als konsti-

²¹Humberstones Vorschlag von 'thetic' und 'telic' Richtungen zu sprechen, ist jedoch im Deutschen schwer wiederzugeben. Daher verwende ich 'konativ' und 'kognitiv', obwohl vor allem letzterer Ausdruck häufig in anderer Weise gebraucht wird.

²²Velleman spricht von „approving one proposition“ [149, S.115] bzw. „to be made true“.

²³Ob es möglicherweise propositionale Einstellungen gibt, die keine Passrichtung haben, scheint mir unplausibel. Auch wenn es nach Searle Einstellungen wie 'bedauern, dass p ' sind, die die „Nullrichtung“ haben, d.h. keine Passrichtung haben, [123, S.24], bildet 'bedauern, dass p ' meines Erachtens eine kognitive Einstellung, da ich nichts bedauern kann, was ich nicht für wahr erachte. Searle gesteht auch selbst zu [123, S.24], dass 'bedauern' eine Überzeugung beinhaltet, die jedoch nicht seiner normativen Formulierung einer Passrichtung genügt.

²⁴Velleman lehnt ab, dass andere mögliche Wahrheitsziele konstitutiv für Überzeugungen sind, z.B. möglichst viele wahre bzw. viel wahre und wenig falsche Überzeugungen zu haben [149, S.250].

4. Überzeugungen, Wünsche und Absichten

tativ erachtet, gibt es nach Velleman in Analogie bei Wünschen nicht. Für Wünsche legt sich Velleman fest, dass es nicht notwendig ist, um im Zustand des Wünschens, dass p , zu sein, dass das Bestehen von p auf 'gut' zu sein oder ein anderes Prädikat anstrebt.²⁵ Er weist daraufhin, dass für 'gut sein' es nicht nötig ist, dass der Akteur eine Einstellung 'dass es gut ist' hat, sondern es nach einem objektiven Standard 'gut' sein kann. Dies ist dann aber kein konstitutives Merkmal eines Wunsches [149, S.118f]. Anscombes Satan dient ihm als Beispiel. Somit gibt es zwei Möglichkeiten, wie es konstitutiv für den Wunsch sein kann 'gut zu sein'. Entweder ist es subjektiv als 'gut' beurteilt oder es ist objektiv 'gut'. Angenommen ein Satan wünscht, dass etwas Schlechtes der Fall ist. Im ersteren Fall, dass es subjektiv als 'gut' beurteilt wird, müsste der Satan urteilen, dass der Gehalt seines Wunsches gut wäre. Dieser ist jedoch aus seiner Sicht schlecht. Oder der Gehalt müsste objektiv gut sein. Er ist jedoch schlecht. Beides scheint unsinnig.²⁶ Somit kann die Passrichtung nicht die Norm postulieren, dass Wünsche bzw. Einstellungen in der Passrichtung Welt-Geist auf das Gute orientiert sein sollten.²⁷ Dass es nicht konstitutiv für einen Wunsch ist, auf seine Befriedigung abzielen, zeigt sich daran, dass bei Wünschen sehr häufig Antagonismen auftreten, die rational sind. Ein Subjekt kann rational wünschen, dass p , und wünschen, dass nicht p .²⁸ Ein Wunsch beinhaltet, den Gehalt als hervorzubringend zu begreifen, jedoch keine Verpflichtung diesen Gehalt auch tatsächlich hervorzubringen. Somit ist ein normatives Verständnis der Passrichtung irreführend, wenn sie für Wünsche und Überzeugung als konstitutiv erachtet wird.

Die Ansicht, dass eine Überzeugung korrekt sein sollte bzw. dass eine konative Einstellung wie ein Wunsch erfüllt sein sollte, ist nach Humberstone eher als eine Hintergrundintention höherer Stufe zu verstehen, die hinter dem Haben einer Überzeugung und dem Haben eines Wunsches steht [85, S.73ff]. Dies könnte konstitutiv für die jeweilige Passrichtung sein und daher konstitutiv für Überzeugungen bzw. Wünsche sein [85, S.74]. Wenn ein Begriff für einen anderen konstitutiv ist, müsste der Akteur

²⁵Das Prädikat 'gut' ist im Falle von Wünschen zu betrachten, da ein 'befriedigt' bzw. 'realisiert' nicht konstitutiv sein kann, da sie es im Allgemeinen gerade nicht sind.

²⁶Der Satan urteilt, dass es für ihn gut wäre, dass das Schlechte der Fall ist. Dann steht man vor dem Problem zu begründen, welche Einstellung 'für ihn gut zu sein' ist und welche Passrichtung diese Einstellung hat. Wiederum kann eine solche Einstellung nicht die Passrichtung definieren.

²⁷Das Prädikat 'gut zu sein' wurde von Velleman als Beispiel gewählt. Generell stellt er in Frage, ob es irgendein konstitutives Prädikat „attractive“, „desirable“ oder ähnliches für jede konative Einstellung gibt [149, S.109]. Eine kognitivistische Ansicht, die Wünsche als bewertende Urteile ansieht, bedarf eines solchen evaluativen Prädikates. Velleman lehnt diese Interpretation eines Wunsches jedoch ab.

²⁸In Abschnitt 4.2 diskutiere ich den Unterschied zwischen paradoxen und antagonistischen Wünschen.

jedoch, um ein Konzept von ‘Überzeugung’ zu haben, ein Konzept von ‘Intention’ haben, siehe hierfür [22, S.37]. Dies scheint eine unplausible Forderung zu sein.²⁹

Wenn die Passrichtung nicht normativ weder als regulative Norm noch als konstitutive Norm zu verstehen ist, muss auf eine Zuordnung zur Passrichtung von als ‘auf gut’ bzw. ‘auf Wahrheit gerichtet’ abgesehen werden. Ebenso wie im Falle der konativen Einstellung, die nicht nur auf ‘gut sein von p ’ bzw. ‘positiv zu p sein’ orientiert ist, möchte ich die kognitive Einstellung nicht verstanden sehen als eine Einstellung, die den Inhalt immer als wahr bzw. den Sachverhalt immer als bestehend begreift. Eine Einstellung ist kognitiv, wenn das Subjekt in irgendeiner Art den Wahrheitswert des propositionalen Gehaltes bzw. den Sachverhalt als bestehend oder nicht bestehend begreift, ungeachtet dessen, ob es ihn korrekt begreift oder nicht. So ist auch ‘disbelief that p ’ eine kognitive Einstellung. Eine kognitive Einstellung bildet den Wahrheitswert des Gehaltes irgendwie ab, so dass es auch sein kann, dass das Subjekt im Unklaren über den Wahrheitswert bzw. über die alethische Modalität des Gehaltes ist. So sind ‘zweifeln, dass p ’ bzw. ‘für möglich halten, dass p ’ ebenfalls kognitive Einstellungen. Mit ‘den Wahrheitswert abbilden’ ist nicht gemeint, dass der Akteur den Gehalt für wahr bzw. falsch erachtet, wie im Falle von ‘belief’ und ‘disbelief’.

Die Passrichtung wird daher nicht rein metaphorisch, sondern deskriptiv in Form einer antreibenden Funktion einer konativen Einstellung und einer abbildenden Funktion einer kognitiven Einstellung verstanden. Im Weiteren stellen sich dann die Fragen: Welche Einstellungen sind kognitiv und welche Einstellungen konativ? Kann es sich dabei um ein Unterscheidungskriterium für den Typen von Zuständen, wie z.B. Überzeugungen und Wünsche, handeln? Wenn für den Typ einer Überzeugung es konstitutiv ist, eine kognitive Einstellung zu sein, kann eine konkrete Überzeugung dann konativ sein kann? Dann ließe sich erklären, dass einer Handlungsmotivation nicht notwendigerweise ein Wunsch zugrunde liegen muss, sondern möglicherweise eine konkrete Überzeugung, die zudem konativ ist.

In der Diskussion, ob Wünsche notwendig für die Motivation von Handlungen sind, wurden immer wieder moralische Urteile als Kandidat für Handlungsmotivationen ins Feld geführt, die die Geist-Welt-Passrichtung haben. Obwohl sie keine Wünsche sind, sind sie in der Lage, den Akteur zu einer Handlung zu motivieren. Sie gehören zu den Zuständen (Einstellungen), die beide Passrichtungen haben, vgl. Platts [108, S.257]. Somit kann eine konkrete Instanz des Zustandstyps ‘Überzeugung’ selbst konativ sein, wenn ihr Inhalt antreibend hinsichtlich p ist.

²⁹Ich denke Humberstone war sich dieses Problems bewusst und hat von einer besonderen Art von *intend to* gesprochen. Jedoch möchte ich diesen Ansatz nicht weiter.

4. Überzeugungen, Wünsche und Absichten

Die strikte Trennung von konkreten mentalen Zuständen anhand der Passrichtung ist daher nicht möglich. Es gibt konkrete Einstellungen, die beide Passrichtungen haben [108]. Den Schluss, den einige ziehen, ist, dass die Passrichtung keine deskriptive Unterscheidung von Zuständen liefert, sondern nur normativ verstanden werden sollte. Dies teile ich nicht. Die normative Formulierung kann, wie eben gesehen, lediglich eine Metapher und nicht konstitutiv für eine Einstellung sein.

Selbst wenn die Passrichtung keine Explikation der Zustandstypen 'Überzeugung' und 'Wunsch' wäre, so kann sie in Vellemans Formulierung doch als ein Charakteristikum verstanden werden, dass einer propositionalen Einstellung zukommt. Auch wenn es schwierig ist, für einen Typ von mentalem Zustand eine Einschränkung der möglichen Passrichtung für alle Instanziierungen des Typs auf diese eine Richtung festzulegen, so gibt es Typen von Zuständen bzw. Einstellungen, bei denen jeder Instanziierung immer eine Passrichtung zukommt, auch wenn einige Instanzen die andere Richtung zusätzlich aufweisen.

Einschlägige Beispiele für die Passrichtung 'Geist-Welt' sind 'akzeptieren als wahr, dass p ', 'überzeugt sein, dass p ', 'zweifeln, dass p ' aber auch 'vorhersehen, dass p ' und 'annehmen, dass p '.³⁰ Genuine Beispiele für die andere Richtung 'Welt-Geist' wären 'wünschen, dass p ', 'verabscheuen, dass p ', 'eine Abneigung dagegen haben, dass p ' und 'hoffen, dass p '. Aber auch wenn den Typen des jeweilige Zustandes eine Passrichtung inhärent ist, schließt es nicht aus, dass konkrete Zustände auch die andere Passrichtung aufweisen.

Es sind nicht nur moralische Urteile, die beide Passrichtungen haben können. Weitere propositionale Einstellung, die beide Passrichtungen haben können, sind meines Erachtens z.B. 'lieben, dass p '. Dies ist eine Einstellung, die sowohl kognitiv als auch konativ zu p ist. Um etwas zu lieben, muss man eine kognitive Einstellung zu p haben, d.h. es als wahr betrachten. Auf der anderen Seite ist es offensichtlich eine konative Einstellung zu p . Auch bei Wünschen und Überzeugungen gibt es Fälle, in denen ein konkreter Wunsch bzw. eine konkrete Überzeugung beide Richtungen inne hat.

Smith spricht sich dagegen aus. Für ihn können Überzeugungen, der Form 'es ist gut, dass p zu tun' Handlungen nur motivieren, weil sie nicht unabhängig von bestimmten Wünschen auftreten. Die Überzeugung, dass es gut ist, p zu bewirken, tritt laut Smith nur dann auf, wenn ein geeigneter Wunsch, der irgendwie in Verbindung zu p steht, ebenfalls vorhanden ist. Smith [129, S.57f] argumentiert wie folgt. Ein „Quasi-Belief“ ist z.B. eine Überzeugung der Form 'glauben, dass p zu tun gut ist'.

³⁰Velleman benennt 'vorstellen' bzw. 'fantasieren, dass p ', als weitere kognitive Einstellung [149, S.247].

4.1. Überzeugungen und Wünsche

Es ist eine Überzeugung und hat somit eine eindeutige Passrichtung. Was das Beispiel laut Smith zeigt, ist nicht, dass diese Überzeugung eine „hybrid kind“ eines mentalen Zustandes ist, sondern dass „they are beliefs that the subject can have only if he has certain *desires*“ [129, S.58]. Um eine solche Überzeugung (Quasi-Belief) überhaupt zu bilden, muss in dem Subjekt bereits ein geeigneter Wunsch vorliegen. Dies scheint unplausibel, zumal Smith selbst zeigt, dass wir uns über unsere Wünsche täuschen können [129, S.49].

Allerdings kann nicht nur eine konkrete Überzeugung eine konative Einstellung sein und somit beide Passrichtungen aufweisen. Auch ein Wunsch kann beide Passrichtungen aufweisen. Der Satz ‘ich wünsche, dass sich nichts ändert’ repräsentiert einen Wunsch. Sein Gehalt impliziert jedoch, dass der Akteur hinsichtlich seines Wunschgehaltes eine kognitive Einstellung hat. Um sich zu wünschen, dass sich nichts ändert, muss er eine Abbildung bzw. Vorstellung der Welt und zwar dessen haben, was gerade der Fall ist, was er wünscht, dass es sich nicht ändern möge. Dennoch liegt nicht notwendigerweise eine Überzeugung vor, was er konkret nicht ändern möchte, sich nicht ändern darf etc. Ob der Wunsch allein jedoch handlungsmotivierend ist, darf bezweifelt werden. Zeigen soll er, dass es möglich ist, dass eine Instanz des Zustandstyps ‘Wunsch’ beide Passrichtungen haben kann. Demzufolge wäre dieser Wunsch noch kein Gegenbeispiel zu (P1).

Velleman meint den Grund, warum mentale Zustände zwei Passrichtungen haben können, in der folgenden Unterscheidung zu finden. Er gibt die Unterscheidung zwischen kognitiv und konativ an. Aber er unterscheidet weiterhin noch Zustände, die ‘direktiv’ oder ‘rezeptiv’ sind.

„In my view, this notion [direction of fit Anm.d.Verf.] conflates two different distinctions. One is the distinction between the cognitive and the conative — the distinction between accepting, or regarding as true, and approving, or regarding as to be made true. The other is a distinction between the receptive and the directive, which are two different ways of attempting to accept what’s true — namely, by accepting so as to reflect the truth, and by accepting so as to create the truth. If these distinctions are conflated under the heading ‘direction of fit’, then one and the same mental state can appear to have two different directions of fit, since a subject can attempt to accept what’s true by accepting something so as to make it true. The resulting state is cognitive rather than conative, but directive rather than receptive: it’s directive cognition. I would claim that this state of directive cognition is the state of intending to act.“ [149, S.195 Fn. 55]

4. Überzeugungen, Wünsche und Absichten

Ich spreche nicht nur von Zuständen, die eine Passrichtung haben, sondern von propositionalen Einstellungen im Allgemeinen, da auch mentale Ereignisse eine Passrichtung aufweisen können. Diese ist wie folgt zu verstehen. Nicht jede propositionale Einstellung kennzeichnet das Vorliegen eines Zustandes. Bei der Akzeptanz der Wahrheit einer Proposition handelt es sich um einen mentalen Akt, der in einem mentalen Zustand kulminiert. Das mentale Ereignis der Bildung und das mentale Ereignis des Vorliegens des Zustandes unterscheiden sich. Die Unterscheidungen zwischen direktiv-rezeptiv und kognitiv-konativ, die Velleman vornimmt, beruhen auf diesem Unterschied zwischen der Passrichtung bei der Bildung des Zustandes und der Passrichtung beim Vorliegen des Zustandes. Ein mentaler Zustand ist kognitiv oder konativ, seine Bildung ist entweder direktiv oder rezeptiv zur Wahrheit des Gehaltes.³¹

Dass sich rezeptiv und direktiv auf die Bildung des Zustandes und konativ und kognitiv in der ersten Formulierung auf das Vorliegen des Zustandes bezieht, wird an Vellemans Zitat deutlich. Wenn Velleman schreibt [149, S.195 Fn. 55], dass ein und derselbe mentale Zustand auf zwei Arten betrachtet werden kann nämlich dahingehend, ob der Akteur den Inhalt bereits als wahr begreift bzw. als etwas, das hervorzubringen ist. Macht es einen Unterschied, ob er gerade dabei ist, die Einstellung zu bilden, oder ob sie bereits vorliegt. Er kann den Gehalt im ersteren als etwas Wahres begreifen und bei letzterem als etwas wahr zu 'Kreierendes' erachten und umgekehrt.³² Bei der Unterscheidung zwischen rezeptiv und kognitiv formuliert Velleman, dass der Akteur etwas 'versucht' zu begreifen „attempting to accept“.³³ Dieses 'Versuchen' deutet auf einen Vorgang bzw. einen Akt hin, der eine Bildung eines mentalen Zustandes beinhaltet. Dagegen ist bei der Unterscheidung von konativ und kognitiv der mentale Zustand laut Velleman bereits eingetreten. Der Akteur hat, dass p , bereits als wahr oder als wahr zu machend begriffen.

Jedoch trifft 'rezeptiv' bzw. 'direktiv' keine Unterscheidung, ob die Bildung ein

³¹Ich spreche abkürzend nur von 'wahr'. Für andere Prädikate des Gehaltes, die eine kognitive bzw. rezeptive Einstellung dem Akteur vermittelt, gilt dies ebenfalls, z.B. für 'falsch' und 'möglicherweise nicht wahr', 'möglicherweise nicht falsch' etc. Analog spreche ich für die konativen und direktiven Einstellung nur von die Proposition 'wahr zu machen' bzw. den Sachverhalt 'hervorzubringen' anstatt alle Prädikate explizit zu erwähnen, wie 'zu vermeiden', 'zu verhindern' etc.

³²Hierbei setzt 'erachten' keine Überzeugung oder andere kognitive Einstellung voraus. Es ist wie „grasp“ und „regarding“ bei Velleman zu verstehen. Dieses Erachten bzw. Begreifen ist nur das Verstehen der Proposition und eine Art Zuordnung, ob diese wahr wie ein bestehendes „Faktum“ oder etwas Hervorzubringendes wie ein „Faciendum“ ist.

³³Auch wenn Velleman von „accepting“ spricht, ist nicht gemeint, dass er die kognitive Einstellung 'Akzeptanz, dass' hat, sondern dass er bereits die Proposition als so oder so begriffen hat.

4.1. Überzeugungen und Wünsche

mentaler Akt oder ein mentaler Vorgang ist. Es gibt mentale Vorgänge, die direktive Einstellungen sind, und mentale Akte, die rezeptiv sind, bzw. umgekehrt. Ein Wunsch wäre eine Einstellung, die sich meist direktiv bildet, da der Akteur angetrieben ist, etwas als erfüllt erachten zu wollen, d.h. den Wunsch zu bilden, ohne dies als Faktum oder nicht Faktum in irgendeiner Form in Betracht zu ziehen. Der vorliegende Wunsch ist konativ, da er den Gehalt als wahr zu machend begreift. Aber eine Wunschbildung ist nicht notwendigerweise ein mentaler Akt, siehe Abschnitt 4.2.

Bildet man die Überzeugung, dass die Sonne scheint, aufgrund eines kurzen Blickes in den Himmel, liegt im Falle der Überzeugung, dass die Sonne scheint, beim Anblick des blauen Himmels eine rezeptive und kognitive Einstellung vor. Die Bildung der Überzeugung erfolgt aufgrund des Begreifens der Proposition als wahr und im gebildeten Zustand wird der Gehalt so verstanden, dass er die Welt abbildet, wie sie ist. Somit ist diese Überzeugung kognitiv und der Vorgang der Bildung rezeptiv. Der vorliegende Zustand bzw. die Einstellung ist insgesamt kognitiv.

Problematisch wird die Sache, wenn beim Vorliegen des Zustandes und bei dem Vorgang der Bildung des Zustandes verschiedene Passrichtungen vorliegen. Z.B. kann ein vorliegender Zustand selbst kognitiv sein, aber die Bildung unter direktivem Aspekt stattgefunden haben. Bei der Akzeptanz der Wahrheit, wie im Falle des Richters Abschnitt 3.3 S.91, kann man von einer direktiven, kognitiven Einstellung sprechen. Der Richter hat als wahr akzeptiert, dass der Angeklagte unschuldig ist. Er erfasst 'dass der Angeklagte unschuldig ist' als wahr. Somit ist die Einstellung kognitiv hinsichtlich ihres Vorliegens. Aber diese Einstellung bildet er nicht rezeptiv. Er empfängt sie nicht bzw. reflektiert nicht über deren Wahrheit, sondern sie ist direktiv. Sie wird gebildet, indem er sie für sich als wahr zu machend begreift. Er 'kreiert' diese Wahrheit für sich in Form der Akzeptanz. Die Einstellung ist somit rezeptiv hinsichtlich ihrer Bildung.

In diesem Falle stimme ich Velleman zu, dass beide Passrichtungen dem Zustand zugeschrieben werden können. Im Falle der Akzeptanz steht aber die Wahrheit von p im Vordergrund, da die Wahrheit von p 'kreiert' wurde, nicht weil die Einstellung konativ zu p ist, sondern konativ zu p als wahr zu begreifen ist. Das Muster, von dem der Akteur möchte, dass die Welt ihm folge, ist p als wahr zu 'kreieren', um p als wahr zu begreifen. Also wird die Wahrheit von p kreiert, um eine kognitive Einstellung zu bilden. Daher würde ich im Falle der Akzeptanz die Einstellung als kognitiv erachten, da der erreichte Zustand kognitiv und die Bildung nur direktiv auf die Bildung eines kognitiven Zustandes und nicht auf dessen Gehalt gerichtet ist.

In den obigen Beispielen wird deutlich, warum der Wunsch, nichts zu ändern, bzw.

4. Überzeugungen, Wünsche und Absichten

die Überzeugungen, dass etwas wünschenswert ist, beide Passrichtungen haben und zwar in dem Sinne, dass bei der Bildung das Begreifen der Proposition die entgegengesetzte Richtung wie beim Zustand selbst hatte. Der Wunsch ‘dass sich nichts ändert’ überträgt das Muster, dass sich nichts ändert, auf die Welt. Der Akteur begreift ‘dass sich nichts ändert’ als etwas Hervorzubringendes. Bei der Wunschbildung muss das Subjekt jedoch für den Wunsch reflektieren, welches Muster er auf die Welt überträgt. Bei diesem Wunsch reflektiert der Akteur somit über den bestehenden Sachverhalt, dass sich nichts ändern soll. Im Akteur findet ein kognitiver Vorgang statt, der dem Akteur den bestehenden Sachverhalt abbildet. Im gebildeten Wunsch ist dieser bestehende Sachverhalt, dass sich nichts ändern soll, genau das, was der Akteur als das auffasst, was wahr zu machen ist. Bei der Wunschbildung ist die Beziehung zu dem Sachverhalt bzw. zur Proposition ‘dass sich nichts ändert’ rezeptiv. Er reflektiert über die Wahrheit, welche Sachverhalte bestehen. Im vorliegenden Zustand ist der Wunsch jedoch konativ also antreibend, da der Sachverhalt als hervorzubringend begriffen wird. Der vorliegende Wunsch ist konativ. Über die Bildung ist er eindeutig rezeptiv und somit eine kognitive Einstellung hinsichtlich der Bildung und eine konative hinsichtlich des Vorliegens.

Die Unterscheidung zwischen kognitiv/konativ und direktiv/rezeptiv ist möglich, indem ich zwischen dem Vorliegen des Zustandes und dem Bilden des Zustandes unterscheide. Statt der Begriffe ‘rezeptiv’ und ‘direktiv’ verwende ich daher ‘konativ’ bzw. ‘kognitiv hinsichtlich der Bildung’ der Einstellung bzw. des Zustandes. Solange die Richtung der Einstellung zwischen dem Subjekt und dem Gehalt im Vordergrund steht, wird abkürzend von kognitiver bzw. konativer Einstellung gesprochen.

Am Beispiel der Überzeugung, dass p wünschenswert ist, lässt sich dies zeigen. Für die Überzeugung muss der Akteur die Proposition ‘dass p wünschenswert ist’ als wahr erfassen. Wenn der Akteur die Proposition als wahr erfasst, hält er p somit für wünschenswert. Also ist dies eine kognitive und konative Einstellung zu p . Dass er p für wünschenswert hält, kann rezeptiv geschehen, indem er über die Wahrheit von ‘dass p wünschenswert ist’ reflektiert. Wenn diese Reflexion dazu führt, dass er p für wünschenswert hält, muss eine konative Einstellung zu p bereits vorliegen. Wenn diese Einstellung zu p nicht vorliegt, dann kann diese Überzeugung nicht rezeptiv, sondern nur direktiv gebildet worden sein, indem der Akteur es als hervorzubringend begreift ‘dass es wünschenswert ist, dass p ’. Demzufolge glaubt er, dass es wünschenswert ist, dass p . Dann wäre die obige Überzeugung eine direktive Einstellung und somit konativ hinsichtlich der Bildung. D.h. sie kann sowohl als konativ als auch kognitiv betrachtet werden.

4.1. Überzeugungen und Wünsche

Ein beliebiger Zustand bzw. eine beliebige Einstellung im Allgemeinen kann sowohl kognitiv als auch konativ sein, wenn sich die Passrichtung bei der Bildung bzw. beim Vorliegen unterscheidet. Jeder propositionalen Einstellung kommt eine Passrichtung zu, da die Proposition auf die eine oder die andere Art begriffen werden muss, um die Einstellung zu bilden. Im Falle der Akzeptanz ist die Einstellung hinsichtlich des Vorliegens des Zustandes kognitiv und hinsichtlich des Bildens konativ. Hierbei tritt jedoch die Besonderheit auf, dass das ‘Konative’ dabei den Akteur zu der Bildung eines kognitiven Zustandes antreibt. Nachdem wir die für die Passrichtung paradigmatischen Zustände ‘Überzeugung’ und ‘Wunsch’ betrachtet haben, werde ich im Folgenden den Zustand der Intention betrachten. Dabei sei für den Moment vorausgesetzt, dass der mentale Zustand der Intention einen propositionalen Gehalt hat.

Eine Intention zum Handeln ist nach Velleman eine direktive, kognitive Einstellung und der erreichte Zustand ist kognitiv und nicht konativ. Wie ist dies zu verstehen und warum ist eine Intention keine Überzeugung? Eine Intention, dass p , wird gebildet, indem der Akteur p als einen Sachverhalt begreift, den es hervorzubringen gilt. Es ist somit eine Einstellung, die direktiv also konativ hinsichtlich ihrer Bildung ist. Die Bildung erfolgt aufgrund des Hervorbringens einer Wahrheit. Wenn der Akteur die Intention geformt hat und der Zustand, dass der Akteur intendiert, dass p , vorliegt, begreift der Akteur eine Proposition als wahr, aber nicht die Proposition p begreift er als wahr, sondern die Proposition, dass er p hervorbringt. Dies macht die Intention sowohl zu einer kognitiven Einstellung, dass er den intendierten Sachverhalt hervorbringt, als auch zu einer konativen Einstellung, dass p .

Velleman bezeichnet Intentionen als eine bestimmte Form einer Überzeugung in [149, Kapitel 9]. Jedoch bereits in [149, S.195 Fn. 55] setzt Velleman fort, dass er nicht an dem Label „belief“ für eine Intention festhält, auch wenn er sie im Folgenden als Überzeugung bezeichnet. In [150, S.206] sagt er, dass eine Intention eine kognitive Verpflichtung zur Wahrheit der intendierten Proposition beinhaltet. In [150, S.209 Fn. 10] sagt er, dass dies offen lässt, ob diese Verpflichtung bereits eine Überzeugung ist.

Ich greife dies auf, um zwei Dinge festzuhalten. Zum einen ist eine Intention immer auch eine kognitive Einstellung. Aber sie ist keine Überzeugung, sondern eine bestimmte Form von Akzeptanz. Der Akteur akzeptiert als wahr, wenn er etwas intendiert, dass er etwas wahr machen wird. Wenn der Akteur intendiert, hat er akzeptiert, den intendierten Sachverhalt hervorzubringen, siehe Abschnitt 4.3. Das bedeutet zum zweiten, dass ich im Gegensatz zu Velleman, aber eigentlich aufgrund seiner Analyse von direktiv, die kognitive Verpflichtung nicht derart verstehe, dass der Akteur verpflichtet ist, die Wahrheit der intendierten Proposition zu erfassen. Er legt sich nicht

4. Überzeugungen, Wünsche und Absichten

fest zu akzeptieren bzw. zu glauben, dass p wahr ist. Vielmehr ist ein Akteur aufgrund dessen, dass die Einstellung direktiv ist, verpflichtet, die Wahrheit zu erfassen, dass er die intendierte Proposition wahr macht.³⁴ Da ein Akteur meist um die eigene Fehlbarkeit weiß, impliziert dies nicht, dass er glaubt, diesen Sachverhalt tatsächlich hervorzubringen. Es impliziert nicht, dass er glauben sollte, dass er diesen Sachverhalt hervorbringen wird.³⁵

Diese Form der Akzeptanz des Akteurs ist keine Form einer Überzeugung. Es ist eine kognitive Einstellung, die dem Akteur den Sachverhalt, dass er p wahr machen wird, als wahr abbildet. Velleman selbst kann dies nicht als Überzeugung benennen. Jedenfalls kann diese Überzeugung nicht konstitutiv für die Intention sein. Wenn eine Intention die Überzeugung beinhaltet, den intendierten Sachverhalt für wahr zu halten, bzw. den Sachverhalt für wahr zu halten, dass der Akteur den intendierten Sachverhalt wahr macht, folgt aus Vellemans Verteidigung der Korrektheitsnorm für Überzeugungen, dass es konstitutiv für die Überzeugung ist, dass der Akteur den intendierten Sachverhalt wahr machen sollte. Demzufolge ist es auch für die Intention konstitutiv, dass der Akteur den intendierten Sachverhalt wahr machen sollte. Ich bin mir nicht sicher, ob Velleman eine solche Norm tatsächlich vertreten würde oder ob dies einer der Gründe ist, warum er an dem Label der Überzeugung nicht notwendigerweise festhält. Zusammengefasst ist eine Intention eine Einstellung, die beide Passrichtungen aufweist. Sie beinhaltet, eine bestimmte Proposition als wahr zu akzeptieren. Die Akzeptanz ist jedoch nicht als eine Form einer Überzeugungen geltend zumachen. Fest steht, dass eine Intention eine konative und immer auch kognitive Einstellung ist, den intendierten Sachverhalt hervorzubringen. Welche Eigenschaften Intentionen noch zukommen, darauf gehe ich in Abschnitt 4.2 bzw. Abschnitt 4.3 ein.

Auf (P1) und die Motivation einer Handlung zurückkommend, kann (P1) vorerst gerettet werden, in dem für die Motivation einer Handlung eine konative Einstellung zu der Handlung vorliegen muss, wobei unerheblich ist, ob die Einstellung konativ aufgrund ihrer Bildung oder im Vorliegen des Zustandes ist. Die Einstellung muss jedoch nicht notwendigerweise positiv zu p sein. Insbesondere muss der Akteur nicht wünschen, dass p . Im Beispiel der Fahrerflucht aus dem vorangegangenen Kapitel habe ich dafür argumentiert, dass der Akteur gehandelt hat, ohne dass eine Proeinstel-

³⁴Diese Form von „Commitment“ ist nicht normativ gemeint in dem Sinne, dass der Akteur einer bestimmten Pflicht folgt. Es ist lediglich, ein ‘Sich-Festlegen’ auf eine bestimmte Wahrheit gemeint.

³⁵Die Diskussion, inwieweit eine Überzeugung einer Intention zugrunde liegt oder mit ihr einhergehen sollte, führe ich in Abschnitt 4.3 bei der Ablehnung der Interaktionsschemata von Intentionen- und Überzeugungsoperatoren aus.

lung für die Fahrerflucht im Akteur vorlag.

Im Akteur liegt im Falle einer Handlung eine konative Einstellung vor. Jedoch muss dies nicht der Wunsch zu dem Sachverhalt sein, den er hervorbringt, ob es überhaupt ein Wunsch ist, wage ich zu bezweifeln. Im Falle der Fahrerflucht ist dies wie folgt zu verstehen. Der Akteur hat den Unfall. Er spürt den Antrieb schleunigst, von dem Unfallort fortzufahren, und entscheidet sich, auf das Gaspedal zu treten und wegzufahren. Er muss jedoch nicht wünschen, Fahrerflucht zu begehen. Er sorgt dafür, dass er Fahrerflucht begeht, aber er hat keinen Wunsch in diesem Fall, die Fahrerflucht zu begehen. Möglicherweise hat er absichtlich auf das Gaspedal getreten, aber nicht jede seiner Folgen des auf das Gaspedal Tretens, die ihm als Handlungen zugesprochen werden, sind tatsächlich absichtlich. Sie sind unabsichtlich, aber ihm zuzuschreiben. In diesem Fall ist es nicht der Wunsch zur Fahrerflucht, der laut (P1) vorliegt. Aber welcher Wunsch und welche zu diesem Wunsch passende Mittel-Zweck-Überzeugung ist es, die die Handlung motiviert? Er müsste den Wunsch haben wegzufahren und glauben, dass auf das Gaspedal zu treten, diesen Wunsch erfüllen würde. Wenn jedoch eine spontane Entscheidung gefällt wird, scheint dies eher intuitiv aus dem Gefühl heraus zu erfolgen, als aufgrund einer Analyse der Form 'ich trete auf das Gaspedal, um wegzufahren, weil ich mir wünsche wegzufahren'. Eine solche Analyse mag im Nachhinein möglich sein, ihm zu erklären, was passiert ist, nachdem er verstanden hat, was passiert ist. Dies ist eine Erklärung, aber ist es auch die Handlungserklärung, die ursächlich zur Fahrerflucht geführt hat. Ich würde bestreiten, dass der Wunsch und die Mittel-Zweck-Überzeugung sich gebildet haben, bevor er handelte. Es ist ihm als seine Handlung zuzuschreiben.

Ein weiteres Beispiel gegen (P1) ist der Fall, wo Wünsche durch Handlungen zu erfüllen sind, die keiner Mittel-Zweck Überzeugung bzw. nur einer trivialen Mittel-Zweck Überzeugung bedürfen, wie im Falle des Wunsches, eine Basishandlung zu vollführen. Ich möchte mich strecken. Ich tue es einfach, dehne mich und höre wieder auf. Nach meiner Bewegung ist mir diese auch bewusst. Ich habe es auch absichtlich getan. Die einzige Mittel-Zweck Überzeugung, die dafür notwendig ist, um den Wunsch, dass ich mich strecke, zu erfüllen ist die Überzeugung, dass ich, in dem ich mich strecke, mich strecke, $\alpha \text{ bel} : (\alpha \text{ dstit} : p \supset \alpha \text{ dstit} : p)$. Aber ist diese Überzeugung notwendig, um mich zu strecken?³⁶ Ich denke nicht. Es gibt Handlungen, die

³⁶Dieses Gegenbeispiel ist anders als das von Smith zurückgewiesene Gegenbeispiel geartet, wo die Mittel-Zweck-Überzeugung dem Akteur fehlte, weil es unabhängig von den Einstellungen des Akteurs eine Art Grund im Sinne von Tatsache gegeben hätte, die ihn zum Handeln möglicherweise gebracht hätte, wenn er Kenntnis davon gehabt hätte.

4. Überzeugungen, Wünsche und Absichten

ich gänzlich aus einem Wunsch heraus ausführen kann, weil ich die Handlung um ihrer Selbstwillen ausführen möchte, ohne dass es der Überzeugung bedarf, dass der Wunsch durch diese Handlung befriedigt wird. Entgegen des Überzeugung-Wunsch-Modells, wie in (P1) verstanden, ist für eine Motivation zum Handeln u.U. keine besondere Form einer Mittel-Zweck Überzeugung erforderlich.

Zusammengefasst habe ich zwei Dinge in diesem Abschnitt bisher gezeigt. Zum einen ist es für eine Handlung nicht notwendig, dass immer ein Wunsch bzw. eine Überzeugung der Handlung voraus- bzw. mit ihr einhergeht. Es ist nicht sinnvoll, zu einem konkreten Sachverhalt eine Interaktion zwischen dem Dafür-Sorgen des Akteurs für diesen Sachverhalt und einer Überzeugung oder einem Wunsch, dessen Gehalt zu p in Beziehung steht, zu fordern. Ein Akteur sorgt für mehr Sachverhalte mit einer Handlung, als er glaubt bzw. wünscht. Ich möchte aus dem Überzeugung-Wunsch-Modell (P1) der Handlungsmotivation keine notwendigen Zusammenhänge zwischen den Wünschen bzw. Überzeugungen eines Akteurs und seinen Handlungen ableiten. Meines Erachtens gibt es keine durch (P1) motivierten Interaktionsschemata zwischen dem Überzeugungs- und Handlungsoperator sowie zwischen dem Wunsch- und Handlungsoperator. Falls der Akteur für etwas sorgt, muss im Akteur nicht notwendigerweise eine bestimmte Überzeugung bzw. ein bestimmter Wunsch vorliegen, die ihn zu der Handlung motivieren. Es ist jedoch sowohl eine konative als auch eine kognitive Einstellung nötig, die in Form von Wunsch und Überzeugung oder einer Einstellung, die beide Passrichtungen hat, vorliegen kann. Auch kann es sein, dass ein einziger mentaler Zustand einer Überzeugung oder eines Wunsches eine Handlung motivieren kann.

Auf der anderen Seite habe ich begründet, dass es zwei Arten der Passrichtung gibt, die darauf hindeuten, dass Wünsche und Überzeugungen zwei verschiedene Typen mentaler Zustände sind. Denn auch wenn es möglich ist, dass sowohl eine konkrete Überzeugung als auch ein konkreter Wunsch beide Passrichtungen haben kann, gibt es jeweils eine Passrichtung, die jeder Instanz eines Wunsches bzw. einer Überzeugung zukommt. Soweit ich weiß, wird dies von niemandem bestritten.

Im Folgenden werde ich die Einstellung der 'Überzeugung' betrachten und diskutieren, welche Interaktionen Überzeugungen untereinander eingehen bzw. nicht eingehen und welche Semantik sich für den Überzeugungsoperator am ehesten eignet. Die Einstellungen 'Wunsch' und 'Absicht' werden in Abschnitt 4.2 und in Verbindung mit Handlungen in Abschnitt 4.3 diskutiert.

Überzeugungen

Eine Überzeugung, dass p , eines Akteurs ist eine propositionale Einstellung. Der Gehalt einer Überzeugung ist eine Proposition, die der Akteur in der gegebenen Situation, in der der Zustand der Überzeugung vorliegt, für wahr hält. Somit wird eine Überzeugung als eine Relation zwischen einem Subjekt, einer Situation und einer Proposition verstanden.³⁷ Die Relation spiegelt wieder, dass das Subjekt α überzeugt ist, dass die Proposition p wahr in einer Situation s ist, d.h., dass in s ein Sachverhalt besteht, der der Proposition korrespondiert. Zudem ist die Überzeugung ein mentaler Zustand. Die Bildung der Überzeugung, das Beibehalten der Überzeugung und das Aufgeben einer Überzeugung (Ereignisse) sind von dem Zustand der Überzeugung verschieden, siehe Abschnitt 3.3.

In der doxastischen Logik werden die Wahrheitsbedingungen für das Haben der Überzeugung eines Subjektes angegeben. Es wird ein Operator für jedes Subjekt eingeführt, der beschreibt, dass α eine Überzeugung, dass p , hat.

Hintikka schlug vor, den Operator durch eine Relation zu interpretieren, die jedem Subjekt in jeder Situation eine Menge von Situationen zuordnet, die, mit dem was das Subjekt glaubt, kompatibel sind [81]. Das Subjekt besitzt der Interpretation zufolge genau dann die Überzeugung, dass p , wenn p wahr in allen doxastisch zugänglichen Situationen ist. Dies hat zur Folge, dass, falls dem Subjekt zumindest eine Situation doxastisch zugänglich ist, die Menge aller propositionalen Gehalte, die das Subjekt in einer Situation glaubt, konsistent sind.

Es ist jedoch nicht der Fall, dass, wenn ein Subjekt eine Überzeugung bildet, alle ihm doxastisch zugänglichen Situationen in die Überzeugungsbildung einbezogen werden. Vielmehr ist es der Fall, dass unsere Überzeugungen sich meist nur auf einen Teil der doxastisch zugänglichen Situationen beziehen, was wiederum zu Überzeugungen mit inkohärenten und einander sogar widersprechenden Gehalten führen kann.

Davon unterscheide ich Überzeugungen, die von Velleman als „perverse“ beschrieben werden und deren Gehalt selbst widersprüchlich ist. Vellemans Argument gegen die Unmöglichkeit einer solchen Überzeugung lautet wie folgt:

„The reason why one cannot believe such a proposition is not that one is incapable of accepting it. The reason is rather that the only way of accepting such a self-evident falsehood would be to accept it irrespective of its truth; and accepting a proposition irrespective of its truth wouldn't amount to believing it.“ [149, S.118]

³⁷Anstelle der Proposition kann ein nicht-transienter Sachverhalt gesetzt werden.

4. Überzeugungen, Wünsche und Absichten

Da eine Überzeugung die geglaubte Proposition dem Akteur als wahr darstellt und dabei so ausgerichtet ist, dass der Akteur mittels der Überzeugung immer ein wahres Bild der Welt erhält, ist die Bildung von Überzeugungen mit kontradiktorischem Gehalt nicht zielfördernd im Sinne, dass der Akteur nicht rational im Falle einer solchen Überzeugung sein kann. Welche Funktion hat eine Überzeugung mit kontradiktorischem Gehalt? Der Akteur hält den Gehalt der Überzeugung für wahr, obwohl er widersprüchlich ist.³⁸

Dies und die Unterscheidung zwischen konfligierendem und inkonsistenten Überzeugungen lässt mich zu dem Schluss kommen, dass inkonsistente Überzeugungen nicht rational möglich sind. Wenn ein Akteur glaubt, dass p und dass $\neg p$, liegen eigentlich zwei Überzeugungen vor, $(\alpha \text{ bel}:p \wedge \alpha \text{ bel}:\neg p)$, wenn der Akteur für das Halten seiner Überzeugungen rational ist. Es ist jedoch nicht rational möglich, die Überzeugung, dass p und nicht p zu haben, $\alpha \text{ bel}:(p \wedge \neg p)$.

Was ich im Folgenden begründen möchte, ist, dass ein Akteur auch dann rational sein kann, wenn er konfligierende Überzeugungen hat. Dies ist einer der Hauptgründe, in Überzeugungs-, Wunsch- und Intentionslogiken die relationale Semantik aufzugeben. In der von Rao und Georgeff vorgeschlagenen *BDICTL** ist sowohl für Wünsche als auch für Überzeugungen ausgeschlossen worden, dass diese konfligierenden Gehalt haben können. Auch wenn diese Ansicht für Akteure von vielen hinsichtlich Überzeugungen noch geteilt wird, ist dies im Fall der Wünsche zu restriktiv. Dies ist ein Grund die Semantik, in der die Wunsch- und Überzeugungsoperatoren interpretiert werden, zu hinterfragen und nach Alternativen zu suchen.

In der in [64, 65, 66] aufgestellten Logik *BDICTL** werden alle Operatoren, die sich auf die verschiedenen Einstellungen (Wünsche, Überzeugungen, Absichten) eines Akteurs zu einer Aussage φ beziehen, auf dieselbe Art und Weise behandelt. Sie werden in einer Semantik ausgewertet, die aus relationalen Strukturen besteht.³⁹ Sätze

³⁸Vellemans Argumentation wendet er nur auf transparente Widersprüche an, „belief cannot be transparently perverse. That is, one cannot believe a proposition that presents itself as false“ [149, S.118]. Wenn Überzeugungsgehalte nicht transparent widersprüchlich sind, wie z.B. es kann jemand glauben, dass Willy Brandt Kanzler war und Herbert Frahm kein Kanzler war, so handelt es sich möglicherweise nicht um inkonsistente, sondern um zwei konfligierende Überzeugungen. Vellemans Beispiel für einen „perverse“ Überzeugungsgehalt ist: „I am five inches taller than I really am.“ Ich würde „really“ streichen. Denn „really“ als ‘aktual’ verstanden laut Vellemans Fußnote würde diese Aussage als falsch identifizieren und wäre somit keine tatsächliche Inkonsistenz, sondern nur eine Überzeugung mit falschem Gehalt.

³⁹In Abschnitt 5.1 gehe ich auf die technischen Details ein. In diesem Abschnitt steht die Motivation von Semantik und Axiomen im Vordergrund.

4.1. Überzeugungen und Wünsche

der Form ‘der Akteur α glaubt, dass φ ’ sind in einer Situation (w, m) wahr, falls φ wahr in allen Welten w' zum Zeitpunkt m ist, die kompatibel mit dem sind, was α in (w, m) glaubt. Dasselbe gilt für die Intentionen und Wünsche eines Akteurs. Ein Akteur α wünscht bzw. beabsichtigt, dass p , in einer Situation (w, m) genau dann, wenn p wahr in allen Situationen (w', m) , die kompatibel mit dem sind, was α in der Situation (w, m) wünscht bzw. beabsichtigt. Dies ist die Übertragung der Vorschläge Hintikkas auf Wunsch- und Intensionsoperatoren.

Für Überzeugungen werden alle kompatiblen Welten in der Menge $B_m^w(\alpha)$ zusammengefasst oder als Relation gelesen: Die Welt w' steht in der Relation $B_m(\alpha)$ zu w . Was in der Situation (w, m) geglaubt wird, ist kompatibel mit jeder Situation (w', m) in der Welt w' . Es wird weder allgemein ausgeschlossen noch verlangt, dass die Welt w zu sich selbst in Relation $B_m(\alpha)$ steht. Es ist somit nicht notwendig, dass eine Überzeugung, die der Akteur α in einer gegebenen Situation (w, m) hat, auch in w wahr sein muss. Wenn die Menge $B_m^w(\alpha)$ als die Menge derjenigen Welten aufgefasst wird, die kompatibel zu dem sind, was α in w glaubt, bzw. die doxastisch zugänglich sind, so scheint es intuitiv strittig, dass ausgerechnet w , also die Welt auf der unsere Evidenzen beruhen, nicht zu dieser Menge gehört.

Würde man Kompatibilität auf diese Art verstehen, dass die Welt kompatibel mit unseren Evidenzen sein muss, müsste im Falle einer relationalen Struktur die Faktivität des Glaubensoperators folgen. Da wir sowohl Überzeugungen bilden, die mit der aktuellen Situation im Widerspruch stehen, als auch solche, die kompatibel mit der aktuellen Situation sind, ist die Forderung nicht haltbar, dass die aktuelle Situation in der Menge der Situationen ist, die kompatibel mit dem ist, was wir glauben. Die aktuelle Situation wäre nicht kompatibel mit unseren Überzeugungen, sobald der Akteur *eine* falsche Überzeugung hat. Dies scheint unplausibel. Die meisten oder zumindest einige Überzeugungen eines Akteurs sollten in der aktuellen Situation wahr und mit ihr kompatibel sein, wenn er rational ist, diese Überzeugungen zu haben. Es wird immer auch Überzeugungen eines rationalen Akteurs geben, die nicht korrekt sind. Warum sollte aufgrund einer nicht korrekten Überzeugung nicht darauf beharrt werden können, dass die aktuelle Situation kompatibel zu einigen, aber möglicherweise nicht zu allen Überzeugungen ist?

Ein anderer Punkt bei der Redeweise von zu den Überzeugungen kompatiblen Situationen ist, dass eine beliebige Überzeugung eines Akteurs in allen Situationen wahr sein muss, die mit dem, was der Akteur glaubt, kompatibel sind. Wenn $\alpha \text{ bel}:\varphi$ und $\alpha \text{ bel}:\psi$, muss $\varphi \wedge \psi$ in jeder Situation erfüllt sein, die kompatibel ist, mit dem was α glaubt. Demzufolge müssen alle Überzeugungen eines Akteurs konsistent sein oder

4. Überzeugungen, Wünsche und Absichten

die Menge der kompatiblen Situationen muss leer sein. Im letzteren Fall wäre jedoch der Akteur von allem überzeugt, was sicherlich nie der Fall ist. Intuitiv ist es rational möglich, Überzeugungen zu haben, die in völliger Übereinstimmung mit der aktuellen Situation stehen, und im selben Moment Überzeugungen zu haben, die dieser aktuellen Situation widersprechen.

So kann es der Fall sein, dass jemand davon überzeugt ist, dass alles, was in der Bibel steht, wahr ist, dass er glaubt, dass in der Bibel steht, dass Wasser in Wein verwandelt werden kann. Jedoch glaubt er, dass es nicht möglich ist, Wasser in Wein zu verwandeln. Dies ist offensichtlich möglich. Aber ist es für einen Akteur noch rational, solche Überzeugungen zu haben?

Ein Akteur kann glauben, dass Geld den Charakter eines jeden Menschen ruiniert, und gleichzeitig glauben, dass sein plötzlicher Geldgewinn seinen eigenen Charakter nicht ruiniert. Die Frage wäre: Kann es Situationen geben, die zu jeder dieser Überzeugung kompatibel sind? Es müsste eine Welt sein, in der sowohl 'Jesus verwandelt Wasser in Wein' als auch 'es ist nicht möglich, Wasser in Wein zu verwandeln' wahr sind. Ein weiterer Grund, sich von einer Menge von zu Überzeugungen kompatiblen Situationen zu verabschieden, ist der der Nichtentscheidbarkeit von Überzeugungen. So kann ein Akteur glauben, dass seine Entscheidung an bestimmter Stelle richtig gewesen ist, ohne dass es möglich wäre zu entscheiden, ob dies nun kompatibel zu seinen Überzeugungen ist, die er zudem in dieser Situation hat.

Ein Beispiel wäre die Goldbachsche Vermutung. Es ist im Moment nicht möglich zu entscheiden, ob sie kompatibel mit den Peano Axiomen ist oder nicht. Nehmen wir zwei Mathematiker an, die beide von der Gültigkeit der Peano Axiome überzeugt sind. Der eine glaubt jedoch, dass die Goldbachsche Vermutung falsch ist, der andere, dass sie wahr ist. Für einen von beiden sind seine Überzeugungen somit inkompatibel. Es gibt keine Situation, die beide Gehalte erfüllt. Die Menge der möglichen Welten, die dazu kompatibel sind, ist somit leer. Ergo müsste einer der Akteure in einem *BDICTL** Modell von allem überzeugt sein.

Somit stellt sich heraus, dass verschiedene, gleichzeitige Überzeugungen mit Situationen in Welten kompatibel sind, wobei die Welten untereinander es nicht sein müssen. Darauf zieht die Behandlungen des Begriffes 'Welt' und 'Kompatibilität' in der *bdi-stit* Logik ab. Überzeugungen beruhen auf Evidenzen, Erfahrungen, Konventionen, ... etc. Nun ist es möglich, dass nicht alle vorliegenden Evidenzen zu einer Überzeugungsbildung beitragen. Zumeist wird lediglich ein Teil der Evidenzen berücksichtigt, wenn ein Akteur eine Überzeugung bildet. Dann basieren verschiedene Überzeugungen zum selben Moment auf verschiedenen Situationen und Welten.

4.1. Überzeugungen und Wünsche

Eine Welt besteht dabei aus verschiedenen Situationen, die zu mindestens einer, aber nicht zu allen, Überzeugungen kompatibel sind.⁴⁰ Die Situationen müssen untereinander nicht mehr kompatibel sein. Es ist möglich, dass in zwei verschiedenen Welten zwei Situationen existieren, die Sachverhaltseigenschaften erfüllen, die einander ausschließen. Eine Situation, in der der Sachverhalt besteht ‘dass Jesus Wasser in Wein verwandelt’ und eine Situation, in der der Sachverhalt ‘dass Wasser in Wein verwandelt werden kann’ durch weitere bestehende Sachverhalte präkludiert wird, können somit beide in Welten liegen, die kompatibel zu dem sind, was der Akteur glaubt.

Der Begriff der ‘Welt’ in *BDICTL* und der Begriff der ‘Welt’ von der im folgenden Kapitel zu entwickelnden *bdi-stit* Logik sind verschieden.⁴¹ Ein Satz, der Form ‘der Akteur α glaubt, dass φ ’ ist genau dann wahr, falls es eine Welt gibt, deren sämtliche Situationen die Wahrheit von φ unterstützen. Die Welt wird als kompatibel mit einer Überzeugung verstanden und die Situationen in der Welt als dem Akteur doxastisch zugänglich. Ein Überzeugungssystem eines Akteurs wird durch eine Menge von Welten beschrieben, wobei eine jede Welt zu einer bestimmten Menge von Überzeugungen kompatibel ist.⁴²

Da es noch eine andere Welt geben kann, die kompatibel mit einer Überzeugung ist, deren Gehalt durch die Situationen in dieser Welt als wahr repräsentiert wird, ist es insbesondere möglich, dass in einer gegebenen Situation der Akteur sowohl φ als auch $\neg\varphi$ glaubt. Dies tritt ein, wenn zwei Welten existieren, wobei in einer Welt in allen Situationen φ wahr ist, in der anderen Welt hingegen in jeder Situation $\neg\varphi$ wahr ist. Kann dies für einen Akteur überhaupt rational sein? Ja. Ich werde zeigen, dass dies für Überzeugungen rational sein kann. Dies soll an weiteren Beispielen verdeutlicht werden, wie z.B. der Goldbachschen Vermutung. Im Falle von Wünschen stellt sich diese Frage nicht. Es ist der Fall, dass Akteure antagonistische Wünsche haben. Im Allgemeinen wird dies weder als rational noch als irrational bewertet.

Im folgenden Beispiel liegen zwei konfligierende Überzeugungen vor, ohne dass der Akteurin Rationalität abgesprochen wird. Eine Frau entscheidet sich aufgrund eines Kinderwunsches, auf ihre Karrierepläne zu verzichten. Jahre später sitzt sie im Kreise ihrer Familie und wird von einer ihrer Kommilitonin besucht, die mittlerweile

⁴⁰Ein solcher Ansatz wird von [55, 56, 146] verfolgt, um das Problem der Logischen Allwissenheit in der epistemischen Logik zu lösen, vgl. Abschnitt 5.1.

⁴¹Der Begriff ‘Welt’ ist nicht mit dem eines ‘Geschichtsverlaufs’ gleichzusetzen. Eine Welt ist eine Menge an Situationen, die eine bestimmte Einstellung des Akteurs wiedergeben, in dem sie den Gehalt der Einstellung als wahr repräsentieren.

⁴²Für die Umsetzung dieser Nachbarschaftsemantik siehe Abschnitt 5.2.

4. Überzeugungen, Wünsche und Absichten

eine erfolgreiche Geschäftsfrau ist und sich an gesellschaftlichem Ansehen erfreut. In dieser Situation glaubt die Frau mit einem Blick auf ihre Kinder, die sie sehr liebt, dass es die richtige Entscheidung war. Gleichzeitig bereut sie die Entscheidung. Sie glaubt, dass die Entscheidung falsch war, obwohl sie gleichzeitig glaubt, dass sie nicht falsch sei. Sie hat sich in dieser Situation sich widersprechende Überzeugungen. Sie glaubt jedoch nicht, dass die Entscheidung falsch und nicht falsch war. Dies wäre irrational. Verschiedene Erfahrungen und Situationen, in diesem Fall die sie liebenden Kinder auf der einen und der Erfolg der Freundin auf der anderen Seite, führen zu *konfligierenden* Überzeugungen.

Die Idee, Akteuren sich widersprechende Überzeugungen rational zu zuschreiben, ist nicht neu vgl. [92, 146, 158]. Es sollte jedoch zwischen der Zuschreibung zweier Überzeugungen, die sich widersprechen, oder der Zuschreibung einer widersprüchlichen Überzeugung unterschieden werden. In dem Aufsatz von Fagin und Halpern *Belief, Awareness, and Limited Reasoning* [55] wird ein Beispiel aufgegriffen, das der Physiker Eugene Wigner in [161] beschrieb. Die Quantenmechanik und die Allgemeine Relativitätstheorie basieren auf mathematischen Konzepten, die sich gegenseitig ausschließen. Beide Konzepte basieren auf bestimmten Annahmen, die, wenn man auf dem jeweiligen Gebiet arbeitet, als Konvention festgelegt sind. Unterschiedliche Konventionen können, wie im Fall der Physik, zu möglicherweise widersprüchlichen Überzeugungen führen. Die Physiker stellen sich die Frage, ob es möglich ist, diese beiden Theorien in einer gemeinsamen Theorie zusammenzubringen. Aber bis dato geht die Physik von zwei miteinander nicht vereinbaren Theorien aus. Fagin und Halpern sprechen von sogenannten „frames of mind“. Der Akteur oder der Physiker wird als eine „society of minds“ aufgefasst, wobei jeder „frame“ eine Menge von konsistenten Überzeugungen unterstützt. Zwei verschiedene Frames können jedoch konfligierende Überzeugungen unterstützen. In meiner Terminologie wäre eine solcher Frame im Modell durch eine Welt repräsentiert.

Die Idee der verschiedenen „frames of mind“ ist nicht neu und wurde in anderer Form von Thomas Kuhn [90] aber auch von Kazimierz Ajdukiewicz [3] diskutiert. Kuhn spricht von verschiedenen Paradigmen. Verschiedene Paradigmen, die auf der einen Seite die Relativitätstheorie auf der anderen Seite die Quantenmechanik beschreiben, können dabei Begriffe, wie Energie, Raum und Zeit verschieden behandeln, so dass sie innerhalb eines Paradigmas konsistent aber nicht inkommensurabel mit anderen Paradigmen sind. Ajdukiewicz behandelt dasselbe Phänomen allerdings spricht er von „Begriffsapparaturen“.

„... Wir wollen hier nämlich die Behauptung aufstellen [...], dass nicht

4.1. Überzeugungen und Wünsche

nur einige, sondern alle Urteile, [...] durch die Erfahrungsdaten noch nicht eindeutig bestimmt sind, sondern von der Wahl der Begriffsapparatur abhängen [...] Diese Begriffsapparatur können wir aber so oder anders wählen, [...] So lange sich jemand einer bestimmten Begriffsapparatur bedient, so lange wird ihm die Anerkennung gewisser Urteile von den Erfahrungsdaten aufgezwungen. Dieselben Erfahrungsdaten zwingen ihn aber nicht absolut zur Anerkennung dieser Urteile, denn er kann zu einer anderen Begriffsapparatur greifen, auf deren Boden dieselben Erfahrungsdaten ihn nicht mehr zur Anerkennung jener Urteile zwingen.“
[3, S.259]

Im weiteren Verlauf des Artikels erörtert er die klassische Physik und die relativistische Physik als Beispiele für verschiedene Begriffsapparaturen. Die Konventionen in der jeweiligen Theorie führen zu Behauptungen, die in der jeweils anderen Theorie zu Inkonsistenzen führt. Dennoch hält der Physiker an seinen Überzeugungen fest, die auf verschiedenen Konventionen basieren. Er wechselt nur seinen „frame of mind“ oder mit anderen Worten seinen Blickwinkel und den Hintergrund, auf dem seine Überzeugungen basieren.

Ein weiteres Beispiel für konfligierende Überzeugungen könnte ein Physiker im Opera-Experiment liefern. Einer der Wissenschaftler des Experimentes hat die Grundsätze der speziellen Relativitätstheorie verinnerlicht und glaubt felsenfest daran, dass es keine Teilchen gibt, die sich schneller als Licht bewegen. Insbesondere ist er davon überzeugt, dass sich Neutrinos nicht schneller als Licht bewegen. Auf der anderen Seite ist er nicht nur von der Korrektheit seiner Vorgehensweise, sondern auch von dem Versuchsaufbau und den Messinstrumenten und ihrer Korrektheit überzeugt. Ich nehme an, dass diese Überzeugungen nicht nur auf Wunschdenken beruhen, sondern gerechtfertigt darauf, dass er immer sorgfältig sich selbst, aber auch die Messinstrumente und die Arbeit seiner Kollegen überprüft hat.

Nach den ersten Messungen steht fest, dass die Neutrinos schneller als erwartet, die Strecke zwischen Cern und Gran Sasso zurückgelegt haben. Aufgrund seiner Überzeugung, dass nichts schneller als Licht ist, sammelt er Beweise bzw. untersucht den Versuchsaufbau nach Fehlern. Nach und nach schließt er eine mögliche Fehlerquelle nach der anderen aus. Unfreiwillig sammelt er immer mehr Evidenzen dafür, was er eigentlich widerlegen wollte, nämlich dass nichts schneller als Licht ist. Jedoch stellt er fest, dass die Messung korrekt ist und die Neutrinos so schnell waren, wie sie gemessen wurden. Er kommt aufgrund dieser Evidenzen zu der Überzeugung, dass die gemessenen Neutrinos schneller als Licht waren. Er kann jedoch seine Überzeugung,

4. Überzeugungen, Wünsche und Absichten

dass kein Teilchen schneller als Licht ist, nicht einfach aufgeben. Dieser Konflikt in seinen Überzeugungen ist jedoch nicht irrational. Irrational wäre er erst dann, wenn er sich auf seinen Überzeugungen ausruhen würde. Die Aufgabe eines Subjektes, das feststellt, dass es konfligierende Überzeugungen hat, ist sicherlich zu prüfen und zu reflektieren, worauf seine Überzeugungen basieren und nach weiteren Evidenzen für die Korrektheit der einen oder der anderen Überzeugung zu suchen.⁴³

Dies ist ein weiterer Grund anzunehmen, dass das alleinige Haben von konfligierenden Überzeugungen nicht irrational ist, da ein Subjekt durch das Haben von konfligierenden Überzeugungen gerade dazu angetrieben wird, nach weiteren Beweisen zu suchen und somit die Qualität seiner Überzeugung zu prüfen bzw. die Überzeugung selbst zu hinterfragen.⁴⁴ Ein solcher Antrieb hat in der Vergangenheit viele Wissenschaftler dazugebracht, umzudenken und neue Wege zu bestreiten.

Was sich in diesem Beispiel bereits andeutet, basiert auf der Unterscheidung von Kontrolle, die ein Akteur über seine Überzeugungen hat, aber vor allem aufgrund der Kontrolle, die er über seine Überzeugungen nicht hat. In Abschnitt 3.3 habe ich bereits diskutiert, dass, eine Überzeugung zu bilden, kein mentaler Akt eines Akteurs ist. Dass es sich bei den mentalen Ereignissen des Hervorbringens, Beibehaltens und Aufgebens einer Überzeugung nicht um mentale Akte handeln kann, begründet, warum Interaktionsaxiome zurückzuweisen sind, die postulieren, dass der Akteur dafür sorgen soll, eine bestimmte Überzeugung zu haben. Es gibt somit keine Interaktionsaxiome zwischen dem Überzeugungs- und dem Handlungsoperator, z.B. in der Form $\alpha \text{ bel}:\varphi \supset \alpha \text{ dstit}:\neg\alpha \text{ bel}:\neg\varphi$, dass eine Handlung, eine Überzeugung auszubilden, beschreibt. Auch wenn diese auf den ersten Blick rational erscheinen mögen, sind sie nicht sinnvoll.

Sie sind sicherlich auch keine konstitutiven Normen für Überzeugungen und können aufgrund von fehlender Kontrolle nicht als regulative Normen verstanden werden. In diesem Abschnitt werde ich kurz auf die Arten von Kontrolle über doxastische Einstellungen zu sprechen kommen, um festzuhalten, dass es nicht möglich ist, Überzeu-

⁴³So veröffentlichten Wissenschaftler im März 2012, dass es einen technischen Fehler gegeben hat. Die spezielle Relativitätstheorie ist nicht widerlegt. Und der Physiker hat, wenn er den Artikel gelesen hat, ein Paar konfligierender Überzeugungen weniger.

⁴⁴Die Qualität einer Überzeugung kann sich hierbei sicherlich nicht nur an der Wahrheit orientieren. Der Status der Rechtfertigung der Überzeugung, ob sie auf adäquaten Gründen basiert, ob sie kognitiv zugänglich sind, ob die guten Gründe auch derart sind, dass sie zur Korrektheit der Überzeugung führen etc. spielen hierbei sicherlich ebenso eine, wenn nicht sogar die größere Rolle. Ein Ansatz, was aus epistemischer Perspektive wünschenswert wäre, findet sich z.B. bei [5], was auf der anderen Seite nicht zu wünschen ist, z.B. auch in [102, Kapitel 5].

gungen direkt willentlich zu steuern.

Ein Beispiel für eine solche unmögliche Forderung in Form einer regulativen Norm habe ich eben genannt. Es ist nicht möglich, von einem Akteur zu fordern, dass er dafür Sorge, dass er keine konfligierenden Überzeugungen habe, da er keine direkte inhaltsgerichtete Kontrolle über seine Überzeugungen hat. Auch ein Axiom (D_b) für Überzeugungen $\alpha \text{ bel}:\varphi \supset \neg \alpha \text{ bel}:\neg\varphi$ ist somit nicht als regulative Norm zu verstehen. Dass es auch nicht als konstitutiv bzw. aufgrund deskriptiver Aspekte besteht, habe ich anhand der Fälle motiviert, in denen der Akteur konfligierende Überzeugungen hat ($\alpha \text{ bel}:\varphi \wedge \alpha \text{ bel}:\neg\varphi$). Um Schemata der Form $\alpha \text{ bel}:\varphi \supset \alpha \text{ dstit}:\neg\alpha \text{ bel}:\neg\varphi$ zurückzuweisen, werde ich auf die dafür fehlende Kontrolle bei der Überzeugungsbildung eingehen, um diese als mentalen Akt zu zählen.

Doxastische Kontrolle

Die Diskussion um den sogenannten Doxastischen Voluntarismus geht in Anfängen auf Descartes zurück, der noch als Befürworter der Kontrolle über Überzeugungen auftrat, über Hume [86],⁴⁵ Williams [162], der den doxastischen Voluntarismus als begrifflich unmöglich nachweisen wollte,⁴⁶ bis zu den Rückweisungen von direkter, voluntativer Kontrolle über Überzeugungen und die Unterscheidung verschiedener Kontrollarten, die hinsichtlich von Überzeugungen ausgeübt werden können. Diese finden sich z.B. bei Alston [4]. Ich werde die Unterscheidung aufgreifen, die Nottelmann in [102] vorgeschlagen hat.

Nottelmann zählt die verschiedenen Arten auf, wie doxastische Kontrolle ausgeübt werden kann. Er unterscheidet vier Paare voneinander ausschließenden Kriterien, wie Kontrolle über Überzeugungen ausgeübt werden könnte, und begründet, warum manche Arten selten einem Akteur zur Verfügung stehen. Seine Klassifizierung umfasst die folgenden Unterscheidungen [102, S. 87]

- 1. positiver und negativer,

⁴⁵„But as it is impossible, that that faculty [„imagination“ Anm.d.Verf.] can ever, of itself, reach belief, 'tis evident, that belief consists not in the nature and order of our ideas, but in the manner of their conception, and in their feeling to the mind.“ [86] (Appendix zu Section VII On the nature of the idea or belief). Wie Velleman scheint Hume 'Vorstellung' und 'Überzeugung' als zwei gleich gerichtete Einstellungen zu begreifen, die das Subjekt die Welt und ihre Objekte erfassen helfen („conceive“), die sich aber fundamental unterscheiden.

⁴⁶Eine Kritik und Widerlegung seiner Argumentation findet sich bei Winters [164]. Ein anderes Argument für die begriffliche Unmöglichkeit wird von Scott-Kakures geliefert [122]. Jedoch ist es ebenfalls zurückgewiesen worden, siehe z.B. [113].

4. Überzeugungen, Wünsche und Absichten

- 2. erzeugender und abbrechender,
- 3. direkter und indirekter,
- 4. inhaltsgerichteter und eigenschaftsbezogener

doxastischer Kontrolle. Nottelmann sieht in der Kontrolle, eine Fähigkeit etwas zu tun, um seine Überzeugungen zu beeinflussen. Die sechzehn Kategorien, die sich aus diesen Unterscheidungen ergeben, möchte ich nicht im Einzelnen aufzählen. Wenn die 3. und 4. Kategorie festgewählt wird, z.B. auf direkte (3.) und inhaltsgerichtete (4.) Kontrolle, werden die Unterschiede in den ersten beiden Kategorien deutlich [102, S.102]. Dann würde die vier entstandenen Klassen für eine feste Proposition p (inhaltsgerichtet) das Folgende charakterisieren. Positiv erzeugende Kontrolle entspricht der Fähigkeit die Überzeugung, dass p , direkt (willentliche Ausführung bzw. effektive Entscheidung zum Ausführen) zu formen. Negativ erzeugende Kontrolle entspricht der Fähigkeit willentlich von der Überzeugungsbildung, dass p , zurückzutreten. Positiv abbrechende Kontrolle ist demzufolge die Fähigkeit die Überzeugung, dass p , willentlich aufzugeben. Und die negativ abbrechende Kontrolle gibt die Fähigkeit an, die Überzeugung, dass p , willentlich nicht aufzugeben, sondern willentlich beizubehalten [102, S.102].

Da eine abbrechende Kontrolle nicht einmal im Falle der meisten Handlungen angenommen wird, würde ich auf diese vier Kategorien verzichten und nur die Bildung bzw. Erzeugung von doxastischen Einstellungen in Betracht ziehen. Zwischen positiver und negativer Kontrolle möchte ich sprachlich nicht ständig unterscheiden müssen, so dass ich statt von Überzeugungsbildung auch von der Bildung einer doxastischen Einstellung spreche. Doxastische Einstellungen zu p umfassen dann die Überzeugung, dass p , die Überzeugung, dass nicht p , und das Enthalten einer Überzeugung zu p .

Haben wir direkte oder indirekte Kontrolle über unsere Überzeugungen bzw. doxastischen Einstellungen? Nottelmann versteht direkte und indirekte Kontrolle wie folgt:

„I shall say that an agent *directly* controls the doxastic occurrences that she may, under some appropriate description of them, bring about simply by intending to bring them about. In other words, doxastic occurrences under direct control are those occurrences, which amount to *belief-influencing actions*. *Indirect* doxastic control is then naturally construed as doxastic control, which is not direct, e.g. the control we enjoy over those *doxastic consequences* of our belief-influencing actions, which are not themselves such actions.“ [102, S. 89]

Er unterscheidet somit direkte von indirekter Kontrolle, in dem er zwischen der Bildung einer doxastischen Einstellung in Form einer Handlung oder in Form einer Konsequenz einer Handlung unterscheidet. Direkte Kontrolle ist somit nicht nur die Fähigkeit irgendetwas zu tun, sondern tatsächlich hinsichtlich der doxastischen Einstellung die Fähigkeit zur effektiven Entscheidung.

Eine effektive Entscheidung „effective decision“ unterscheidet Nottelmann von einer bewertenden Entscheidung „value-decision“ [102, S.135]. Nottelmann meint „Deciding to believe“, z.B. von Ginet [69], wird von vielen im Sinne einer bewertenden Entscheidung interpretiert. Eine solche Entscheidung sei bei Überzeugungen jedoch keine effektive Entscheidung für die Überzeugung, sondern die Entscheidung, ob das Haben der Überzeugung von Wert sei. Dies führt jedoch nicht zu einer Überzeugungsbildung. Nottelmann vergleicht es mit effektiven bzw. bewertenden Entscheidungen für Handlungen. So wie das Bewerten, dass eine Handlung zu vollziehen gut wäre, nicht allein dazu führt, die Handlung tatsächlich zu vollziehen, führt auch eine bewertende Entscheidung zu glauben nicht dazu, dass die doxastische Einstellung angenommen wird. Ein Akteur hat somit die Möglichkeit, eine bewertende Entscheidung hinsichtlich seiner Bildung doxastischer Einstellungen deliberativ zu bilden. Dies ist jedoch keine Art der Kontrolle, die dazu führt, dem Akteur die Überzeugungsbildung als mentalen Akt zuzuschreiben.⁴⁷

Ob es ‘effektive Entscheidung’, ‘Intention’ oder ‘Handlungsmotivation’ genannt wird, das Argument ist analog. Die effektive Entscheidung zu einer doxastischen Einstellung ist die Entscheidung für das Hervorbringen dieser Einstellung. Wenn die effektive Entscheidung anders hätte ausfallen können, muss für das Hervorbringen im

⁴⁷In Anlehnung an Davidson ist für Nottelmann etwas eine Handlung, wenn es unter einer Beschreibung absichtlich ist [102, S.88]. ‘Absichtlich’ bei Nottelmann bedeutet, dass es unter *der* relevanten Beschreibung von dem Akteur intendiert wurde. Ein doxastisches Ereignis *direkt* zu kontrollieren, ist laut Nottelmann, die Bildung der doxastischen Einstellung im Falle einer inhaltsgerichteten Kontrolle absichtlich zu vollziehen bzw. die Eigenschaften (Qualität) der doxastischen Einstellung im Falle der eigenschaftsbezogenen Kontrolle absichtlich hervorzubringen. Im ersteren Fall heißt dies, die doxastische Einstellung zu intendieren. Eine effektive Entscheidung ist laut Nottelmann, die Intention zu haben, „*deciding to perform the relevant action in the sense of intending to perform it*“ [102, S.135]. Dies führt unwillkürlich in einen Regress wie in Abschnitt 3.3 gesehen. So vertritt Davidson auch nicht, dass intentional handeln, bedeutet, die Handlung auch zu intendieren [48, S.102]. Einer absichtlichen Handlung muss nach Davidson kein ‘intending’ vorausgehen. Es ist nach Davidson möglich, absichtlich zu handeln, ohne zu intendieren, so zu handeln [48, Essay 5]. ‘Intendieren’ ist nur ein Werturteil ähnlich zu ‘want’ bei Davidson. Zudem ist Davidson ein Vertreter der von Nottelmann kritisierten Belief-Desire Theorie von Intentionen [102, S.140]. Jedoch ist Nottelmanns Bezugnahme auf Intentionen nicht notwendig, um gegen direkte inhaltsgerichtete doxastische Kontrolle zu argumentieren.

4. Überzeugungen, Wünsche und Absichten

Akteur eine wie auch immer geartete Motivation (Überzeugung, Wunsch etc.)⁴⁸ vorliegen, warum er genau diese Einstellung zu p direkt hervorbringt. Aber was außer der doxastischen Einstellung, dass p , selbst motiviert den Akteur die Einstellung hervorzubringen? Eine Überzeugung kann nicht irrespektiv von der Betrachtung des Wahrheitswertes des Gehaltes gebildet werden. Die Motivation zu der Bildung der doxastischen Einstellung müsste wiederum eine Einstellung sein, die den Wahrheitswert betrachtet. Selbst wenn die konative Motivation hinter der Bildung der doxastischen Einstellung nicht auf den Wahrheitswert gerichtet ist,⁴⁹ muss z.B. im Falle der Überzeugung, dass p , nicht nur eine kognitive Einstellung in der Form vorliegen, dass der Akteur p als wahr erachtet, sondern der Zustand in dem der Akteur die Überzeugung bildet, muss kausal dafür verantwortlich sein, dass die Überzeugung gebildet wird.

Die Entscheidung muss also auf dem internalen Gesamtzustand des Subjektes basieren. Die Entscheidung kann jedoch nicht ursächlich auf den Gesamtzustand zurückzuführen sein, wenn der Akteur gleichzeitig sich entscheiden könnte, eine andere doxastische Einstellung anzunehmen, da diese Einstellung unter seiner direkten inhaltsgerichteten Kontrolle ist. Die Ursache der Wirkung $\alpha\text{bel}:p$ kann nicht Ursache der Wirkung $\neg\alpha\text{bel}:p$ sein. Also wäre eine Möglichkeit, dass der Akteur in Form einer Akteurskausalität diese Entscheidung herbeigeführt hat. Aber auch im Falle einer Akteurskausalität ist anzunehmen, dass der Akteur irgendwelcher Gründe bzw. Motivation für diese Entscheidung bedarf, da es sich sonst um eine reine Zufallsentscheidungen handeln würde. Es müsste also in demselben Gesamtzustand, der ursächlich und motivierend für die Entscheidung war, die Überzeugung, dass p , hervorzubringen, ebenfalls motiviert werden, die Überzeugung, dass nicht p , hervorzubringen, bzw. motiviert werden, die Überzeugung, dass p , aufzugeben.

Auch wenn keine begriffliche Unmöglichkeit nachgewiesen werden kann, ist dies höchst unplausibel. Derselbe Gesamtzustand, der den Akteur motiviert, eine Proposition als wahr (Bestehen des Sachverhaltes) anzusehen, müsste ihn gleichsam motivieren, sie als nicht wahr bzw. falsch anzusehen.⁵⁰

⁴⁸Eine Intention ist keine Motivation zu der intendierten Handlung, sondern die Entscheidung zur Handlung ist bereits gefällt. Eine Intention kann somit nur eine Motivation für eine weitere Handlung aufgrund einer Mittel-Zweck-Beziehung sein, siehe für eine ausführliche Diskussion Abschnitt 4.2.

⁴⁹Instrumentelle Gründe, weil der Akteur wünscht, dass es wahr ist (Wunschdenken), reicht für diese Fälle nicht aus. Ebenso würde eine bewertende Entscheidung nicht zu einer Handlung führen, selbst wenn der Akteur glaubt, dass es für das Vorliegen der doxastischen Einstellung nur der Bildung der doxastischen Einstellung bedarf, bringt er die doxastische Einstellung nicht hervor, vgl. die Diskussion solcher „trivialen“ Überzeugungen [102, S.141].

⁵⁰Wenn ein Akteur eine solche Form von Kontrolle hätte, wäre es somit auch unmöglich, dass er konf-

Auch wenn Nottelmanns Argumentation, dass direkte inhaltsgerichtete Kontrolle darüber beschrieben wird, dass man die Intention, eine bestimmte doxastische Einstellung zu haben, erfolgreich umsetzt, nicht nötig bzw. möglicherweise irreführend ist, zeigt die dahinter stehende Idee, dass eine inhaltsgerichtete direkte Kontrolle über die Bildung von doxastischen Einstellung psychologisch unmöglich ist. Das Argument von Nottelmann [102, S.140f] ist nicht bedroht. Jedoch sollte er auf den Begriff 'Intention' verzichten und nur von Entscheidungen sprechen. Ich werde im Folgenden darauf eingehen, dass das Fassen der Intention selbst eine Entscheidung ist, und ebenfalls wie Nottelmann das Belief-Desire-Modell für Intentionen ablehnen. Allerdings werde ich ablehnen, dass bei jeder absichtlichen Handlung eine Intention zu eben dieser Handlung vorlag.

Die weiteren Kategorien, die für die Kontrolle über doxastische Einstellung noch zur Verfügung stehen, sind für die Rechtfertigung der Interaktionsschemata, die ich diskutiere, aus den folgenden Gründen nicht relevant. Zum einen zeigt Nottelmann, dass die positiv und negativ erzeugenden indirekten eigenschaftsbezogenen Kontrollarten die einzigen beiden Kategorien der doxastischen Kontrolle sind, die rechtfertigen, das Überzeugungssystem eines Subjektes als rational zu bewerten. Somit ist es nicht möglich, einen Akteur als nicht rational zu bewerten, wenn er eine Rationalitätsforderung verletzt, die durch ein Axiomenschema beschrieben wird, dass er in der Situation dafür sorgen soll, eine bestimmte Überzeugung zu haben. Die Axiomenschemata, die zumeist eine oder mehrere Überzeugungen voraussetzen, aufgrund derer eine andere Überzeugung gebildet werden sollte, sind somit zurückzuweisen. Wenn eine bestimmte doxastische Einstellung (direkt) zu einer bestimmten Proposition (inhaltsgerichtet) gebildet bzw. nicht gebildet werden soll, so müsste der Akteur über direkte oder zumindest inhaltsgerichtete Kontrolle verfügen. Die „control over the properties of a belief relevant to the belief’s level of epistemic desirability“ [102, S.90] spielt eine untergeordnete Rolle.

Mangels der Kontrolle des Akteurs über seine Überzeugungen ist es nicht möglich, ein Axiomenschema, das Stit-Operatoren enthält, um die Überzeugungsbildung zu beschreiben, als regulative Norm zu verstehen. Es kann von niemanden etwas verlangt werden, das er nicht kontrollieren kann. Auch für doxastische Kontexte lege ich mich auf ein „Sollen impliziert Können“ [84] fest. Somit kann kein Axiom der Form $\alpha bel:p \supset \neg \alpha dstit:\alpha bel:\neg p$ begründet werden. Ein Akteur sollte nicht als nicht rational bezeichnet werden, nur weil er glaubt, dass p , und nicht dafür sorgt, dass er glaubt,

ligierende Überzeugungen hätte, da die Motivation zu einem mentalen Akt nicht die Motivation gegen den mentalen Akt sein kann.

4. Überzeugungen, Wünsche und Absichten

dass nicht p , da er keine Kontrolle darüber hat, die eine oder die andere Überzeugung direkt willentlich aufzugeben.⁵¹

Es ist nur noch zu klären, inwieweit indirekte inhaltsgerichtete Kontrolle auftritt und ob sie geeignet ist, Axiomenschemata, die die Rationalität von Akteuren widerspiegeln, zu motivieren. Bei indirekter inhaltsgerichteter Kontrolle unterscheidet sich wieder erzeugende und abbrechende von positiver und negativer Kontrolle. Vor Augen zu halten ist, dass es sich bei inhaltsgerichteter Kontrolle immer um einen bestimmten Sachverhalt handelt, der der Inhalt der Einstellung sein soll. Dass das Ausüben-Können von negativer erzeugender inhaltsgerichteter indirekter Kontrolle dazu führt, den Akteur für seine Überzeugungen verantwortlich zu machen, hält Alston fest.

„This suggests that even if propositional attitudes are not under our effective voluntary control, we might still be held responsible for them, provided we could and should have prevented them; provided there is something we could and should have done such that if we had done it we would not have had the attitude in question.“ [4, S.137f]

Aber wenn wir möglicherweise für Überzeugungen, die wir unter inhaltsgerichteter indirekter Kontrolle erlangen, verantwortlich gemacht werden,⁵² sind die Fälle, in denen ein Akteur tatsächlich inhaltsgerichtete direkte Kontrolle über seine Überzeugungsbildung verübt, so selten,⁵³ dass sie auch nicht hinreichen, ein Axiom der folgenden Form zu unterstützen: Wenn der Akteur glaubt, dass φ , sollte er irgendwie dafür sorgen, dass er nicht glaubt, dass nicht φ , $\alpha \text{ bel}:\varphi \supset \alpha \text{ dstit}:\neg\alpha \text{ bel}:\neg\varphi$. Auch Axiome, die eine bestimmte Implikation zwischen einer doxastischen Einstellung und dem indirek-

⁵¹Mittels temporaler Operatoren ist es möglicherweise beschreibbar, dass der Akteur, wenn er konfligierende Überzeugungen hat, in einem zukünftigen Moment dafür sorgen soll, dass er keine konfligierenden Überzeugungen mehr hat. Für ein und demselben Moment sind solche Schemata zurückzuweisen.

⁵²In [102] wird dies abgelehnt.

⁵³So würde wohl jeder vertreten, dass es möglich ist, sich entscheiden zu können zu glauben, dass das Licht an ist, wenn er es eben durch Betätigen des Lichtschalters selbst eingeschaltet hat [58, S. 672]. Die Kontrolle über diese Überzeugung wäre indirekt und inhaltsgerichtet. Man hat nicht direkte Kontrolle über den mentalen Zustand, der die Überzeugung beinhaltet, sondern Kontrolle über die Evidenz, dass das Licht an ist. Ist dies jedoch eine rationale Forderung, dass Akteure Sachverhalte allein mit der Intention hervorbringen, eine wahre Überzeugung bzgl. des hervorgebrachten Sachverhaltes zu bilden? Dies scheint nicht die Funktion einer Überzeugung zu sein. Eine Überzeugung soll die Welt dem Akteur abbilden. Es ist nicht Sinn und Zweck der Überzeugungsbildung, dass der Akteur die Welt so formt, dass er wahre Überzeugungen hat. Es ist auch kein Gegenargument, dass wir meist Wissen darüber haben, was wir tun. Die Handlung wurde selten bzw. fast nie mit der alleinigen Intention vollzogen, an das Bestehen des Zustandes nach der Handlung zu glauben.

ten bzw. irgendwie gearteten Hervorbringen einer weiteren doxastischen Einstellung fordern, sind nicht haltbar, um die Rationalität eines Akteurs zu beschreiben.

Zusammengefasst kann es für Überzeugungen bestimmte Kontrollarten hinsichtlich der Bildung von doxastischen Einstellungen geben, nämlich eigenschaftsbezogene indirekte Kontrolle, die im Sammeln von weiteren Informationen, der Steigerung der eigenen Konzentrationsfähigkeit bzw. dem Reflektieren über die eigenen Einstellungen bestehen kann. Ich kann Überzeugungen über Fische erlangen wollen und mir ein Fischbuch kaufen und lesen. Aber diese Form der Kontrolle über Überzeugungen lässt es nicht zu, Axiome zu postulieren, die für beliebige, aber feste Propositionen Zusammenhänge zwischen verschiedenen doxastischen Einstellungen als rational bzw. nicht rational postulieren.

Axiomenschemata für den Überzeugungsoperator

Im Folgenden werden einige Axiome und Regeln für Überzeugungsoperatoren kurz diskutiert, die bereits häufiger in verschiedenen epistemischen und doxastischen Logiken vertreten wurden, die keine Stit-Operatoren enthalten. Ich werde kurz begründen, warum ich viele dennoch nicht übernehmen möchte.

$$(K_b) \alpha \text{ bel}:(\varphi \supset \psi) \supset (\alpha \text{ bel}:\varphi \supset \alpha \text{ bel}:\psi)$$

Abgeschlossenheit unter geglaubter Implikation — Diese Form der Abgeschlossenheit ist oft abgelehnt worden, da sie erfordert, dass die Überzeugungen eines Subjektes abgeschlossen in dem Sinne sind: Wenn ein Akteur die Überzeugung hat, dass eine bestimmte Implikation gilt, deren Antezedens der Gehalt einer anderen Überzeugung ist, dann glaubt der Akteur den Sukzedens. Es ist sicherlich kein empirischer Fakt, dass ein Akteur alle seine geglaubten Implikationen abschließt. Im Gegenteil es ist eher die Regel bei normalen Akteuren, dass diese Abgeschlossenheit der Überzeugungen nicht vorliegt. Es wird daher häufig zwischen Formen von expliziten und impliziten Überzeugungen unterschieden. Im Falle der expliziten Überzeugung erfüllt das Subjekt diese Abgeschlossenheitsbeziehung seiner Überzeugungen nicht. Explizit wird als dem Akteur zugänglich gedeutet. Die Unterscheidung in explizite und implizite Überzeugungen möchte ich nicht vornehmen. Eine implizite Überzeugung ist quasi keine Überzeugung. Wenn aus einem Gehalt oder mehreren Gehalten von Überzeugungen folgen würde, dass der Akteur Grund hat, p für wahr zu halten, dann liegt diese Überzeugung, dass p , implizit vor. Tatsächlich liegt die Überzeugung nicht vor, sondern sie könnte vorliegen, da der Akteur hinsichtlich dessen, dass p , es als wahr

4. Überzeugungen, Wünsche und Absichten

erachten sollte bzw. könnte. Da ich den Überzeugungsoperator aber nicht als ‘der Akteur hat Grund dafür zu glauben, dass p ’ verstehe, sondern dem Akteur mit $\alpha \text{ bel}:p$ die Überzeugung zugesprochen wird, betrachte ich nur explizite Überzeugungen.

Das Schema (K_b) führt wie das äquivalente Theorem (A_b) zu Forderungen, die normale Subjekte nicht erfüllen. Das Theorem (A_b) fordert, wenn der Akteur von zwei Propositionen überzeugt ist, dass er auch von der Konjunktion überzeugt ist. Da ich am Anfang des Abschnitts argumentiert habe, dass Akteure konfligierende Überzeugungen nicht nur haben können, sondern dabei auch nicht ihre Rationalität verlieren, lehne ich (K_b) bzw. (A_b) ab. Für weitere Argumentation verweise ich auf (A_b) .

$$(D_b) \alpha \text{ bel}:\varphi \supset \neg \alpha \text{ bel}:\neg\varphi$$

Dieses Axiom ist aus bereits dargelegten Gründen nicht annehmbar. Rationale Akteure können konfligierende Überzeugungen haben.

$$(T_b) \alpha \text{ bel}:\varphi \supset \varphi$$

Die Faktivität von Überzeugungen ist unbestritten zu stark. Akteure besitzen falsche Überzeugungen. Und obwohl ein Akteur eine Überzeugung, dass p , nur bilden sollte, wenn er p für wahr hält, wäre eine Norm unangebracht, die dem Akteur verbietet zu glauben, dass p , wenn p falsch ist.

$$(A_b) (\alpha \text{ bel}:\varphi \wedge \alpha \text{ bel}:\psi) \supset \alpha \text{ bel}:(\varphi \wedge \psi)$$

Zu diesem Axiom hatte ich mich bereits bei (K_b) geäußert. Dass eine solche Forderung nicht rational ist, kann in gewisser Weise aus einer Argumentation von Hattiangadi und Bykvist vertreten werden [37, S.281]. Wenn die Rationalität eines Akteurs bezüglich Überzeugungen darauf ausgerichtet ist, dass der Akteur möglichst viele und vor allem korrekte Überzeugungen bildet, dann sind Schemata, die Überzeugungen als agglomerierend beschreiben, nicht zielförderlich.

Hattiangadi und Bykvist ziehen sogenannte *blindspots* in Betracht.⁵⁴ Blindspots sind Propositionen, die nicht wahrheitlich glaubbar sind, wie z.B. die Proposition, dass es regnet und niemand glaubt, dass es regnet. Man kann voneinander getrennte Überzeugungen haben, dass es regnet und dass niemand glaubt, dass es regnet. Aber die Konjunktion zu glauben, ist nicht wahrheitlich möglich. Sie beschreiben sie nicht nur als eine psychologische Unmöglichkeit, sondern auch als eine logische Unmöglichkeit. Da Überzeugungen auf Wahrheit gerichtet sein sollten, wenn der Akteur als rational bewertet wird, ist es nicht sinnvoll zu postulieren, dass zwei Überzeugungen,

⁵⁴Die Beschreibung von *blindspots* findet sich bei Sorensen [131].

4.1. Überzeugungen und Wünsche

die jede für sich wahrheitlich glaubbar sind, zu einer Überzeugung führen, die nicht wahrheitlich glaubbar ist.

Wenn der Akteur glaubt, dass es regnet, und er glaubt, dass niemand glaubt, dass es regnet, müsste er laut (A_b) glauben, dass es regnet und dass niemand glaubt, dass es regnet, oder eine der beiden Überzeugungen aufgeben. Die Überzeugungen zu glauben, dass es regnet, und zu glauben, dass niemand glaubt, dass es regnet, können sehr wohl in einem Akteur vorliegen. Sie sind auch nicht konfligierend, da ihre Gehalte konsistent sind. Es kann der Fall sein, dass es regnet, aber niemand glaubt, dass es regnet. Aber die Schlussfolgerung, dass ein Akteur immer die Konjunktion der Gehalte glauben sollte, nur weil die Gehalte konsistent sind, sollte wie dieser *blindspot* zeigt, abgelehnt werden. Denn in diesem Fall würde die Forderung dazu führen, dass eine Überzeugung mit falschem Inhalt geglaubt wird. Wird sie nicht geglaubt, sondern eine der beide vorliegenden Überzeugungen abgelehnt, z.B. dass niemand glaubt, dass es regnet, dann führt dies nicht zu weiteren korrekten Überzeugungen.

Hattiangadi und Bykvist wollten damit nicht zeigen, dass eine Rationalitätsforderung wie Axiom (A_b) widerlegt werden kann, sondern eine Korrektheitsnorm wie die folgende kritisieren: Wenn man sich fragt, ob p , dann soll man dann und nur dann glauben, dass p , wenn p wahr ist [37, S.282]. So kann man sich fragen, ob es regnet und niemand glaubt, dass es regnet. Es kann sogar wahr sein, dass es regnet und niemand glaubt, dass es regnet. Dennoch sollte man die Konjunktion nicht glauben. Sie vertreten dafür eine Annahme, die dem Schema (A_b) auf den ersten Blick ähnelt: Wenn jemand glauben soll, dass p , und glauben soll, dass q , dann soll er auch glauben, dass p und q . Dies scheint plausibel, jedoch benennen Hattiangadi und Bykvist es als nicht unkontrovers. Dieses Schema anzunehmen, ist jedoch nicht nötig, um mittels *blindspots* zu zeigen, dass (A_b) nicht angenommen werden muss. Da ich ausschließen würde, dass es für eine einzelne Proposition nicht der Fall ist, dass man die Proposition glauben soll, wäre dieses Prinzip trivialerweise erfüllt.

Bisher habe ich sowohl für (K_b) und (A_b) das Schema als eine Forderung verstanden, die an den Akteur gestellt wird, wenn er als rational für seine Überzeugungen bewertet werden soll. Somit würden sich diese Schemata nicht auf deskriptiven, sondern auf normativen Aspekten für Überzeugungen begründen. Sie beinhalten ein implizites Sollen. Der Akteur soll, wenn er glaubt, dass p , und glaubt, dass q , glauben, dass $p \wedge q$. Dieses implizite Sollen kann mit engem Skopus oder weitem Skopus verstanden werden. Wenn also ein solches Schema wie (A_b) zurückgewiesen werden soll, ist nicht nur das Schema mit enger Lesart des Sollens, sondern auch mit weiter Lesart zurückzuweisen.

4. Überzeugungen, Wünsche und Absichten

Broome [33] spricht von „detaching“ und „non-detaching“ Lesarten normativer Forderungen. Die enge Lesart wäre somit von (K_b) : Wenn ein Akteur glaubt, dass p und glaubt, dass $p \supset q$, dann sollte er glauben, dass q .⁵⁵ Dieses „sollte“ ist für manche ein Grund gewesen, einem Akteur zuzuschreiben, dass sein Überzeugungssystem die Schemata (K_b) bzw. (A_b) erfüllt.

Ich habe gezeigt, dass dieses „sollte“ mit engem Skopus von einem Akteur nicht zu verlangen ist, da er keine Kontrolle über seine Überzeugungsbildung hat. Somit ist es nicht als regulative Norm zu verstehen. Es kann nicht als implizite Norm verstanden werden, da *blindspots* zeigen, dass zwei nicht-konfligierende Überzeugungen vorliegen können, die zusammen nicht wahrheitlich glaubbar sind. Also sollte die Konjunktion nicht geglaubt werden. Dies würde jedoch (A_b) mit engem Skopus verlangen.

Hattiangadi und Bykvist betrachten (K_b) dann nochmal mit weitem Skopus der folgenden Form $\bigcirc(K_b)$. Für einen Akteur sollte es so sein: Wenn er glaubt, dass p , und er glaubt, dass $p \supset q$, dann glaubt er, dass q . Dies würde jedoch bedeuten, dass es so sein sollte, dass ein Akteur für feste beliebige Propositionen p, q nicht glaubt, dass p , oder nicht glaubt, dass $p \supset q$, oder glaubt, dass q . Hattiangadi und Bykvist halten folgendes Beispiel fest. Sei q die Menge aller wahren Propositionen. Dann ist es psychologisch für einen Menschen nicht möglich zu glauben, dass q .

Diesem Ansatz folgend ist eine Argumentation gegen $\bigcirc(K_b)$ denkbar. Wenn der Akteur den Begriff der materiellen Implikation verinnerlicht hat, ist es durchaus möglich, dass für irgendein wahres p der Akteur glaubt, dass p , und glaubt, dass dieses p jede wahre Proposition und somit auch die Konjunktion über alle Wahrheiten materiell impliziert. Diese Implikation kann ein Akteur glauben, ohne die jeden Sachverhalt der Konjunktion aller wahren Propositionen tatsächlich zu erfassen. Genauso wie er glauben kann, dass $a=a$ ist, ohne einen Begriff von ‘a’ zu haben. Beispielsweise kann ein Akteur glauben, dass supersymmetrische Teilchen genau dann R-paritätsverletzend zerfallen, wenn sie R-paritätsverletzend zerfallen, ohne dass er erfasst, was R-paritätsverletzend bzw. supersymmetrisches Teilchen bedeutet. Insbesondere muss er nicht glauben, dass supersymmetrische Teilchen R-paritätsverletzend zerfallen.

⁵⁵Gegen diese Form des Prinzips (K_b) haben Harman [73], Broome [33] und Wallace [156], indem er auf Broome verwies, hingewiesen. Broome stellt fest, dass die Überzeugung, dass p , bzw. die Überzeugung, dass $p \supset q$ nicht dazu führt, zu glauben, dass q , sondern dass es auch möglich ist, eine der ersten beiden Überzeugungen aufzugeben. Broome [35] liefert ein Beispiel. Ich glaube, dass es regnet, und ich glaube, dass, wenn es regnet, der Schnee schmilzt. Soll ich dann glauben, dass der Schnee schmilzt? Mittlerweile ist es Sommer.

4.1. Überzeugungen und Wünsche

Ein Akteur kann einen Begriff von materieller Implikation haben und darüber Überzeugungen bilden. Dennoch ist es ihm nicht möglich jede Implikation, die er glaubt, dass sie besteht, derart zu verstehen, dass es ihm ermöglicht, die implizit enthaltene Wahrheit von den propositionalen Gehalten seiner Überzeugungen zu glauben, wie z.B. die Konjunktion aller wahren Propositionen. Ich glaube, dass p die Konjunktion aller wahren Propositionen zu diesem Zeitpunkt impliziert, wenn ich glaube, dass p . Dennoch ist es mir nicht möglich, die Konjunktion alle wahren Propositionen zu glauben. Demzufolge habe ich nach $\circ(K_b)$ nur die Möglichkeit entweder meine Überzeugung, dass p die Konjunktion aller wahren Propositionen impliziert, oder jede andere korrekte Überzeugung aufzugeben. In diesem Fall ist $\circ(K_b)$ nicht zu halten und auch eine Variation in Form von $\circ(A_b)$ ist nicht haltbar. Wenn (K_b) bzw. (A_b) nicht durch deskriptive Zusammenhänge begründet sind und auch normative Aspekte von Überzeugungen diese Forderung als unplausibel darstellen, gibt es keinen Grund diese Schemata in einer Logik für Überzeugungen zu implementieren.

Im Folgenden werde ich die Axiomenschemata häufig nach normativen Gesichtspunkten beurteilen, da sie als psychologisches Phänomen bzw. aufgrund deskriptiver Aspekte sehr selten begründet sind. Dafür ist es jedoch notwendig, von weitem und engen Skopus des impliziten „Sollens“ zu sprechen, auch wenn in den Schemata selbst kein deontischer Operator benannt wird. Dies ist auch nicht nötig. Es geht hierbei nur um die Frage, ob diese Schemata für Operatoren durch normative Überlegungen begründet werden können. Es geht nicht darum, eine deontische Logik aufzustellen. Wenn weder ein deskriptiver Zusammenhang noch ein normativer Zusammenhang ausgemacht werden kann, werden die Schemata als nicht sinnvoll zurückgewiesen.

$$\begin{array}{l} (4_b) \alpha \text{ bel}:\varphi \supset \alpha \text{ bel}:\alpha \text{ bel}:\varphi \\ (I_b) \alpha \text{ bel}:\alpha \text{ bel}:\varphi \supset \alpha \text{ bel}:\varphi \end{array}$$

Positive Introspektion (4_b) und Infallibilität (I_b) – Ob und vor allem inwieweit ein Akteur privilegierten Zugang zu seinen eigenen Überzeugungen bzw. mentalen Zuständen im Allgemeinen hat, ist eine weitgefächerte Diskussion in der Philosophie des Geistes, auf die ich hier nicht näher eingehen kann. Obiges Axiom würde postulieren, dass ein Akteur von jeder seiner Überzeugung glaubt, sie zu haben. Dies wäre Schema (4_b). Die andere Richtung (I_b) besagt, dass er nicht von einer Überzeugung glaubt, dass er sie hat, wenn er sie nicht hat. Eine solche Infallibilität bezüglich seiner eigenen Überzeugungen bzw. mentalen Zustände postuliert kaum jemand. Selbst überzeugte Anhänger, dass es einen privilegierten Zugang zu den eigenen mentalen Zuständen gibt, geben zu, dass es möglich ist, sich zu irren [75].

4. Überzeugungen, Wünsche und Absichten

So vertreten viele, die der Meinung sind, dass man sich über seinen mentalen Zustände oder über einen bestimmten Teil seiner mentalen Zustände, beispielsweise Wahrnehmungszustände, nicht täuschen kann, dass es lediglich immer möglich sein muss, über Reflexion, Introspektion in Form von innerem Sinn etc. erfolgreich Zugang zu seinen Überzeugungen zu verschaffen.⁵⁶ Man muss sich dessen jedoch nicht stets bewusst sein und somit müssen die höherstufigen Überzeugungen nicht immer gebildet vorliegen.

Auf Ryles Regressargument war ich bereits im Abschnitt 3.3 eingegangen. Ryle [117] ist der Meinung, dass auch Selbsterkenntnis ein mentaler Prozess ist. Würde nun jemand postulieren, dass ein Subjekt zu jedem seiner mentalen niederstufigen Zustände einen mentalen höherstufigen Zustand mit dem Gehalt bildet, dass er in dem niederstufigen Zustand sei, würde jede Zustandsbildung in einen Regress führen. Hierbei verstehe ich höher- bzw. niederstufiger in dem Sinne von Überzeugung, dass er die Überzeugung hat, dass er die Überzeugung hat, usw. ‚dass p . Somit vertritt heutzutage niemand eine solch starke Infallibilitätsthese über das Erkennen und Vorliegen von Überzeugungen über die eigenen Zustände vertreten, wie (I_b) beschreibt.

Die These, dass für jeden mentalen Zustand das Subjekt erkennen kann, z.B. (4_b), ob es sich in dem mentalen Zustand befindet oder nicht, wird jedoch oft kritisiert. So wird die Transparenz, dass die eigenen mentalen Zustände dem Subjekt transparent sind, von Williamson verworfen. Er hat die These der Transparenz auf Luminosität abgeschwächt und anschließend ein Gegenargument geliefert. Luminos ist ein Zustand (Typ), wenn für ihn gilt, dass, wenn ein Subjekt in dem Zustand (konkret) ist, das Subjekt auch erkennen kann, dass es in dem Zustand ist. Williamson hat dann gezeigt, dass kein mentaler Zustandstyp luminos ist. Somit ist nicht jeder konkrete mentale Zustand dem Akteur auch transparent. Für dieses Argument siehe [163, Kapitel 4].

Somit werde ich weder die starke Infallibilitätsthese in Form von (I_b) noch die starke Selbsterkenntnisthese in Form von (4_b) vertreten. Ein Gegenbeispiel für die sogenannte positive Introspektion (4_b) wäre eine Situation, in der der Akteur glaubt, dass φ , aber nicht glaubt, dass er glaubt, dass φ . Es gibt zwei Arten der Erwiderung darauf.

⁵⁶Armstrong verglich die mentalen Zustände, die einem zwar gewahr sein können, aber erst bewusst werden, wenn man den inneren Sinn darauf richtet, mit einer Autofahrt. Man fährt quasi im „Autopiloten“. Ein Mann fährt, er sieht entgegenkommende Autos, er sieht den Straßenrand, er glaubt, dass Platz genug zum Überholen ist. Er ist sich dieser Zustände gewahr, denn er überholt und weicht den Autos aus. Dennoch glaubt er noch nicht, dass er in jedem dieser Zustände ist. Erst wenn er seinen inneren Sinn darauf richtet, was er sieht und was er glaubt, bildet er die höherstufigen Überzeugungen. Diese sind korrekt. Dieses Szenario von Armstrong ist nach [67] zitiert. In [67] sind verschiedene Ansätze, wie Selbsterkenntnis zu verstehen ist, dargestellt.

4.1. Überzeugungen und Wünsche

Zum einen besteht ein Konflikt mit Ryles Regress Argument. Zum anderen gibt es Überzeugungen, die ein Subjekt nicht haben möchte. Angenommen das Subjekt hat eine rassistische Überzeugung $\alpha\text{ bel}:\varphi_r$. Dann ist es dennoch möglich, dass es von sich glaubt, diese rassistische Überzeugung nicht zu haben $\alpha\text{ bel}:\neg\alpha\text{ bel}:\varphi_r$.⁵⁷ Wenn man als Vertreter der Infallibilitätsthese beide Richtungen und somit auch die Richtung (I_b) annimmt, erhält man den Widerspruch, dass das Subjekt die rassistische Überzeugung, dass φ_r , hat und nicht hat.

Ein weiteres Beispiel gegen die positive Introspektion findet sich in abgewandelter Form bei Gertler [68].⁵⁸ Kate hat eine langjährige Freundin, der sie vertraut, wenn es um ihre psychologischen Einstellungen geht, da diese sie sehr gut kennt. Diese Freundin äußert glaubhaft, dass sich Kate schon lange wünscht, auf dem Land zu leben. Kate glaubt ihrer Freundin. Aber die Freundin irrt. Kate ist sich zwar dessen nicht bewusst. Aber sie hat keinen Wunsch, auf dem Land zu leben. Dass es sich bei dem Gegenbeispiel gegen Introspektion um einen Wunsch und nicht um eine Überzeugung handelt, ist nicht von Bedeutung. Das Beispiel könnte wie folgt geändert werden. Kates Vater wird ein Mordprozess gemacht. Kate glaubt ungewollt und unbewusst aufgrund von Gerüchten und Äußerungen anderer, dass ihr Vater schuldig ist. Die Mutter, die felsenfest von der Unschuld ihres Mannes und von der Loyalität ihrer Tochter ausgeht, versichert Kate glaubhaft, dass Kate an die Unschuld ihres Vaters glaubt. Kate glaubt ihr nur zu gern und glaubt, dass sie glaubt, dass ihr Vater unschuldig ist. Somit glaubt Kate, dass er schuldig ist, und glaubt nicht, dass er unschuldig ist, obwohl sie glaubt, dass sie glaubt, er sei unschuldig.

Ein ähnlich geartetes Schema wie (4_b) kann ebenfalls durch ein solches Beispiel zurückgewiesen werden, $\alpha\text{ bel}:\varphi \supset \neg\alpha\text{ bel}:\neg\alpha\text{ bel}:\varphi$. Eine Freundin versichert Kate glaubhaft, dass sie die rassistische Äußerung der Lehrerin, dass φ_r , nicht glaubt. Kate glaubt, dass sie nicht glaubt, dass φ_r . Aber tatsächlich ist die Äußerung der Lehrerin auf fruchtbaren Boden gefallen und Kate glaubt, dass φ_r , auch wenn sie möglicherweise nicht wünscht, dass sie dies glaubt.

Dass ein Subjekt sich selbst über seine Überzeugungen täuschen kann, führt Velleman an [149, S.279f] und widerlegt somit (I_b). Präsident Reagan hat seine Geschichten

⁵⁷Dies wäre lediglich ein Gegenbeispiel gegen $\alpha\text{ bel}:\varphi \supset \neg\alpha\text{ bel}:\neg\alpha\text{ bel}:\varphi$ und noch nicht ausreichend gegen die positive Introspektion.

⁵⁸Ich habe das Beispiel Gertlers abgeändert. Im Original glaubte Kate, dass sie wünscht, auf dem Land zu leben, aber sie wünschte, nicht auf dem Land zu leben. Da sich Kate beides wünschen kann, auf dem Land zu leben und gleichzeitig sich wünschen, nicht auf dem Land zu leben, wäre dies kein Gegenbeispiel.

4. Überzeugungen, Wünsche und Absichten

über seine Zeit beim Film durch wiederholtes Erzählen immer weiter ausgeschmückt. Irgendwann glaubt er selbst, dass er glaubt, dass diese Geschichten wahr sind. Jedoch täuscht er sich, er glaubt es selbst nicht. Er glaubt zwar, dass er glaubt, dass seine Geschichte wahr ist. Tatsächlich glaubt er aber nicht, dass seine Geschichte wahr ist.

Zusammengefasst wird heutzutage selten bestritten, dass Subjekte sich über ihre Überzeugungen bzw. über ihre mentalen Zustände im Allgemeinen täuschen können. Daher werden die Axiome (I_b) – (4_b) zurückgewiesen. Bei den Beispielen ist auch nicht zu erkennen, warum die Subjekte im Annehmen der Überzeugung als nicht rational eingestuft werden sollten. Die Funktion eines mentalen Zustandes absolviert dieser im Allgemeinen unabhängig davon, ob der Akteur über das Vorliegen des mentalen Zustandes Überzeugungen bildet oder nicht. In jedem oben genannten Beispiel können die Rahmenbedingungen so zugespitzt werden, dass das Subjekt nicht anders glauben kann und dies für jede Stufe der Überzeugung.

$$(5_b) \quad \neg \alpha \text{ bel} : \varphi \supset \alpha \text{ bel} : \neg \alpha \text{ bel} : \varphi$$

Negative Introspektion — Dieses Axiom ist seit Hintikka [81] in vielen doxastischen Logiken zu finden. Es fordert nicht nur Reflexion über seine Überzeugungen bzw. Introspektion des Akteurs, sondern dass er auch Einstellungen über Propositionen hat, die er nicht glaubt. Dies ist unplausibel. Wenn ein Akteur keinerlei Zugang zu einer Proposition hat, somit keine doxastische Einstellung zu p bildet, ist es unplausibel, dass er glaubt, irgendeine doxastische Einstellung zu p zu haben.

$$\begin{array}{l} (RN_b) \quad \vdash \varphi \Rightarrow \vdash \alpha \text{ bel} : \varphi \\ (RM_b) \quad \vdash \varphi \supset \psi \Rightarrow \vdash \alpha \text{ bel} : \varphi \supset \alpha \text{ bel} : \psi \\ (M_b) \quad \alpha \text{ bel} : (\varphi \wedge \psi) \supset (\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : \psi) \end{array}$$

Logische Allwissenheit bzw. deduktive Abgeschlossenheit – Das Problem der logischen Allwissenheit war auch Hintikka bereits bekannt [82]. Seitdem ist immer wieder diskutiert worden, inwieweit sich dieses Problem lösen lässt.⁵⁹ Das Problem besteht hinsichtlich zweier Annahmen, die mit (RN_b) und (RM_b) auseinander zu halten sind. Die Regel der Necessitation (RN_b) besagt, dass jede Tautologie geglaubt wird. Die Regel der Monotonie (RM_b) besagt, dass die Überzeugungen unter logischer Implikation abgeschlossen sind. Falls der Akteur glaubt, dass φ , und aus $\varphi \psi$ allgemeingültig folgt, dann glaubt der Akteur, wenn er glaubt, dass φ , auch dass ψ unabhängig davon, ob der Akteur die Implikation glaubt. Der Akteur ist hinsichtlich seiner Überzeugungen deduktiv abgeschlossen. Daher spricht man auch von deduktiver Abgeschlossenheit. Die

⁵⁹Eine Auswahl an Ansätzen findet sich in [55].

Regel (RM_b) wird von vielen verworfen. Es ist nicht nötig, alle Tautologien zu glauben. Es gibt genügend Ansätze in der epistemischen bzw. doxastischen Logik, die eine solche Regel nicht beinhalten.

Die Regel (RM_b) ist diffiziler zu bewerten. Es gibt Ansätze, die diese Regel in doxastischen Logiken nicht beinhalten. Ich werde im Folgenden einen Ansatz benennen, für den ich jedoch argumentiere, dass er nicht weniger Probleme hat, als (RM_b) selbst aufwirft. Daran anschließend werde ich anhand eines Aufsatzes von Stalnaker dafür plädieren, dass (RM_b) eine akzeptable Idealisierung eines Akteurs in einer doxastischen Logik ist.

Ein Ansatz, bei dem sich die Regel (RM_b) als nicht allgemeingültig erweist, lässt zur Bewertung von Sachverhalten Situationen zu, in denen Sachverhalte bestehen, die sich gegenseitig ausschließen, sogenannte unmögliche Situationen bzw. Nicht-Standardsituationen [56, Abschnitt 9.4].⁶⁰ Ein solches Nicht-Standard-Modell von möglichen, unmöglichen Situationen unterteilt somit syntaktisch in zwei Arten nicht nur von Situationen, sondern auch von Implikationen — eine logische Implikation \Rightarrow , die jede Situation berücksichtigt, und eine Implikation \supset , die nur Standardsituationen in Betracht zieht. Dies führt dazu, dass es möglich ist, (RM_b) zu widerlegen, da es möglich ist, dass $\vdash \varphi \supset \psi$, aber $\not\vdash \alpha \text{bel}:\varphi \supset \alpha \text{bel}:\psi$. Das Problem der logischen Abgeschlossenheit der Akteure tritt erneut auf, sobald der Akteur für seine Überzeugungen nur Standardsituationen betrachtet. Die Regel (RM_b) wäre gültig. Eine Nicht-Standardsituation in einem Überzeugungssystem zuzulassen, bedeutet, dass der Akteur inkonsistente Überzeugungen haben kann. Dies hatte ich bereits abgelehnt. Die Funktion einer Überzeugung ist dem Akteur die Welt bzw. die Situation wahrheitsgetreu wiederzugeben. Dass ein solches Abbild in sich widersprüchlich sein kann und nicht von Rationalitätsforderungen zurückgewiesen wird, halte ich für unplausibel. Dass ein Akteur mehrere Abbilder der Welt hat, die zusammengenommen widersprüchlich sind, tritt wie im Beispiel des Physikers häufiger auf. Daher ist es rational

⁶⁰Da das Bestehen von Sachverhalten verlangt wird, stände dies im Widerspruch dazu, dass sich in einer Situation bestehende Sachverhalte nicht ausschließen dürfen, Abschnitt 2.1. Nun könnte argumentiert werden, dass es sich bei Nicht-Standardsituationen um rein abstrakte Situationen handelt, die im aktuellen Geschichtsverlauf nicht instanziiert werden können und die verschiedene Sachverhalteigenschaften enkodieren und dass in diesen Situationen nicht das Bestehen des Sachverhaltes als wahr, sondern das Enkodieren der Sachverhalteigenschaft, der Proposition, dass der Sachverhalt in dieser Nicht-Standardsituation besteht, den Wahrheitswert zuspricht. Dies würde die Korrespondenz einer Proposition mit dem nichttransienten Sachverhalt (ein Sachverhalt, der in einer Standardsituation besteht) aufheben. Dann würde jedoch die von mir verteidigte Terminologie in Kapitel 2 nicht mit einer Korrespondenztheorie der Wahrheit und einer Nicht-Standardsituationensemantik vereinbar sein.

4. Überzeugungen, Wünsche und Absichten

möglich, konfligierende Überzeugungen zu haben. Aber es ist irrational, inkonsistente Überzeugungen zu haben. Würde man ausschließen, dass nur Standardsituation, mit dem kompatibel sind, was der Akteur glaubt, wäre die Regel (RM_b) gültig.

Es gibt viele Ansätze, wie dem Problem der deduktiven Abgeschlossenheit zu begegnen ist, z.B. indem eine Art Gewähr-Sein einer Überzeugung hinzukommen muss, um als eine explizite Überzeugung zu gelten [55, Abschnitt 4]. Dies verhindert eine deduktive Abgeschlossenheit. Allerdings schließt ein solcher Ansatz ebenfalls aus, dass der Akteur beliebig deduktiv Überzeugungen bilden kann. Nach Stalnaker liegt hierbei das eigentliche Problem der logischen Abgeschlossenheit [135].⁶¹ Wieviel leitet eine Akteur aus seinen Überzeugungen her? Dass ein Akteur deduktiv Überzeugungen bildet, ist unbestritten. Wenn jemand glaubt, dass $2 + 2 = 4$, so ist dafür keine Schlussfolgerung nötig. Am Beispiel $3 * 16 = 48$ ist dies bei manchen Akteuren schon eher der Fall. Glaubte der Akteur, dass $3 * 16 = 48$, oder glaubt er es noch nicht? Wie lange ein Akteur braucht, bis er durch deduktives Schließen, Rechnen etc. soweit gelangt, dass er glaubt, dass $283 * 189 = 53487$, ist sicherlich ungewiss, aber dennoch ist es nicht unmöglich, dass er diese Überzeugung aus Überzeugungen über elementare, arithmetische Grundregeln herleitet. Aber nicht nur mathematische Gesetzmäßigkeiten, sondern jede Form des Schließens führt ein Akteur in bestimmten Rahmen aus, so dass eine Grenze in einem Modell zu ziehen, die dem Akteur eine Menge von festen Überzeugungen zuordnet, die von ihm nicht deduktiv vervollständigt werden können, nicht adäquat ist [135].⁶²

Eine doxastische Logik stellt immer eine Idealisierung der Akteure dar. Wie sollte sich das Überzeugungssystem eines ideellen Akteurs verhalten? Welche Überzeugungen werden gebildet? Wenn ein Akteur glaubt, dass φ , ist es sicherlich nicht möglich, von dem Akteur zu verlangen, dass er alle ableitbaren Konsequenzen ebenfalls glaubt. Begrenzte Ressourcen des Akteurs bilden für einige einen Grund, warum auch für einen ideellen Akteur das Überzeugungssystem nicht deduktiv abgeschlossen ist. Die-

⁶¹ Stalnaker bezieht sich auf deduktive Abgeschlossenheit, wenn er von logischer Allwissenheit spricht.

⁶² Stalnaker argumentiert anhand des „Storage-Sentence-Model“, das zwischen impliziten und expliziten Überzeugungen unterscheidet. Während implizite Überzeugungen deduktiv abgeschlossen sind, sind es explizite Überzeugungen nicht. Sie haben auch einen feinkörnigeren Gehalt als Propositionen. Zwei verschiedene Sätze können derselben Proposition zugeordnet werden, der Akteur muss jedoch nur einen davon glauben. Die Frage, die Stalnaker aufwirft ist: Was macht eine Überzeugung explizit und was macht eine Überzeugung implizit? Und welchen Vorteil hat es, dass implizite Überzeugungen logisch abgeschlossen sind? Wenn auch der ignoranteste Akteur als deduktiv abgeschlossen bezüglich seiner impliziten Überzeugungen betrachtet wird, welchen Wert hat dann eine deduktive Schlussfolgerung? [135, S.426] Eine Unterscheidung macht daher keinen Sinn.

4.1. Überzeugungen und Wünsche

ser Einwand kann sicherlich nicht entkräftet werden, aber möglicherweise ist er nicht relevant. Angenommen jemand hätte die Ressourcen und Kapazitäten alle erforderlichen deduktiven Schlüsse zu ziehen. Welche Schlüsse zöge er? Er bildet genau die Konsequenzen, die er für sein Interagieren mit der Welt braucht, er würde nicht notwendigerweise weitere Schlüsse ziehen, auch wenn er es von den Ressourcen her könnte [135]. Nur weil ich rechnen kann, impliziert dies nicht, dass ich jede wahre Proposition $x * y = z$ für beliebige Zahlen glaube.

Wenn über Ressourcen gesprochen wird, so ist diese Beschränkung ein Grund, bei der Beschreibung in logischen System eine Simplifizierung und Vereinheitlichung bei der Beschreibung der Akteure vorzunehmen. Es wäre außerdem eine Ressourcenverschwendung, eine Logik aufzustellen, die jedem Akteur in jeder Situation genau eine bestimmte Menge an Propositionen zuschreibt, die keine Axiomatisierung in irgendeiner Weise zulässt.

Somit benennt Stalnaker drei Gründe, die für eine Idealisierung und somit für eine Annahme von (RM_b) sprechen: Zum einen sollte man idealisieren, um Mechanismen hinter der Beobachtung beschreiben zu können. In diesem Fall ist (RM_b) sinnvoll, um zu beschreiben, dass Akteure deduktiv Überzeugungen bilden. Der zweite Punkt ist die Simplifizierung. Um etwas Komplexes abbilden zu können, muss man idealisieren. Den dritten Punkt, den Stalnaker anführt, ist der folgende.

Es wäre für jeden Akteur von Vorteil, alle ableitbaren Konsequenzen eines Gehaltes einer Überzeugung zu glauben [135], wenn er zum Ziel hat, möglichst viel wahre Überzeugungen zu bilden. Dies ist das eigentliche Ideal eines Überzeugungssystems. Stellt die deduktive Abgeschlossenheit nicht ein Ziel eines jeden Überzeugungssystems dar? Stalnaker führt als Gegenargument an, auf das er gleich widerspricht, dass man nur rational ist, wenn man auch an seine eigene Fabilität glaubt. Man soll glauben, dass man falsche Überzeugungen hat. Dies stellt jedoch kein Problem für die Regel (RM_b) dar, sondern nur ein Problem für die Abgeschlossenheit unter geglaubter Implikation (K_b) oder unter materieller Implikation. In der Regel (RM_b) steht nur die beweisbare logische Implikation zur Debatte. Würde ein Akteur gleichzeitig falsche und wahre Überzeugung haben, so tritt im Fall der deduktiven Abgeschlossenheit ein Konflikt auf, der den Akteur indirekt dazu bewegen könnte, seine Überzeugungen zu hinterfragen.

Im Falle von (K_b) kann eine solche Forderung des Abschlusses unter geglaubter Implikation zu einer Fülle falscher Überzeugungen führen, die alle kohärent und nicht konfligierend sind, aber kein Abbild der tatsächlichen Welt liefern. Dies ist bei (RM_b) nicht der Fall. Aufgrund der Allgemeingültigkeit der Implikation führt dies im Falle

4. Überzeugungen, Wünsche und Absichten

von falschen und wahren Überzeugungen irgendwann zu Konflikten, auch wenn die Menge an Überzeugungen nur teilweise an die Welt angepasst ist, werden falsche Überzeugungen schneller als solche deutlich.

So ist mein Punkt für (RM_b) nicht, dass diese Regel eine rationale Forderung an den Akteur darstellt, dass er seine Überzeugung deduktiv abgeschlossen halten sollte. Dies ist nicht möglich, da er solche Überzeugungen nicht direkt willentlich annehmen kann. Die Begründung ist vielmehr, dass es unabhängig davon, was er glaubt, einen guten Grund gibt, diese Überzeugung, die aus dem Gehalt einer anderen Überzeugung folgt, ebenfalls anzunehmen, da diese unter Umständen zu mehr wahren Überzeugungen bzw. eher zu Konflikten führt, die falsche Überzeugungen aufdecken können. Dies unterscheidet die Begründung für (RM_b) von den Ablehnungen der bisherigen Schemata. Es ist nicht nur eine Überzeugung des Akteurs, aufgrund dessen er eine weitere Überzeugung bilden bzw. aufgeben soll, sondern unabhängig vom Akteur gibt es einen Grund, Überzeugungen über die logischen Konsequenzen vorliegender Überzeugungen zu bilden. In Abschnitt 6.1 werde ich daher den Überzeugungsoperator mit (RM_b) und (F_b) axiomatisieren. Das Axiom (F_b) besagt, dass es nicht der Fall ist, dass Akteure Überzeugungen mit kontradiktorischem Gehalt haben. Die Regeln (RN_b) werde ich nicht übernehmen.

Im Folgenden gehe ich auf einige mögliche Interaktionsschemata ein, die ein Überzeugungsoperator eingehen kann, die ich zurückweisen werde. Ich hatte bereits festgehalten, dass es bei einer Handlung nicht notwendigerweise einer Überzeugung bedarf. Ebenso stellt eine Überzeugungsbildung keine Handlung dar. Ein Axiom der Form $\alpha \text{bel}:\varphi \supset \alpha \text{dstit}:\alpha \text{bel}:\varphi$ ist nicht sinnvoll und zwar nicht, weil die Überzeugung einen Zustand und kein Ereignis wiedergibt, sondern deshalb weil das Ereignis der Überzeugungsbildung kein mentaler Akt ist. Dies geht aus der fehlenden direkten inhaltsgerichteten doxastischen Kontrolle hervor, die ein Akteur über die Bildung doxastischer Einstellungen hat.

Ich hatte bei der Diskussion um die Introspektion bereits ein Beispiel genannt, warum ein Akteur sich nicht nur hinsichtlich seiner Überzeugungen, sondern auch seiner Wünsche täuschen kann. Kate hatte nicht den Wunsch, auf dem Land zu leben, aber sie glaubte, dass sie diesen Wunsch hat. Dies wäre ein Beispiel gegen (BDD) .

$$\begin{array}{l} (DBD) \alpha \text{des}:\varphi \supset \alpha \text{bel}:\alpha \text{des}:\varphi \\ (BDD) \alpha \text{bel}:\alpha \text{des}:\varphi \not\supset \alpha \text{des}:\varphi \end{array}$$

Transparenz von Wünschen (DBD) – Dass ein Akteur immer glaubt, dass er den Wunsch hat, wenn er ihn hat, lässt sich anhand eines Beispiels von Smith [129, S.46]

4.1. Überzeugungen und Wünsche

widerlegen. Smith ist dafür beide Schemata zurückzuweisen. Ich skizziere beide Gegenbeispiele gegen (*BDD*) und (*DBD*). John kauft jeden Tag seine Zeitung an einem bestimmten Stand, der nicht auf der kürzesten Strecke zu seinem Büro liegt. Er muss einen Umweg machen und dies aus keinem anderen Grund als dem, die Zeitung an genau dem Zeitungsstand zu kaufen. Die Zeitungen sind überall gleich teuer. John hält sich für keinen eitlen Menschen. Der Wunsch, nach Smith der Grund seines Verhaltens steht, ist jedoch, dass im Zeitungsstand ein Spiegel angebracht ist, und Johns ihm nicht transparenter Wunsch, sich immer wieder in dem Spiegel zu betrachten, dazu führt, dass er die Zeitung an diesem und keinem anderen Stand kauft. Würde man John auf diesen Wunsch hin ansprechen, würde er es vehement leugnen und nichts dergleichen in Betracht ziehen. Er ist überzeugt, dass er diesen Wunsch nicht hat [129, S.46].

An dieses Beispiel gegen (*DBD*) schließt Smith ein Gegenbeispiel gegen (*BDD*) an [129, S.47]. Ein junger Mensch glaubt, dass er sich wünscht, dass er Musiker wird. Jeden Tag übt er. Er lernt Noten. Er merkt, dass er üben will, dass er spielen möchte. Er spielt sehr gern seiner Mutter etwas vor, die sein größter Fan ist und felsenfest an die Karriere ihres Sohnes glaubt. Doch dann stirbt die Mutter und mit ihr der scheinbare Wunsch, Musiker zu werden. Denn eigentlich war es nie sein Wunsch, Musiker zu werden. Sondern der Wunsch, seiner Mutter zu gefallen, für sie zu üben, zu spielen, und ihren Wunsch zu erfüllen, dass er Musiker wird, hat ihn angetrieben. Dabei hatte er selbst nie den Wunsch gehabt, Musiker zu werden [129, S.47], sondern er glaubte immer nur, dass er den Wunsch hatte, Musiker zu werden.

Es gibt möglicherweise noch andere Vorschläge für Interaktionen zwischen Überzeugungs-, Wunsch- und Handlungsoperatoren. Ich habe die elementarsten und häufig vorkommenden Schemata zurückgewiesen. In [66] diskutieren Rao und Georgeff noch Beziehungen zwischen dem Vorliegen von Überzeugungen, Wünschen und Absichten, so z.B. dass aus einer Überzeugung, dass p , bereits folgt, dass p , bzw. wünscht, dass möglich, notwendig bzw. einfach nur p [66, S.321]. Solche Schemata sind durch eine geeignete Wahl der Relation bzw. Nachbarschaftssemantik über einer Klasse von Rahmen sicherlich als allgemeingültig zu beschreiben, aber schwer zu motivieren. Eine Überzeugung, dass p , beinhaltet nicht notwendigerweise einen Wunsch zu p . Bzw. schließt ein Wunsch, dass p , nicht aus, dass die Überzeugung, dass p , ebenfalls vorliegt.

Oft vertreten werden die Schemata $\alpha des:\varphi \supset \alpha bel:\neg\varphi$ und $\alpha des:\varphi \supset \neg\alpha bel:\varphi$. Man kann sich etwas nur wünschen, wenn man glaubt, dass es nicht bereits erfüllt ist, bzw. unsicher ist, ob es wahr ist [22, S.42]. Boghossian meint, dass man es sich sonst nicht wünscht, sondern sich nur darüber freut. Einige Vertreter des normativen

4. Überzeugungen, Wünsche und Absichten

Verständnis der „Passrichtung“ formulieren, dass ein Wunsch darauf abzielt, erfüllt zu werden [7, 123]. Dies scheint zu implizieren, dass der Wunsch beim Bilden noch nicht erfüllt ist. Dies kann zwei Gründe haben, entweder ein Wunsch kann sich nur auf Zukünftiges beziehen oder die Überzeugung innehaben, dass der gewünschte Sachverhalt noch nicht vorliegt, wie auch in [22]. Im letzteren Fall wäre letztgenanntes Schema begründet.

Dies ist jedoch, an sehr einfachen Beispielen zu widerlegen. Man kann sich wünschen, dass man gesund bleibt, auch wenn man glaubt, dass man gesund bleibt. Sicherlich ist man in diesen Fällen erfreut darüber, gesund zu bleiben. Man kann sich dennoch wünschen, gesund zu bleiben, auch wenn man glaubt, gesund zu bleiben. Ein Schema der Form $\alpha des:\varphi \supset \neg\alpha bel:\varphi$ ist nicht haltbar. Das Schema $\alpha des:\varphi \supset \alpha bel:\neg\varphi$ ist möglicherweise häufig erfüllt. Aber es spielt keine konstitutive Rolle für Wünsche. Es kann sein, dass ich etwas wünsche und gar keine doxastische Einstellung zu p habe, z.B. wenn ich fantasie, was wahr sein könnte.

Für den Überzeugungsoperator gibt es meines Erachtens keine erforderlichen Interaktionsaxiome mit den Wunsch- und Handlungsoperatoren. Dass es meines Erachtens auch keine Interaktion mit dem Intentionoperator gibt, dafür argumentiere ich in Abschnitt 4.3. Als Axiom und Regeln für den Überzeugungsoperator kommen nur die Regeln (RM_b) und das Axiom (F_b) in Frage. Da ich die Regel der Necessitation und andere Schema wie (A_b) ausschließen möchte, ist eine Interpretation in einer relationalen Semantik nicht möglich. Daher verwende ich für den Überzeugungsoperator eine Nachbarschaftssemantik. Dieselbe Semantik verwende ich für den Wunschoperator, dessen Eigenschaften im folgenden Abschnitt motiviert werden. Besonderes Augenmerk wird darauf gerichtet sein, Wünsche von Absichten zu unterscheiden.

4.2 Wünsche und Absichten

Dieser Abschnitt beschäftigt sich mit den Einstellungen ‘Wunsch’ und ‘Abscht’. Es wurde bereits mehrfach begründet, warum es sich bei Absichten um mentale Zustände handelt, die durch mentale Akte beschrieben wurden. Bei Wünschen handelt es sich im Allgemeinen um mentale Zustände, die sich aufgrund mentaler Vorgänge im Akteur bilden. Dieser Abschnitt ist den logischen Gesetzmäßigkeiten gewidmet, denen Absichten und Wünschen unterliegen bzw. unterliegen sollten. ‘Sollten’ bezeichnet hierbei, dass ein Akteur als rational für seine Wünsche und Absichten angesehen wird. Anhand der gültigen Schemata, die Absichts- und Wunschoperator erfüllen, lässt sich die Verwendung der jeweiligen Semantik für den Operator in der in Kapitel 6 zu postu-

lierenden *bdi-stit* Logik motivieren. Zuvor werde ich kurz auf die Rolle eingehen, die Wünsche und Absichten bei einer Handlung spielen. Diese Rolle ist einer der Gründe, Wünsche und Absichten als unabhängige mentale Zustandstypen anzusehen.

Antagonistische Wünsche und konfliktfreie Intentionen

Im letzten Abschnitt ist ein Beispiel angegeben, das zeigt, dass sich ein Akteur hinsichtlich seiner Wünsche täuschen kann. Es ist eine weit verbreitete Ansicht, dass es ebenfalls möglich ist, dass Wünsche in einem Akteur konfliktieren können.⁶³ Es gibt sicherlich kaum jemanden, der dies verneinen würde. In [144] stellen van der Hoek und Wooldridge fest, dass

“[i]mplemented *BDI* agents require that desires be *consistent* with one another, although *human* desires often fail in this respect”.

Dies ist zum Beispiel zu beobachten, wenn schwierige Entscheidungen anstehen. Ein Akteur hat die Möglichkeit, mit einer Spende einer Niere seinem Bruder das Leben zu retten. Es ist nur natürlich anzunehmen, dass in dem Akteur zwei Wünsche auftreten, der eine seinem Bruder die Niere zu spenden, um ihm das Leben zu retten, der andere, die Niere nicht zu spenden, um möglichst mit zwei gesunden Nieren weiterzuleben. Dass es die Rationalität eines Akteurs nicht bedroht, antagonistische Wünsche zu haben, hält z.B. Searle [123, S.136f] fest und formuliert weiter, dass es nicht die Wünsche sind, die aus Rationalitätsgründen nicht-konfliktierend sein dürfen, sondern die Absichten eines rationalen Akteurs.

Dies sind die Gründe, im folgenden Kapitel eine Logik zu postulieren, die Formeln als erfüllbar klassifiziert, die nicht nur konfliktierende Überzeugungen und antagonistische Wünsche, sondern auch konfliktfreie Intentionen beschreibt. Jedoch muss ebenso wie zwischen konfliktierenden und inkonsistenten Überzeugungen zwischen antagonistischen und paradoxen Wünschen unterschieden werden. Im obigen Beispiel hat der Akteur zwei Wünsche. Aber er hat nicht den Wunsch, die Niere zu spenden und nicht zu spenden. Antagonistische Wünsche wie auch konfliktierende Überzeugungen gehören zu verschiedenen Welten, die zeigen, was der Akteur hervorzubringen

⁶³In [61] diskutiert Frankfurt meherer Beispiele für konfliktierende Wünsche erster Stufe, z.B. den unwillig Drogenabhängigen, der seine Droge zu nehmen wünscht und nicht zu nehmen wünscht. Bratman in [25, S.20] möchte einen Nachttisch essen und möchte ihn gleichzeitig auch wieder nicht essen. Searle möchte einen ganzen Mittwoch lang in Sacramento sein und hat gleichzeitig den Wunsch denselben Mittwoch in Berkeley zu verbringen [123, S.137]. Beide sehen darin keinen Fehler bzw. keinen Widerspruch im Akteur. Sie betonen jedoch, dass dies für Intentionen nicht der Fall sein sollte.

4. Überzeugungen, Wünsche und Absichten

wünscht. Der Akteur wird hinsichtlich seiner Wünsche ebenfalls als eine „society of minds“ im Sinne von Halpern und Fagin verstanden [56]. Dabei muss insbesondere die Abgeschlossenheit⁶⁴ von Überzeugungen und Wünschen eingeschränkt sein. Im Falle der Überzeugungen ist gezeigt worden, falls ein Akteur die Überzeugung hat, dass φ , und die Überzeugung, dass ψ , so kann es keine Forderung an den Akteur sein, dass allein daraus folgt, dass er die Überzeugung, dass $\varphi \wedge \psi$, besitzen sollte. Sonst würde aus konfligierenden Überzeugungen folgen, dass ein Akteur eine paradoxe Überzeugung hat. Da Akteure antagonistische Wünsche jedoch keine paradoxen Wünsche haben können, muss eine solche Abgeschlossenheit für Wünsche abgelehnt werden.

Im Falle der Wünsche handelt es sich nicht um einen Konflikt im Akteur mit der Wirklichkeit, sondern eher um Gegenspieler, die im Akteur vorliegen, und darum konkurrieren, welcher handlungswirksam wird. D.h. welcher der Wünsche ist graduell stärker und kann eine Handlung oder eine Absicht motivieren. Im Falle von Wünschen spreche nicht von konfligierenden, sondern von antagonistischen Einstellungen. Ein Wunsch ist oft nicht nur eine konative Einstellung, sondern die Bildung des Wunsches war zumeist direktiv. Die Bildung des Wunsches geschah unter der Prämisse, dass der Akteur den Gehalt des Wunsches als etwas wahr zu kreierendes begriff.

Intentionen eines rationalen Akteurs verhalten sich anders als Wünsche und Überzeugungen. Es ist nicht rational möglich, dass ein Akteur beabsichtigt, dass φ , und gleichzeitig die Absicht hat, dass $\neg\varphi$. Der Grund hierfür liegt in der Tatsache, dass ein Akteur eine Absicht entwickelt, wenn er sich bereits aktiv für den Gehalt entschieden hat. Die These, dass Intentionen konfliktfrei sein sollten, wird von vielen vertreten. Die meisten berufen sich dabei darauf, dass es sich bei Intentionen um eine bestimmte Art von Überzeugungen handelt oder eine Überzeugung in einer Intention involviert ist, die dazu führt, dass die Gehalte der Intentionen nicht konfligieren sollten. Dies vertritt z.B. Harman [73] oder Wallace [156]. In [73] legt sich Harman fest, dass es sich bei einer Intention, um eine Überzeugung des Akteurs handelt, dass er eine bestimmte Handlung vollziehen wird. In [156] vertritt Wallace, dass es sich bei Intentionen nicht um eine Überzeugung handelt, aber dass eine Intention die Überzeugung beinhalten sollte, dass das Beabsichtigte auf gewisse Weise möglich sein sollte.

Für Velleman [150] beinhalten Intentionen, wie bereits erwähnt, eine Form kognitiver Verpflichtung, die ihn zu der Aussage veranlassen, Intentionen als eine Art Überzeugung zu bezeichnen. Searle formuliert, dass der Akteur im Falle der Absicht für die erfolgreiche Umsetzung der Handlung diese für möglich halten sollte [123, S.136],

⁶⁴Ein Problem der Logischen Allwissenheit und einer der Gründe für Halpern, Fagin u.a. sich für diese Form der Semantik zu entscheiden.

bzw. dass der Akteur, wenn er beabsichtigt, etwas zu tun, auch glauben muss, dass es ihm möglich ist, dies zu tun [123, S.55].

Wenn jedoch keine rationalen Schemata für Überzeugungen der Form (D_b) und (A_b) angenommen werden, ist es nicht möglich, die Rationalität der Schemata (D_i) $\alpha \text{ dint}:\varphi \supset \neg\alpha \text{ dint}:\neg\varphi$ bzw. (A_i) $\alpha \text{ dint}:\varphi \wedge \alpha \text{ dint}:\psi \supset \alpha \text{ dint}:(\varphi \wedge \psi)$ darüber zu begründen, dass es sich bei Intentionen um Überzeugungen handelt bzw. dass Intentionen notwendigerweise Überzeugungen involvieren. Dass die Rationalitätsforderungen an Intentionen sich aufgrund der Rationalitätsforderungen an Überzeugungen ergeben, bezeichnet Bratmann [27, S.30] als Kognitivismus.⁶⁵

Zum zweiten habe ich bereits im letzten Abschnitt dafür argumentiert, dass eine Intention zwar eine kognitive Einstellung ist, die im gebildeten Zustand die Akzeptanz beinhaltet, den beabsichtigten Sachverhalt hervorzubringen bzw. die beabsichtigte Proposition wahr zu machen. Aber sie beinhaltet nicht notwendigerweise eine Überzeugung, dass dieser Sachverhalt tatsächlich oder möglicherweise hervorgebracht wird. Allerdings ist es nicht möglich, aufgrund der beinhalteten Überzeugung zu fordern, dass Intentionen nicht konfliktieren dürfen. Es werden weitere Argumente benötigt, um Schemata der Form (D_i) und (A_i) zu untermauern. Bratman begründet in [27] ohne Rückgriff auf Überzeugungen, warum Intentionen nicht konfliktieren sollten bzw. warum ein Schema der Form (A_i) eine sinnvolle Forderung für Intentionen ist. Darauf werde ich in Abschnitt 4.3 zurückkommen.

Wünsche und Absichten - Antworten auf Warum? und Wie?

Im Folgenden werde ich den Unterschied zwischen Intentionen und Wünschen hinsichtlich ihres Gehaltes, ihrer Bildung und ihrer Funktion darstellen.

Im obigen Beispiel hat der Bruder die beiden konfliktierenden Wünsche bezüglich der Spende seiner Niere. In dem Moment, wo er die Absicht hat, die Niere zu spenden, hat er sich entschieden. Die Aussage ‘er hat die Absicht, die Niere zu spenden, und er hat die Absicht, die Niere nicht zu spenden’ wird als widersprüchlich angesehen oder zumindest auf die Weise, dass der Akteur sich eines Irrtums bzw. Fehlers schuldig macht [123]. Er sollte entweder die eine oder die andere Intention haben. Diese Entscheidung betrifft jedoch nicht seine Wünsche. Sie können weiterhin beide vorliegen,

⁶⁵Mit ‘Kognitivismus’ ist an dieser Stelle nicht die Position in der Metaethik [27, Fußnote 4] gemeint. In der Metaethik verneint Kognitivismus zumindest eine der zwei Behauptungen des Non-Kognitivismus, der besagt, dass moralische Urteile weder wahr noch falsch sind und demzufolge moralische Urteile auch keine Überzeugungen darstellen.

4. Überzeugungen, Wünsche und Absichten

ohne dass sich der Akteur eines Irrtums oder Fehlers schuldig macht. Ein grundlegender Unterschied zwischen Absichten und Wünschen ist die Forderung der Konfliktfreiheit von Intentionen im Gegensatz zu Wünschen. bzw. Überzeugungen. Während Wünsche und Überzeugungen uns mehr oder weniger zustoßen, entstehen Absichten erst nach Reflexion über den beabsichtigten Sachverhalt bzw. die beabsichtigte Proposition und unterliegen der Entscheidungsgewalt des Akteurs. Er übt eine Art von Kontrolle darüber aus, was er beabsichtigt, die ein Akteur über Überzeugungen und Wünsche nicht ausüben kann. Mit Rücksicht auf das „Sollen impliziert Können“-Prinzip sind Forderungen an die Rationalität des Akteurs nur zu stellen, wenn dieser der Forderung auch nachkommen kann.

Im Gegensatz zu einer indirekten eigenschaftsbezogenen Kontrolle über Überzeugungen und einer im Wesentlichen ähnlichen Kontrolle über unsere Wünsche, wird in der Literatur dem Akteur hinsichtlich seiner Intentionen mehr Kontrolle zugeschrieben. Zu der Kontrolle, die wir über Überzeugungen und Wünsche haben, schreibt Davidson in [48, Essay 4]

„The point isn't that desires and beliefs aren't ever in an agent's control, but rather that coming to have them isn't something an agent does.“ [48, S.73]

Somit bezieht sich Davidson auf die Bildung einer Überzeugung und eines Wunsches, die keine Form eines mentalen Aktes ist. Dennoch können diese eine Handlung rationalisieren. Überzeugung und Wunsch sind dann ursächlich für eine Handlung [48, Essay 1]. Im Falle einer Handlung liegt bei Davidson u.U. ein primärer Grund zur Handlung vor, der ursächlich für die Handlung ist, und der eine bestimmte Kombination aus Wunsch und Überzeugung darstellt [48, Essay 1-4]. Liegt ein primärer Grund vor, bezeichnet Davidson das Ereignis als absichtlich durch den Akteur hervorgerufen. Für Davidson besteht jedoch ein Unterschied zwischen „intentional“ (absichtlich) und „pure intending“ (beabsichtigen). Dies ist vielen zu passiv hinsichtlich des Akteurs, als dass es ihnen für eine Handlungsmotivation ausreicht. Velleman kritisiert an Davidsons Ansatz:

„[T]he story fails to include an agent – or, more precisely, fails to cast the agent in his proper role. In this story, reasons cause an intention, and an intention causes bodily movements, but nobody – that is, no person – does anything. [...] the person serves merely as the arena for these [psychological and physiological] events: he takes no active part.“ [149, S.123]

Wallace schreibt, dass eine weitverbreitete Annahme die folgende sei,

„[...] that motivations are not the kinds of states that are under our direct control. To be motivated in a certain way is to be subject to a desire of some kind, and desires in turn are psychological states with respect to which we are ultimately passive. In the terms of the essentially empiricist conception of motivation, desires may be considered the active principles in the human psychological economy, insofar as they provide an original impetus to action. But we are not active in regard to our desires. At the end of the day, whether or not we are subject to a given form of motivation is simply not up to us;“ [155, S.222].

Wallace selbst ist der Meinung, dass Wünsche zwar nicht unter unserer direkten Kontrolle liegen. Aber einem solchen ‘Ausgeliefertsein der Motivation’ beim Durchführen von Handlungen wie in eben genannten Zitat stimmt Wallace nicht zu. Motivationen für unsere Handlungen müssen derart sein, dass es möglich ist, mittels deliberativer Reflexion unsere Gründe bzw. Motivationen für unsere Handlungen direkt zu kontrollieren [155, S.219]. Daher kann eine Motivation nicht nur passiv sein. Neben Wünschen und Überzeugungen, die laut Wallace passiv sind, auch wenn sie möglicherweise durch bewertende Urteile entstanden sind, müssen ‘Volitionen’ teil einer jeden Handlungsmotivation sein.⁶⁶

Da Handlungen als aktiv durch das Subjekt hervorgebracht beschrieben werden, muss für Wallace eine Motivation aus drei Punkten bestehen:

- Wünsche, die zu dem Zeitpunkt in dem Akteur vorliegen. Diese beeinflussen den Akteur unterschiedlich stark [155, 157]. Er hat jedoch keine direkte Kontrolle über sie.
- Überzeugungen, was das Beste wäre zu tun, d.h. die praktischen Gründe und bewertenden Urteile. Auch über Überzeugungen hat ein Akteur nur bedingt Kontrolle. Nach Wallace können sich diese beiden Punkte sogar widersprechen

⁶⁶Wallace formuliert, dass sowohl Zustände als auch Ereignisse passiv sind. Meiner Meinung nach kann jedoch nur ein Ereignis passiv oder aktiv sein, je nachdem ob bei dem Ereignis ein Akteur involviert ist und ob das Ereignis dem Akteur als Handlung zugeschrieben werden kann oder nicht, siehe S.3.2. Zustände sind weder passiv noch aktiv. Es ist falsch anzunehmen, dass Zustände deshalb passiv sind, weil der Akteur nicht als aktiv beschrieben wird. Es gibt Zustände und Ereignisse, auf die weder die Beschreibung aktiv noch passiv zutrifft. Meines Erachtens weist die Kontrolle, die ein Akteur bei der Bildung, beim Aufgeben und Beibehalten der Überzeugungen und Wünsche nicht hat, auf das hin, was Wallace mit der Bezeichnung ‘passiv’ meint. Es ist jedoch ein Fehler ‘passiv’ mit dem Fehlen von Kontrolle gleichzusetzen. Etwas als ‘passiv’ zu bezeichnen, kann eine Form von Kontrolle beinhalten, die man jedoch möglicherweise nicht ausübt.

4. Überzeugungen, Wünsche und Absichten

[157], z.B. kann jemand urteilen, dass man kein Stück vom Kuchen mehr essen sollte. Jedoch besteht gleichzeitig der Wunsch fort, noch ein weiteres Stück zu essen [157, S.179].⁶⁷

- Volitionen, die den Akteur als aktiv kennzeichnen, da sie unter seiner direkten Kontrolle sind und eine Art der Proeinstellung bzw. selbstgewählter Wunsch sind. Diese unterscheiden seine volitionale Theorie von anderen Theorien der Motivation von Handlungen.

Unter 'Volitionen' fasst er Intentionen („intentions“), Entscheidungen („decisions“) und die Möglichkeiten zu wählen („choices“) [157, S.176]. Diese Volitionen sind weder auf Wünsche noch auf bewertende Urteile und Überzeugungen noch auf eine Kombination aus beiden zurückzuführen. Obwohl Wallace Volitionen als Proeinstellung auffasst und er daher vertritt, dass ein Akteur immer das tut, was er am meisten will („truistic dictum“) [157, S.172], sind Volitionen im Gegensatz zu Wünschen unter der direkten Kontrolle des Akteurs [157]. Es sind nicht nur Zustände bzw. bloße Vorgänge, wie z.B. Wünsche, die einem Akteur zustoßen. Volitionen sind primitive Phänomene dessen, was wir unter Handlungen verstehen [157, S.176]. In [155] formuliert Wallace, dass der Akteur nicht nur ein Subjekt seiner Wünsche sein kann, wenn er als rational bzw. irrational bewertet werden soll, sondern „we will need, beyond that, to establish that A had the power to generate a desire of this kind directly“ [155, S.236]. Nach Wallace sind Handlungserklärungen für Fälle von Akrasia ein Hinweis darauf, dass nicht nur Wünsche und Überzeugungen ursächlich für unsere Handlungen sein können, da wir uns offensichtlich entgegen unseren Wünschen und Überzeugungen entscheiden können. Dies macht eine inhaltsgerichtete oder direkte Kontrolle darüber notwendig, was wir tun und was wir intendieren [155, 157].

Somit sind für Wallace Intentionen nicht auf Überzeugungen und Wünsche zurückführbar. Dass er sie gleichzeitig als Zustände und primitive Phänomene von Handlungen beschreibt, kommt meiner Interpretation nahe. Eine Intention zu fassen, ist ein mentaler Akt. Das Vorliegen der Intention kennzeichnet ein Ereignis, das ein bestimmter Zustandstyp in diesem gewissen Zeitraum besteht, der sich nicht auf eine

⁶⁷Wallace gesteht zu, dass diese zwei Punkte oft zusammenfallen. Nagel vertritt, dass motivierende Gründe durch praktische Überlegungen hinsichtlich der eigenen zukünftigen Wünsche bzw. durch moralische Urteile motiviert sein können [101], welche nicht notwendigerweise durch zum Zeitpunkt der Handlung präsente Wünsche abgedeckt sein müssen. Für Wallace bedeutet das jedoch, „this variant gives us a nonaccidental connection between practical judgement and motivation, but at the considerable cost of depriving practical judgement of its potential autonomy from given desire“ [157, Fußnote 15 S.175].

Kombination aus Wünschen und Überzeugungen zurückführen lässt. ‘Direkte’ Kontrolle über den Zustand hat ein Akteur meines Erachtens nur insoweit, dass er eine inhaltsgerichtete Kontrolle beim Fassen der Intention ausübt, die er bei der Bildung von Überzeugung und Wunsch nicht hat. Eine inhaltsgerichtete direkte Kontrolle wird jedoch unabhängig von den Wünschen und Überzeugungen eines Akteurs für Intentionen als problematisch erachtet. Ich gehe daher nicht soweit zu sagen, dass der Akteur zu einer Intention etwas von den in ihm vorliegenden mentalen Zuständen kausal Unabhängiges beisteuern muss. Eine Intention zu fassen, wird ebenso wie eine Handlung durch den vorliegenden mentalen Gesamtzustand des Subjekts motiviert. Die Kontrolle beim Fassen der Intention ist ebenso inhaltsgerichtet wie die Kontrolle über seine Handlungen.⁶⁸

Wallace hat recht, wenn er Volitionen von Wünschen und Überzeugungen aufgrund der Kontrolle unterscheidet, die wir über die einen haben und über die anderen nicht haben. Aber meines Erachtens irrt Wallace, wenn er meint, dass Volitionen zur Handlungsmotivation hinzukommen müssen. Wenn sie selbst primitive Phänomene von Handlungen sind, können sie nicht Teil einer Handlungsmotivation sein, da sie selbst motiviert werden müssten. Alternativ kann man Akteurskausalität annehmen, was mir unplausibel erscheint. Unabhängig davon gibt es einen weiteren Grund anzunehmen, dass Volitionen bzw. zumindest Intentionen keine Form einer Handlungsmotivation sind.

Eine Handlungsmotivation sollte eine Antwort auf die Frage geben, warum der Akteur dies oder jenes getan hat. Wie Searle vollkommen zurecht bemerkt [123, S.138],

⁶⁸Ein Problem, welches Wallace volitionale Theorie erklären muss, ist, was eine Handlung bzw. Volition verursacht. Wenn ein Akteur eine Volition direkt unter seiner Kontrolle hat und diese nicht durch seine mentalen Zustände und Ereignisse verursacht werden, die in dem Akteur vorliegen, dann vertritt Wallace eine Art der Akteurskausalität, wie sie bspw. Chisholm vertreten hat in [43, S.70] bzw. in [40]. Dies ist nicht notwendig. Bratman unterscheidet Wünsche und Überzeugungen von Intentionen, ohne sich auf Akteurskausalität zu berufen. Er veranschlagt im Rahmen einer „Planning-Theory“, die nicht nur Zustände zu dem gegebenen Zeitpunkt in Betracht zieht, sondern den indirekten Einfluss des Akteurs über einen längeren Zeitraum beschreibt. Dabei ist es möglich, Einstellungen zu seinen Intentionen einzunehmen, die selbstreferentiell sind [23, Kapitel 10]. Fasst der Akteur zu einem Zeitpunkt den Plan, ein bestimmtes Ziel zu erreichen, ist es meistens der Fall, dass der Plan noch nicht vollständig ist. Bratman schlägt vor, solche Pläne für das Erreichen von Intentionen einzuführen, indem ein höherstufiger Wunsch Ziel des Planes ist. So ist es möglich, eine Rechtfertigung für den gewünschten Inhalt bzw. den Wunsch selbst zu erwerben. Dies wäre meines Erachtens in Nottelmanns Terminologie eine Form der inhaltsgerichteten indirekten Kontrolle, die wir über Intentionen hätten. Bratman nennt diese Form der Kontrollmöglichkeit selbst-regulierende Strategien. Damit wären Intentionen ebenfalls nicht zurückführbar auf Wunsch- und Überzeugungskombinationen, wie sie die Standardtheorie vorschreibt.

4. Überzeugungen, Wünsche und Absichten

würde aber niemand auf die folgende Frage mit der Erwähnung der Absicht antworten: Warum hat er den Arm gehoben? Weil es seine Absicht war, den Arm zu haben. Vielmehr hatte er einen Grund, das zu tun, also die Absicht zu formen, den Arm zu heben. Wenn ich den Arm hebe, hebe ich den Arm. Wenn ich beabsichtige, den Arm zu heben, dann motiviert mich etwas zu dieser Absicht. Aber nicht die Absicht motiviert mich, den Arm zu heben. Es ist umgekehrt, was mich motiviert, den Arm zu heben, das motiviert mich auch, die Intention zu bilden, den Arm zu heben.⁶⁹

Wünsche und Intentionen unterscheiden sich nicht nur, über die Art und Weise wie sie gebildet werden und ob Kontrolle bei der Bildung des jeweiligen Zustandes ausgeübt wurde. Sie können Antworten auf verschiedene Fragen hinsichtlich der Handlungen eines Akteurs geben. Während Wünsche bzw. Ziele eines Akteurs Antworten auf die Frage geben, warum der Akteur eine bestimmte Handlung vollzogen hat bzw. warum er diese Handlung beabsichtigt, gibt der Gehalt der Intention eine Antwort darüber, *wie* der Akteur die Handlungen vollziehen wird. Warum hebst du den Arm? Ich möchte den Herrn gegenüber grüßen. Wie grüßt du den Herrn? Ich beabsichtige, den Arm zu heben.

Ob es sich um eine Absicht bzw. um einen Wunsch handelt, lässt sich bereits am Gehalt des jeweils vorliegenden Zustandes erkennen. Der propositionale Gehalt von Absichten bezieht sich auf (zukünftige) Handlungen [123, S.138].⁷⁰ Eine Absicht beinhaltet immer eine Absicht zum Handeln, wohingegen Wünsche völlig losgelöst von den Handlungen eines Akteurs auftreten. Beispielsweise kann ich wünschen, dass es schneit. Aber jemand, der behauptet, dass er beabsichtigt, dass es schneit, würden wir als nicht rational bezeichnen oder als jemand, der das Konzept 'Absicht' bzw. 'schneien' nicht verstanden hat.

Im Gegensatz zu einem bloßen Wunsch unterscheidet sie, dass ein reines Erfüllt-Sein des beabsichtigten Sachverhaltes nicht ausreicht, um die Absicht tatsächlich als erfüllt anzusehen. Eine Absicht ist erst erfüllt, wenn der Sachverhalt mit der beabsichtigten Handlung hervorgebracht wurde. Die implizit beabsichtigte Handlung muss zu dem Erfüllt-Sein des Sachverhaltes führen, um eine Absicht als erfüllt zu betrachten.

⁶⁹In [7, § 51] diskutiert Anscombe, dass jemand auf die 'Warum'-Frage antworten kann, weil es seine Absicht war. Dies zeigt nach Anscombe jedoch nicht, dass es nicht eine andere Motivation hinter dieser Absicht gab, sondern dass derjenige gerade nur an seinem gegenwärtigen Tun und nicht an der Frage, warum er es tut, interessiert ist. Dennoch hat er nach Anscombe, wenn er es absichtlich tut, eine Motivation hinter dieser Absichtlichkeit.

⁷⁰Der Grund, den Searle angibt, ist die kausale Selbstbezüglichkeit des Gehaltes einer Intention. Ich kann nur beabsichtigen, was ich mit der Absicht verursachen kann bzw. was meine Einstellungen mir als möglicherweise hervorzubringen darstellen.

Dies zeigen die folgenden Beispiele.

Beispiel eins geht auf Chisholm zurück [40, S.37]. Der Neffe möchte seinen Onkel töten. Er steigt mit der Absicht in sein Auto, zu seinem Onkel zu fahren und ihn dort mit einer Axt zu töten. Er ist so fokussiert auf sein Ziel, dass er auf seiner Fahrt zu dem Onkel unabsichtlich einen Menschen überfährt. Dieser Mensch stirbt und es stellt sich heraus, dass dieser sein Onkel war. Was ist erfüllt? Das Ziel bzw. der Wunsch, seinen Onkel zu töten, ist sicherlich erfüllt. Was jedoch nicht erfüllt ist, ist die Absicht, seinen Onkel zu töten. Die Absicht, seinen Onkel zu töten, beinhaltet, wie er beabsichtigte, seinen Onkel umzubringen, nämlich ihn mit einer Axt zu erschlagen. Auch wenn sein Ziel erreicht ist, das zum Formen der Absicht führte, so ist die Absicht nicht erfüllt und sie hat auch nicht ursächlich zum Erreichen des Zieles beigetragen. Die beabsichtigte Handlung hat nicht stattgefunden.

Meine Folgerung wäre, dass für eine Absicht, den Onkel zu töten, der Wunsch vorliegt, den Onkel zu töten. Aber eine Absicht beinhaltet mehr als den Wunsch bzw. das Ziel, den Sachverhalt hervorzubringen. Auch die Überzeugung, dass eine bestimmte Handlung ausreicht, den Onkel zu töten, und den Wunsch zu haben, den Onkel zu töten, reichen nicht aus. Der Neffe hätte in diesem Beispiel sogar wissen können, dass, wenn er seinen Onkel überfährt, dieser stirbt, und es sich wünschen können. Dennoch hätte er den Mann immer noch unabsichtlich überfahren. Für das Vorliegen einer Absicht sind ein Wunsch und die Überzeugung, wie das gewünschte Hervorzubringen ist, nicht hinreichend. Eine Absicht beinhaltet zudem die Handlung (das Ereignis) wie der Sachverhalt hervorgebracht werden soll.⁷¹

In [43] verwendet Chisholm ein anderes Beispiel. Ödipus hat den Mann mit Absicht erschlagen, aber er hat nicht absichtlich seinen Vater erschlagen. Für mich ist dieses Beispiel kein Hinweis darauf, dass Absichten ursächlich sein können und dennoch die Handlung nicht absichtlich ist. Es ist vielmehr so, dass der Gehalt einer Absicht ein Sachverhalt bzw. eine Proposition ist. Der Sachverhalt steht aber für das dadurch

⁷¹Chisholm möchte mit diesem Beispiel zeigen, dass Absichten ursächlich für eine Handlung aufgrund äußerer Umstände und nicht vorhersehbarer kausaler Zusammenhänge sind, die Handlung selbst jedoch nicht absichtlich sein muss. Um etwas als absichtlich vom Akteur hervorgebracht zu zählen, muss der Akteur wählen, dies zu tun. Das Beispiel zeigt dies jedoch nicht. Die Absicht, den Onkel zu töten, war nicht die Ursache für dessen Tod, obwohl Searle dem zustimmt. Aufgrund des Zieles, den Onkel zu töten, fasste der Akteur den Plan, zu dem die Absicht gehörte, den Onkel zu töten, aber auch die Absichten, ins Auto zu steigen, loszufahren und ihn mit der Axt zu erschlagen, gehörten dazu. Die Absicht, den Onkel zu töten, ist dafür nur mittelbar ursächlich. So hätte der Akteur unterwegs die Absicht, den Onkel zu töten, aufgeben und stattdessen beabsichtigen können, mit ihm zu reden. Er wäre immer noch absichtlich zu ihm gefahren und hätte unabsichtlich den Onkel überfahren.

4. Überzeugungen, Wünsche und Absichten

charakterisierte Ereignis der beabsichtigten Handlung. Während des Ereignisses, das der Handlung zugrunde liegt, bestehen mehr als nur dieser Sachverhalt, der als Gehalt der Absicht im Akteur vorhanden ist. Die Handlung, seinen Vater zu töten, ist dieselbe, wie den Mann zu töten, nur mit dem Unterschied, dass der erste Sachverhalt nicht beabsichtigt war und der zweite schon. Davidson nennt das, was ich mit Sachverhalt bezeichne, *unter einer Beschreibung*.⁷²

Dass ein Wunsch und eine Überzeugung nicht ausreichen, um dem Akteur eine Absicht bei einem Ereignis zu unterstellen, soll folgendes Beispiel von Searle zeigen. Eigentlich stammt das Beispiel von Davidson [48, Essay 4 S.79]. Dieser wollte damit zeigen, dass es kausale nicht-normale Abweichungen geben kann, die erklären, warum jemand eine Überzeugung und Proeinstellung haben kann und diese das gewünschte Ereignis verursachen, aber dennoch kein primärer Grund für eine Handlung vorliegt. Laut Davidson [48, Essay 1 S.4f] liegt ein primärer Grund für eine Handlung nur dann vor, wenn eine Proeinstellung zu Handlungen mit einer bestimmten Eigenschaft und eine überzeugungsartige Einstellung, dass diese Handlung diese Eigenschaft hat, vorliegen. Ein solcher primärer Grund würde die Absichtlichkeit einer Handlung zeigen und gleichzeitig die Ursache für diese Handlung sein. Wenn keine Handlung vorliegt, liegt kein primärer Grund vor, auch wenn möglicherweise passende Überzeugungen und Wünsche vorliegen, die jedoch keine Handlung verursachen.

Ein Kletterer hat den Wunsch, sich aus der Gefahr zu befreien, und weiß, dass er dafür nur das Seil loslassen muss, an dem sein Partner hängt. Dieser Wunsch und diese Überzeugung zermürben ihn, so dass sie ursächlich sind, dass er das Seil *aus Versehen* loslässt. Er kann sogar wünschen, das Seil loszulassen. Dieser Wunsch kann sogar ursächlich zum Loslassen beigetragen haben, weil der Akteur, erschrocken über das Haben dieses Wunsches, das Seil loslässt. Aber nach Searle muss zudem eine Handlungsabsicht vorliegen. Die Wünsche und Überzeugungen reichen für eine Absichtlichkeit des Loslassens nicht aus [123, S.143]. Davidson spricht ebenso von unabsichtlich. Es ist gar keine Handlung des Akteurs, auch wenn er Wunsch und

⁷²Siehe bereits die Ausführung in Abschnitt 3.3. Wie Searle korrekt bemerkt, ist dies jedoch keine sprachliche Unterscheidung [123, S.133], vgl. Searles Bemerkungen zu Sinn und Unsinn von *Beschreibungen* [123, S.123f]. So spricht Searle von *Repräsentation* der Handlung in einer Absicht. Und weiter „selbst die unabsichtlichen Handlungen (Beispiel: Ödipus heiratet seine Mutter) gibt es nur deshalb, weil es ein identisches Ereignis gibt, das eine absichtlich vollzogene Handlung ist (Ödipus heiratet Jokaste)“ [123, S.111]. Beide Sachverhalte charakterisieren jedoch das Ereignis, dass Ödipus seine Mutter heiratet, auch wenn Ödipus nicht beide beabsichtigt hat. Er hat jedoch dafür gesorgt, seine Mutter bzw. Jokaste zu heiraten. Er entschied sich für das Ereignis ‘zu heiraten’ und vollzog die Handlung zu heiraten. Unbeabsichtigt sorgte er für den Sachverhalt, dass er seine Mutter heiratete.

Überzeugung hatte, da diese kausal nicht ursächlich für die Handlung sind, das Seil loszulassen. Es liegt somit kein primärer Grund, nach Davidson keine Absichtlichkeit, und keine Handlung vor [48, Essay 4 S.79f].

Es stellen sich zwei Dinge heraus. Zum einen folgt aus dem Vorliegen einer bestimmten Überzeugung und eines Wunsches nicht, dass auch eine Absicht vorhanden ist. Zum anderen beinhaltet die Absicht nicht nur, den Sachverhalt hervorzubringen, sondern impliziert bereits, dass und wie der Akteur für ihn sorgen wird, siehe Abschnitt 4.3. Laut Searle und Davidson ist der Akteur ursächlich an dem Bestehen des Sachverhaltes beteiligt, wenn er es absichtlich getan hat. Demzufolge beinhaltet eine Absicht auch immer, die Absicht ursächlich für den Sachverhalt zu sorgen.

Wünsche und Überzeugungen liefern Gründe für die Handlung eines Akteurs. Es besteht jedoch keine unter deskriptiven und normativen Aspekten bestehende notwendige Beziehung zwischen den Wünschen und Überzeugungen eines Akteurs und seinen Handlungen. Wenn keine kausalen Beziehungen zwischen den Überzeugungen und Wünschen eines Akteurs und seinen Handlungen in der Form bestehen, dass eine jede Kombination aus Wunsch und Überzeugung eine Handlung nach sich zieht bzw. ziehen sollte, so kann auch keine solche Beziehung zwischen diesen Überzeugung und Wünschen auf der einen und seinen Intentionen auf der anderen Seite bestehen. Für Davidson ist dies nicht korrekt, da er in einem primären Grund die Absichtlichkeit sieht, die ein Ereignis als Handlung des Akteurs kennzeichnet [48, Essay 1 S.7].

In [48, Essay 5] spricht Davidson dann von Formen des „pure intending“. Dies sind zukunftsgerichtete bzw. vorausgehende Absichten. Er stellt fest, dass diese durch seinen Begriff von ‘absichtlich handeln’ bzw. ‘mit der Absicht handeln’ nicht abgedeckt werden können und dass für diese Intentionen gilt, dass es sich um eine Form von Urteil handelt, dass es wünschenswert ist, eine bestimmte Handlung zu vollführen [48, Essay 5]. Einer Intention muss keine Überzeugung zugrunde liegen, dass man die Handlung durchführen wird bzw. dass sie möglich ist. Jedoch sollte eine Intention immer mit den Überzeugungen konsistent sein, die das Subjekt über das hat, was ist, und das, was sein wird [48, Essay 5 S.100]. Davidson hält für Intentionen fest, dass sie zwar keine Form von „wanting to do“ bzw. „wish“ sind, aber für eine Intention das Vorliegen einer Proeinstellung in Form eines Wunsches („desire“) bzw. eines Wollens („want“) konstitutiv ist [48, Essay 5 S.102]. Indem er eine Intention als ein Urteil auffasst, dass eine Handlung wünschenswert ist, und gleichzeitig als etwas, für das ein Wunsch konstitutiv ist, gibt er Vellemans Charakterisierung als eine direktive und kognitive Einstellung recht, obwohl Davidson explizit widerspricht, dass es eine Form von Verpflichtung ist, dass ich etwas intendiere [48, Essay 5 S.100].

4. Überzeugungen, Wünsche und Absichten

Wie Davidson betont, sind die Absichten eines Akteurs nicht seine Wünsche. Sie stehen jedoch in engem Zusammenhang. Bei einer Absicht handelt es sich um eine konative Einstellung und um die Verpflichtung zu einer Handlung, die nach Davidson von dem Akteur als wünschenswert erachtet wird. Der Akteur sollte die Handlung auch wünschen, die er beabsichtigt. D.h. er sollte den die Handlung charakterisierenden Sachverhalt wünschen, den er beabsichtigt hervorzubringen. Die Absicht, dass φ , sollte den Wunsch, dass φ , beinhalten.

$$(DI) \alpha \text{dint}:\varphi \supset \alpha \text{des}:\varphi$$

Vertreter der Auffassung, dass sich Intentionen auf Wünsche und Überzeugungen zurückführen lassen, sind mit (DI) zumeist einverstanden. So formuliert Audi [8], der eine Intention als eine Kombination von bestimmten Wunsch- („want“) und Glaubeinstellungen auffasst:

„The core, but only the core of the idea is simply this: one intends to do something, say to decline a drink, when one wants to do it more than one wants to do anything one believes is incompatible with doing so, and one believes at least that one probably will do it.“ [8, S.18]

Auch wenn ich nicht zustimme, dass bei einer Intention eine Überzeugung vorliegen muss, dass man etwas tun wird, unterstützt Audi (DI), dass es sich bei einer Intention um eine Einstellung handelt, der der Wunsch zu einer Handlung inhärent ist. Der Ausdruck ‘mehr zu wünschen als alles andere’ zeigt, dass es Sachverhalte gibt, die nach Audi mit dem intendierten Sachverhalt bzw. dem Ergebnis der Handlung in Widerspruch stehen können.

Beispiele, die scheinbar zeigen, dass Absichten kein Wunsch zugrunde liegt, sind solche, die unter Zwang oder aus Rücksicht bezüglich anderer Wünsche entstehen. So könnte in obigem Beispiel jemand behaupten, der Bruder hat nicht den Wunsch, eine Niere zu spenden. Er hat lediglich die Absicht und vollzieht sogar die Handlung, jedoch nicht aufgrund des Wunsches, sondern eines Pflichtgefühls oder unter dem Erwartungsdruck seiner Familie. Dieser Einwand verkennt meines Erachtens die Einstellung auf erster Stufe mit einer Einstellung auf zweiter Stufe. Wenn er die Ab-

$$(4_d) \quad \alpha \text{des}:\varphi \supset \alpha \text{des}:\alpha \text{des}:\varphi$$
$$(IDI) \quad \alpha \text{dint}:\varphi \supset \alpha \text{des}:\alpha \text{dint}:\varphi$$

sicht hat, die Niere zu spenden, so hat er den Wunsch. Was er jedoch nicht haben muss, ist den Wunsch, diesen Wunsch oder diese Absicht zu besitzen. Nur weil ich

einen Wunsch oder eine Absicht habe, impliziert dies nicht diesen Wunsch bzw. diese Absicht auf höherer Stufe zu wünschen. Wird eine Handlung unter Zwang vollzogen, muss sich der Akteur nicht wünschen, sie zu beabsichtigen. Aber auch einer erzwungenen Absicht liegt ein Wunsch zugrunde. Hält mir jemand eine Pistole an den Kopf und zwingt mich, jemanden zu erschießen. So hatte ich die Entscheidung zu schießen oder nicht zu schießen. Ich formte, die Absicht zu schießen. Diese konative Einstellung basiert auf dem Wunsch, überleben zu wollen. Sie ist ursächlich für den Wunsch zu schießen, da ich mich entschied, mich zwingen zu lassen oder nicht.⁷³

Es ist nicht rational, sich etwas zu wünschen, ohne sich die Mittel zu wünschen, die für die Umsetzung nötig sind. Es ist jedoch möglich, gleichzeitig den Wunsch zu haben, diese Mittel nicht einzusetzen. Somit liegen antagonistische Wünsche vor. Der Wunsch ‘zu schießen’ impliziert nicht, dass der Wunsch ‘nicht zu schießen’ nicht vorhanden ist. Ich kann sogar auf zweiter Stufe wünschen, diese Mittel nicht einsetzen zu wollen. Die Absicht ‘zu schießen’ impliziert, dass der Wunsch ‘zu schießen’ vorliegen sollte. Jedoch ist es nicht notwendig, dass der Wunsch 2. Stufe ‘die Absicht zu haben zu schießen’ oder der Wunsch 2. Stufe ‘den Wunsch zu haben zu schießen’ vorhanden sind. Da dies in diesem Beispiel auch rational für den Akteur ist, lehne ich Axiomenschemata (4_d) bzw. (*IDI*) ab. Es ist für Wünsche nicht konstitutiv, dass man sich jeden Wunsch auch wünscht. Es gibt rational keinen Grund, sich jeden Wunsch wünschen zu müssen. Desgleichen gilt für Absichten. Man muss sich eine Absicht nicht wünschen. Somit stellen (4_d) bzw. (*IDI*) keine geeigneten gültigen Schemata auf, um einen Akteur als rational für eine bestimmte Absicht bzw. einen Wunsch zu beschreiben.

$(I_d) \quad \alpha des : \alpha des : \varphi \supset \alpha des : \varphi$ $(DII) \quad \alpha des : \alpha dint : \varphi \supset \alpha dint : \varphi$

Es macht für die Verantwortung, die man dem Akteur für den Schuss zuschreibt, möglicherweise einen Unterschied, ob der Akteur sich auch wünschte, dass er wünschte zu schießen, oder ob er es nur aus dem Zwang heraus tat [60]. Aber wie Searle [123, S.136] festhält, sind ‘für eine Tat verantwortlich sein’ und ‘eine Hand-

⁷³Bei Finlay [59] beinhaltet ein Wunsch, dass p , den Sachverhalt p als bestehend zu wünschen. In gewisser Weise stimmt Finlay Velleman's Einordnung von Wunsch als konative Einstellung zu. Diese Konativität sorgt nach Finlay dafür, dass, wenn ein Akteur ein Ziel bzw. einen Wunsch hat, er den Wunsch hat, den gewünschten Sachverhalt wahr zu machen. Bzw. beinhaltet sie auch, die Umstände zu wünschen, die dieses p wahr machen. „Desiring an end does not necessarily involve motivation towards doing the end, but rather towards facilitating the end“ [59, S.180].

4. Überzeugungen, Wünsche und Absichten

lung absichtlich vollziehen' zwei unterschiedliche und unabhängige Konzepte.⁷⁴

Finlay [59] stellt die Frage, ob ein Wunsch 2. Stufe 'sich zu wünschen, dass man wünscht, dass p ' nichts anderes als der Wunsch 1. Stufe 'sich zu wünschen, dass p ' ist bzw. dass der Wunsch 2. Stufe den Wunsch 1. Stufe voraussetzt, vgl. Schemata (4_d) und (I_d). Aus der Ablehnung von (4_d) ist offensichtlich, dass das erst genannte eine ungeeignete Forderung wäre. Dass der Wunsch 2. Stufe immer den Wunsch 1. Stufe voraussetzt bzw. dass es nicht rational ist, sich etwas in 2. Stufe zu wünschen, was man sich auf 1. Stufe nicht wünscht, lässt sich durch ein Beispiel Frankfurts widerlegen [61, S.9]. Ein Psychotherapeut behandelt Suchtkranke. Er ist ein Gegner von Drogen und hat noch nie den Wunsch verspürt, Drogen zu nehmen. Seine Motivation ein besonders guter Arzt zu sein, lässt ihn sich jedoch wünschen, dass er sich in seine Patienten besser hineinversetzen kann. Seine Mittel-Zweck Überzeugung, dass er sich besser hineinversetzen könnte, wenn er den Wunsch, die Droge zu nehmen, einmal verspüren würde, führt dazu, dass er urteilt, dass es gut wäre, den Wunsch, die Droge zu nehmen, einmal zu verspüren. Die erstgenannte Motivation, ein guter Arzt zu sein, führt dazu, dass er tatsächlich den Wunsch bildet, den Wunsch zu haben, Drogen zu nehmen. Dennoch hat er zu keinem Zeitpunkt den Wunsch, Drogen zu nehmen. Somit ist er rational für den Wunsch, sich zu wünschen, Drogen zu nehmen, aber gleichzeitig nicht wünscht, Drogen zu nehmen. Ein Schema wie (I_d) ist somit ungeeignet.

Für Absichten ist dies offensichtlich. Es kann rational sein, sich eine Absicht zu wünschen, aber die Absicht nicht zu haben. Wie ein Akteur sich wünschen kann, eine Handlung zu vollführen, ohne die Handlung auszuführen, kann ein Akteur sich wünschen, eine Absicht zu haben, ohne die Absicht zu haben. Es ist sicherlich unstrittig, dass (DII) abzulehnen, weniger der Begründung bedarf, als (I_d) abzulehnen.

Am Ende dieses Abschnittes möchte ich auf den Anfang zurückkommen. Ich hatte begonnen, Unterschiede und Gemeinsamkeiten zwischen Absichten und Wünschen aufzustellen. Während es für Wünsche rational nicht relevant ist, ob sie konfliktieren, ist es für die Rationalität von Absichten vonnöten, dass sie nicht konfliktieren. Gleichzeitig zu beabsichtigen, dass p , und zu beabsichtigen, dass nicht p , würde bedeuten, eine Handlung zu planen, die sicherstellt, dass sowohl p als auch nicht p vorliegt. In Abschnitt 4.1 ist festgehalten worden, dass eine Intention eine kognitive Einstellung ist, bei der der Akteur akzeptiert, dass der Sachverhalt von ihm hervorgebracht wird.

⁷⁴Jemand, der Fahrerflucht begeht, handelt möglicherweise nicht absichtlich, ist aber verantwortlich. Jemand, der gezwungen wird, einen Vertrag zu unterschreiben, unterschreibt ihn absichtlich. Er ist jedoch möglicherweise nicht verantwortlich [123, S.136]. Es gibt zwei Arten von „sollte“. Zum einen sollte ein Akteur rational handeln, zum anderen sollte er moralisch richtig handeln.

$(D_d) \quad \alpha \text{ des} : \varphi \supset \neg \alpha \text{ des} : \neg \varphi$ $(D_i) \quad \alpha \text{ dint} : \varphi \supset \neg \alpha \text{ dint} : \neg \varphi$ $\alpha \text{ int} : \varphi \supset \neg \alpha \text{ int} : \neg \varphi$

Im Falle konfligierender Intentionen müsste der Akteur akzeptieren, dass p und nicht p gleichzeitig hervorzubringen sind.

Für Überzeugungen hatte ich eine solche Rationalitätsforderung abgelehnt. Es ist nicht rational, die Konjunktion der Gehalte der Überzeugungen zu glauben, nur weil zwei Überzeugungen vorliegen, siehe S.145 bzw. meine Argumente gegen (K_b) , (A_b) . Für Akzeptanz stellt Cohen fest, ist dies anders [46, S.29]. Für Akzeptanz sollte gelten, dass, wenn jemand p als wahr akzeptiert und $p \supset q$ als wahr akzeptiert, dann sollte er auch q als wahr akzeptieren. Ich stimme Cohen zu. Da die Rationalität einer Akzeptanz von p als wahr nicht an dem Wahrheitswert von p in der tatsächlichen Situation festgemacht wird, wie im Falle einer Überzeugung, ist es rational für den Akteur zu akzeptieren, dass q , wenn bereits p und $p \supset q$ als wahr akzeptiert werden [46, §5]. Demzufolge ist es aufgrund der kognitiven Komponente einer Intention nicht möglich, dass konfligierende Intentionen rational für den Akteur sind.

Dies schließt nicht aus, dass ich ebenfalls den Wunsch im obigen Beispiel habe, nicht zu schießen, weil ich ein friedliebender Mensch bin. Dass antagonistische Wünsche möglich sind, sogar rational möglich sind, hatte ich bereits zu Beginn dieses Abschnittes befürwortet. Audi [8, S.18], Bratman [25], Wallace [157] stimmen alle überein, dass es möglich ist, antagonistische Wünsche zu haben, und dass es *rational* möglich ist, diese zu haben, nur so hat der Akteur die Möglichkeit, zwischen Handlungsalternativen zu wählen. Dass ein Akteur sich für eine Handlung entscheidet und möglicherweise ein Wunsch eine Handlung stärker als ein anderer Wunsch motivieren kann, liegt darin, dass Wünsche im Gegensatz zu Intentionen graduell sind. Auch wenn ein Wunsch graduell ist, eine Absicht ist es nicht. Wenn ein Akteur eine Absicht hat, dann stellt er sich vor, den Sachverhalt hervorzubringen. Er ist nicht mehr oder weniger davon überzeugt, bzw. wünscht es mehr oder weniger. Er nimmt es als wahr an, dass er diese beabsichtigte Handlung vollziehen wird. D.h. auch wenn ich (D_d) ablehne, ist (D_i) von einem rationalen Akteur zu erwarten. Zum anderen lässt die Nichtgraduiertheit der Intentionen ein Wunsch-Überzeugung-Modell für Intentionen unsinnig erscheinen.

Dies hat Auswirkungen auf die Semantik, in der ein Absichts- bzw. Wunschoperator interpretiert werden sollte. Während im Falle des Wunschoperators nur eine

4. Überzeugungen, Wünsche und Absichten

Nachbarschaftssemantik in Frage kommt, aufgrund der Möglichkeit antagonistische Wünsche zu haben, bietet sich für den Intentionoperator eine relationale Semantik wie in *BDICTL** vorgeschlagen an. Dies führt zur Diskussion weiterer Axiome.

(K_i)	$\alpha \text{ int} : (\varphi \supset \psi) \supset (\alpha \text{ int} : \varphi \supset \alpha \text{ int} : \psi)$
(A_i)	$(\alpha \text{ int} : \varphi \wedge \alpha \text{ int} : \psi) \supset \alpha \text{ int} : (\varphi \wedge \psi)$
(K_{dint})	$\alpha \text{ dint} : (\varphi \supset \psi) \supset (\alpha \text{ dint} : \varphi \supset \alpha \text{ dint} : \psi)$
(A_{dint})	$(\alpha \text{ dint} : \varphi \wedge \alpha \text{ dint} : \psi) \supset \alpha \text{ dint} : (\varphi \wedge \psi)$
(K_d)	$\alpha \text{ des} : (\varphi \supset \psi) \supset (\alpha \text{ des} : \varphi \supset \alpha \text{ des} : \psi)$
(A_d)	$(\alpha \text{ des} : \varphi \wedge \alpha \text{ des} : \psi) \supset \alpha \text{ des} : (\varphi \wedge \psi)$

Warum (K_d) und (A_d) zurückzuweisen sind, liegt auf der Hand. Angenommen, ich wünsche, dass es regnet, weil ich möchte, dass meine Gartenblumen gegossen werden, und ich wünsche, dass es nicht regnet, weil ich das schöne Wetter genießen möchte, besteht kein Grund zu wünschen, dass es regnet und nicht regnet. Diesen Sachverhalt als bestehend zu wünschen, ist unsinnig. Die andere Möglichkeit in einem solchen Fall (A_d) zu genügen, wäre den einen oder den anderen Wunsch aufzugeben. Dies aufgrund von Rationalität zu fordern, ist ebenfalls unsinnig. Ich bin rational, beide Wünsche zu haben. Ob es regnet oder nicht, gehört nicht zu den Sachverhalten, die ein Akteur hervorbringt. Somit kann es sinnvoll sein, für beide Alternativen Motivationen zu Handlungen zu haben, die je nachdem, ob es tatsächlich regnet oder nicht, handlungswirksam werden. So ist es im Falle der Gartenblumen rational, weiterhin zu wünschen, dass es regnet, auch wenn er sich in der Sonne liegend wünscht, dass sie scheinen möge. In einer allgemeinen Nachbarschaftssemantik sind (K_d) und (A_d) nicht allgemeingültig.

Dies ist für Absichten nicht sinnvoll. Die Funktion einer Absicht besteht darin, Ziele bzw. Wünsche des Akteurs umzusetzen. Bratman formuliert, dass zukunftsgerichtete Intentionen Pläne darstellen, die ein Akteur entwickelt und die mit Handlungen zu vervollständigen sind, die ein Akteur für die Umsetzung der Pläne durchführen sollte. Beispielsweise fasse ich die Intention, in der Buchhandlung vorbeizufahren, wenn es mein Ziel ist, mir ein Buch aus der Buchhandlung zu besorgen [25]. Wenn es nicht nur beim Wunsch vorbeizufahren bleiben soll, sondern der Akteur dies tatsächlich sicherstellen will, muss er einen Plan fassen, der ihn dahin leitet. Solche Pläne basieren nach Bratman auf den Intentionen bzw. sind selbst Intentionen. Umgekehrt ist eine Absicht Teil eines Planes bzw. selbst ein „partial plan“. Aus Rationalitätsgründen ist der Akteur nach Bratman verpflichtet, für jeden Plan auch die notwendigen Mittel zu intendieren. Z.B. sollte ich überlegen, ob ich das Auto zur Buchhandlung nehme,

und weiter sicher stelle, dass das Auto nicht meinem Mann versprochen ist, bzw. ob ich den Bus nehme und mir dafür den Fahrplan anschau [25, S.29ff]. Dabei betont Bratman, dass dies nur dazu dient, die Handlungen des Akteurs zielgerichtet zu leiten. Dafür stellt er Rationalitätsforderungen an Pläne und Intentionen, zum einen Bewahrung der Konsistenz der Gehalte von Intentionen zum anderen Kohärenz zwischen den Zielen und den notwendigen Mitteln, die jemand beabsichtigt [25, S.31]. Auf letzteres möchte ich später eingehen.

Beide Forderungen begründen sich darauf, dass die intendierten Handlungen zum einen zulässig und zum anderen relevant sein sollen. Ein Akteur intendiert nur, was ihm auch möglich ist umzusetzen. Und er berücksichtigt beim Fassen von Intentionen andere Intentionen und Ziele, die bereits vorliegen [25, S.33f]. Intentionen stellen eine Art „Filter“ dar, welche Optionen, die durch Wünsche und Überzeugungen aufgezeigt werden, relevant und erlaubt sind [25, S.34]. Dies ist einer der Gründe, warum für Bratman Intentionen nicht auf Wünsche und Überzeugungen zurückführbar sind. Diese Rationalitätsforderungen können Wünsche und Überzeugungen nicht sinnvoll erfüllen. Es ist nicht ihre Aufgabe.

„Their role [the role of intentions, Anm. d. Verf.] is to help determine which options are to be considered in the process of weighing conflicting reasons for action, rather than to provide reasons to be weighed in favor of one considered alternative over another. The reasons to be weighed in deliberation remain desire-belief reasons. In this way we go beyond the modest extension of the desire-belief model, giving intentions a direct role as inputs in practical reasoning.“ [25, S.34]

Dies schließt sich daran an, was Searle bei der Diskussion gemeint hat, welche Einstellungen eine Antwort auf die Frage geben, warum ein Akteur etwas tut. Die Antwort auf die Warum-Frage bei Handlungen geben die Wünsche und Überzeugungen. Die Intentionen beantworten hinsichtlich Handlungen eine andere Frage: Wie wird der Akteur seine Ziele erreichen? Die Antwort auf die Wie-Frage ist die Intention zur Handlung. Warum fahre ich in die Buchhandlung, weil ich glaube, dass ich ein Buch brauche, weil ich wünsche, meiner Freundin ein Buch zuschenken etc. Wie komme ich in die Buchhandlung? Wie erreiche ich mein Ziel? Ich beabsichtige, die Uni zu verlassen, mit dem Auto zu fahren, das Auto in der Nähe der Buchhandlung abzustellen und mir ein Buch im Laden auszusuchen.

In gewisser Weise hat Baier dies bereits festgehalten [11]. Sie argumentiert gegen Chisholm [42], dass eine Intention, etwas zu tun, nicht mit der Intention gleichzusetzen ist, dass man etwas hervorbringen wird bzw. dafür sorgen wird, dass man es

4. Überzeugungen, Wünsche und Absichten

tut. Ihre Argumentation schlägt meines Erachtens fehl. Sie beruft sich darauf, dass eine Aussage der Form ‘ich intenziere, den Jungen zur Raison zu rufen’ eine andere Aussage ist als ‘ich beabsichtige, dafür zu sorgen, dass ich den Jungen zur Raison rufe’. Damit hat sich sicherlich recht. Eine Aussage, ob ich etwas beabsichtige, oder die Aussage, dass ich beabsichtige, dafür zu sorgen, unterscheiden sich in Hinsicht darauf, dass ich mit erst genannter implizit sage, dass ich weiß, wie ich gedenke, es zu tun, und in letzterer die Möglichkeit besteht, dass mir die Methode nicht bewusst ist, mit der ich dafür sorgen werde.

Die Idee Chisholms war nicht, einen Unterschied zu verneinen, zwischen dem Aussagen von Absichten und dem Aussagen von Absichten für bestimmte Sachverhalte selbst zu sorgen. Er wollte festhalten, dass eine Absicht immer auch beinhaltet, für den beabsichtigten Sachverhalt zu sorgen. Dennoch liefert Baiers Argument, warum die beiden Aussagen unterschiedlichen Gehalt haben, ein wichtiges Kriterium für Intentionen. Ihr Problem ist, dass ein Ausdruck der Form ‘der Akteur beabsichtigt hervorzubringen, dass φ ’ lediglich eine kausale Zuschreibung des Sachverhaltes φ ist. Dies wäre zu wenig für eine Handlung. Wie kommt die ‘Wie-Frage’ ins Spiel? Baier meint, dass die Frage nach dem ‘wie wird er die Handlung vollziehen’ im ersten Fall beantwortet, aber im zweiten Fall nur aufgeworfen wird. Wenn ich beabsichtige, den Jungen zur Raison zu bringen, unterstellt man mir bereits, dass ich weiß, was ich tun werde. Im Falle der Aussage ‘ich beabsichtige dafür zu sorgen, dass ich den Jungen zur Raison bringe’ steht nach Baier die Frage im Raum, wie ich dafür sorgen werde. Welche Methode ich anwende, um den Jungen tatsächlich zur Raison zu bringen. Dies ist nach Baier eine andere Frage, weil es auf die Frage ein ‘irgendwie’ als mögliche Antwort geben kann. Dieses ‘irgendwie’ erklärt nicht, welche Handlungen bzw. Aktivitäten der Akteur explizit beabsichtigt. Nach Baier ist dieses Erkennen der Dinge, wie ich eine Handlung erfolgreich durchführen kann, essentiell für den Begriff ‘Intention’. Insbesondere trifft dies zu, wenn ‘absichtlich handeln’ also gegenwärtige Intentionen bei einer Handlung betrachtet werden.

Ungewiss ist, ob Chisholm Baier recht geben würde, dass im Beabsichtigen, einen bestimmten Sachverhalt *hervorzubringen*, der Akteur glaubt, dass er den Sachverhalt *irgendwie* hervorbringt, aber selbst nicht genau wissen muss, wie er ihn hervorbringt, jedoch im Beabsichtigen, dass der Sachverhalt *besteht*, bereits der Akteur überzeugt ist, *wie* er den Sachverhalt zum Bestehen bringt. Bei Chisholm muss bei jeder Intention die Überzeugung vorliegen, den Sachverhalt auch hervorbringen zu wollen. Baier hat recht, dass Chisholm nicht hinzufügt, dass eine Intention beinhaltet, dass der Akteur vor Augen hat, wie er es umsetzt. Baiers Beispiel ist überzeugend: ‘Ich

beabsichtige zu gehen' ist nicht notwendigerweise äquivalent mit 'ich beabsichtige dafür zu sorgen, dass ich gehe'. Aber mit letzterer Beschreibung ist nicht gemeint, dass tatsächlich ein anderer Sachverhalt bzw. eine andere Handlung als die Aktivität des Gehens beabsichtigt ist, sondern vielmehr, dass zum Ausdruck gebracht wird, dass es sich beim Gehen um eine Handlung des Akteurs handelt.

An diesen Argumenten von Bratman und in gewisser Weise auch von Baier⁷⁵ wird deutlich, dass eine Absicht nicht nur die Vorstellung beinhaltet, eine Handlung zu vollziehen, sondern auch die Vorstellung, wie die Handlung vollzogen wird. Wenn ein Akteur sich nicht nur verpflichtet, dass er mit einer Handlung einen Sachverhalt hervorbringt, sondern sich auch festlegt, wie er diesen Sachverhalt hervorbringt, sollten die Gehalte der Absichten eines Akteurs in einem Moment konsistent sein. Ein Akteur kann sich nicht rational vorstellen, wie er gleichzeitig mit einer Handlung eine Inkonsistenz hervorbringt. Hierbei spielt die Überzeugung, ob es möglich ist oder nicht, keine Rolle. Allein aufgrund der Funktion einer Intention, die nicht auf Überzeugungen zurückführbar ist, ist somit gezeigt worden, dass die Rationalitätsforderung nach der Konsistenz der Gehalte sinnvoll ist.

Absichten agglomerieren, jedoch nicht monoton

An den bisherigen Argumenten wird Folgendes deutlich. Während ein Akteur sich auf die Fragen, warum habe ich für A gesorgt und warum habe ich im selben Moment für B gesorgt, konfligierende Antworten geben kann, ist dies bei der Beantwortung der Frage, wie habe ich das erreicht, nicht möglich. Wie er es getan hat und wie er es wieder tun wird, sollten konfliktfrei unabhängig davon sein, was er gern möchte. Dies ist nicht nur ein weiterer Grund (D_i) anzunehmen, sondern vor allem ein Grund, dass (K_i) und (A_i) sinnvolle Rationalitätsforderungen sind. Wenn ein Akteur sich vorstellt, wie er bewirkt, dass φ , und wie er bewirkt, dass ψ , sollte er sich vorstellen, wie er $\varphi \wedge \psi$ zusammen bewirkt. Wenn seine Absicht, dass φ , mit der Absicht, dass ψ , nicht im Bereich dessen liegen, was er gleichzeitig hervorbringen kann, sollte er eine der beiden Absichten aufgeben. Diese Forderung stellen (K_i) und (A_i) sicher.⁷⁶

Bratman begründet, dass (A_i) eine rationale Forderung ist, in Analogie zu Vellemans Begründung, dass Überzeugungen agglomerieren [27, S.50f]. Velleman argumentierte dafür, dass Überzeugungen einer Korrektheitsnorm genügen und auf die Welt abzie-

⁷⁵Darauf werde ich im nächsten Abschnitt zurückkommen.

⁷⁶Der Skopus bei der Interpretation von (A_i) bzw. (K_i) der Rationalitätsforderung ist immer weit zu verstehen. Es sollte der Fall sein, wenn der Akteur zwei Absichten hat, dass er die Konjunktion der Gehalte der Absichten beabsichtigt. Wenn dies nicht möglich ist, sollte er eine der beiden Absichten aufgeben.

4. Überzeugungen, Wünsche und Absichten

len. Da die Welt eine Einheit bildet und somit konsistent und kohärent ist, ist es für Überzeugungen rational, konsistent und kohärent zu sein. Bratmans Argument, dass Intentionen agglomerieren sollten. Seine Forderung, dass, wenn ein Akteur beabsichtigt, dass φ , und er beabsichtigt, dass ψ , er ebenfalls beabsichtigen sollte, dass φ und ψ , basiert auf derselben Argumentation. Die Intentionen eines Akteurs sollten (A_i) genügen, weil er nicht gleichzeitig in der realen Welt hervorbringen kann, dass φ , und hervorbringen kann, dass ψ , ohne auch $\varphi \wedge \psi$ hervorzubringen. Nach Bratman ist (A_i) begründet und zwar unabhängig davon, ob ein Akteur eine Überzeugung oder irgendeine andere kognitive Einstellung im Falle einer Intention hat oder nicht [27, S.52]. Der Akteur sollte, wenn er intendiert, dass φ , und intendiert, dass ψ , auch intendieren, dass $\varphi \wedge \psi$, nicht nur weil eine Intention darauf abzielt, ihren Gehalt in einer Welt wahr zu machen, von der der Akteur glaubt, dass sie so oder so ist. Vielmehr ist das Ziel einer Intention, ihren Gehalt nicht nur in der vom Akteur geglaubten Welt wahr zu machen, sondern den Gehalt in der einzigen aktuellen Welt als Teil einer koordinierten Realisierung aller seiner Intentionen wahr zu machen. Dabei spielt es keine Rolle, ob der Akteur glaubt oder nicht glaubt, dass er es tatsächlich tun wird, tun kann oder es überhaupt möglich ist [27, S.52].

Bratman meint, dass es sinnvoll ist, dass die Intentionen eines Akteurs agglomerieren, ohne dass involvierte Überzeugungen dieses Agglomerieren erzwingen. Denn nur wenn Intentionen agglomerieren, ist es gewährleistet, dass ihre Gehalte konsistent sind und zu einem konsistenten Plan beitragen. Die Konsistenz ist notwendig, um in einer einzigen Welt realisierbar zu sein.

„If we think of the aim of each intention in a correspondingly atomistic way – each intention aims at making its specific content true in the world as one believes it to be – it is not clear why one intention should agglomerate with a second intention in the absence of belief in success of both intentions. [...] each intention aims at making its content true as a part of a coordinated realization of one’s planning system, in the world as one believes it to be.“ [27, S.52]

Bratman hält weiter fest, dass nach Vellemans Ansicht die „Einigkeit“ der Welt erklärt, warum Überzeugungen unter dem Druck stehen, zu agglomerieren. Für Bratman ist es nicht die Welt, die diesen Druck aufbaut, dass Intentionen agglomerieren. Intentionen agglomerieren für ihn nicht deswegen, weil sie eine Form von kognitiver Verpflichtung sind, dass die Welt irgendwie geartet ist, sondern weil der Akteur innerhalb seiner Handlungen nur koordiniert und effizient agieren kann, wenn seine Intentionen agglom-

merieren [27, S.52f].⁷⁷

Ich gebe Bratman recht, wenn er meint, dass die Agglomeration von Intentionen nicht auf die Agglomeration von Überzeugungen zurückführbar ist. Er übersieht meines Erachtens, dass ein Grund ist, warum Handlungen nicht koordiniert und effizient ablaufen können, wenn die Intentionen nicht agglomerieren oder inkonsistent sind, dass ein Moment in einem Geschichtsverlauf so beschaffen ist, dass inkonsistente Gehalte nicht hervorgebracht werden können. So stimme ich zu, dass Intentionen unter dem rationalen Druck stehen, konsistent zu sein und zu agglomerieren, was sich in den Schemata (D_i) und (A_i) widerspiegelt. Aber ich würde zurückweisen, dass der Grund dafür *nicht* der ist, dass die Welt konsistent und eine Einheit ist. Denn dies ist genau der Grund.

Velleman fragt: „my intentions must be jointly executable, if I can be agnostic as to whether they will be executed?“ [150, S.206] und „why [...] should an agent be rationally obliged to arrange means of carrying out an intention, if he is agnostic about whether he will in fact carry it out?“ [150, S.205] und „why should my intentions be subject to a requirement of consistency if I can remain cognitively uncommitted to their truth?“ [150, S.206]. Er verlangt somit eine kognitive Einstellung von dem Akteur, wenn er der Meinung ist, dass er etwas intendiert. Diese ist jedoch nicht notwendigerweise eine Überzeugung. Ich werde in Abschnitt 4.3 auf die Frage zurückkommen, inwieweit eine Überzeugung, dass man etwas tun wird oder tun kann, einer Intention zugrunde liegen muss.

In der Summe sind sich Velleman und Bratman einig, auch wenn aufgrund scheinbar unterschiedlicher Ansichten, dass Intentionen agglomerieren müssen. Für „reine“ Wünsche verneint Bratman solche Rationalitätsforderungen. Da diese keine koordinierenden Aufgaben übernehmen, sind Rationalitätsforderungen an Wünsche unangebracht. Er lehnt somit (K_d) und (A_d) ab. Die Koordinierung und Realisierung von Handlungen sind Aufgabe der Intentionen und unterscheiden diese von gewöhnlichen Wünschen [27, S.52].

In Abschnitt 6.1 werde ich daher einen *int*-Operator einführen, der agglomeriert und konsistent aufgrund der Semantik ist, in der er ausgewertet wird. Dennoch ist dieser Operator nicht hinreichend, um alle Rationalitätsforderungen an einen Akteur zu

⁷⁷Bratman versucht das Agglomerationsprinzip für Intention auf die Agglomeration von Überzeugungen zurückzuführen. Dies ist inkonsequent. Da Intentionen nicht auf Überzeugungen zurückzuführen sind, sollten sich die Rationalitätskriterien für Intentionen nicht aus den Rationalitätsforderungen der Überzeugungen ableiten.

4. Überzeugungen, Wünsche und Absichten

(M_i)	$\alpha \text{ int}:(\varphi \wedge \psi) \supset (\alpha \text{ int}:\varphi \wedge \alpha \text{ int}:\psi)$
(RM_i)	$\vdash \varphi \supset \psi \Rightarrow \vdash \alpha \text{ int}:\varphi \supset \alpha \text{ int}:\psi$
(M_{dint})	$\alpha \text{ dint}:(\varphi \wedge \psi) \not\supset (\alpha \text{ dint}:\varphi \wedge \alpha \text{ dint}:\psi)$
(RM_{dint})	$\vdash \varphi \supset \psi \not\Rightarrow \vdash \alpha \text{ dint}:\varphi \supset \alpha \text{ dint}:\psi$
(M_d)	$\alpha \text{ des}:(\varphi \wedge \psi) \supset (\alpha \text{ des}:\varphi \wedge \alpha \text{ des}:\psi)$
(RM_d)	$\vdash \varphi \supset \psi \Rightarrow \vdash \alpha \text{ des}:\varphi \supset \alpha \text{ des}:\psi$

erfüllen. Eine rationale Forderung⁷⁸ ist, dass ein Akteur nicht jeden notwendigerweise bestehenden Sachverhalt intendieren sollte. Da ein Gehalt einer Absicht ein Sachverhalt sein sollte, den der Akteur mit einer seiner Handlungen selbst hervorbringt, werde ich in Abschnitt 6.4 einen spezifizierten Intentionsoperator *dint* einführen. Dieser schließt aus, dass ein Akteur beabsichtigen sollte, was notwendigerweise besteht, vgl. Definition 23 S.296. Diese Definition enthält als Positiv-Bedingung die Forderung, dass die Bedingung des *int*-Operator erfüllt ist. Der *dint*-Operator erfüllt somit die Axiome (K_{dint}) und (A_{dint}) . Die Gehalte der Intentionen, die mit dem *dint*-Operator beschrieben werden, bleiben somit aufgrund des Agglomerationsprinzips konsistent. Einen Unterschied zum *int*-Operator macht die Negativ-Bedingung des *dint*-Operators. Während der *int*-Operator monoton ist, ist es der *dint*-Operator nicht. Ich werde kurz argumentieren, warum Monotonie im Falle eines Intentionsoperators nicht sinnvoll ist.

Vielleicht möchte man annehmen, dass, wenn eine Intention mittels einer relationalen Semantik beschrieben wird, und sowohl das Schema (D_i) als auch (K_i) angenommen wird, sich jemand anschließen würde, dass Monotonie für diesen Operator ebenfalls angenommen werden sollte. Der Meinung bin ich nicht. Es ist nicht notwendig, wenn ich intendiere, dass ich $\varphi \wedge \psi$ hervorbringen werde, auch jedes einzeln beabsichtigen würde. Es kann sein, dass es Sachverhalte gibt, die nur sinnvoll hervorbringen sind, wenn sie zusammen hervorgebracht werden. So macht es Sinn zu intendieren, einen linken Schuh und einen rechten Schuh zu kaufen. Aber allein daraus zu schlussfolgern, dass ich intendiere, nur einen linken Schuh zu kaufen, ist nicht rational. Während es sinnvoll sein kann, zwei Sachverhalte zusammen hervorbringen, ist es nicht notwendigerweise rational, die Sachverhalte auch einzeln hervorbringen. Aus diesem Grund ist eine Forderung wie (M_i) für Intentionen und die korrespondierende Regel der Monotonie für Intentionen (RM_i) nicht sinnvoll.

In gewisser Weise wird dies auch in der Bewertung von Intentionen berücksichtigt.

⁷⁸Siehe Abschnitt 4.3.

Die „Lehre vom Doppeleffekt“ erlaubt, Handlungen zu vollführen, die zu unerwünschten Nebenwirkungen führen, solange die Handlung einem höheren Ziel dient, auch dann, wenn die Nebenwirkungen allein nicht absichtlich hervorzubringen erlaubt sind. Es wird moralisch bzw. möglicherweise in Hinsicht auf die Verantwortung verschieden beurteilt, ob der Akteur einen Sachverhalt allein beabsichtigt hat oder ob er *nur* vorhergesehen hat, dass er eintreten wird, wenn er die Handlung vollzieht.

Es wird eine Handlung intendiert, die, wie der Akteur vorhersieht, zu einem gewünschten Sachverhalt ψ und einem bestimmten Sachverhalt φ führt, der bestimmte negative Wirkungen hat, ohne dass der Akteur den Sachverhalt φ beabsichtigt. In Bratmans Beispiel [25, 27] ist es für einen Bomberpiloten notwendig, eine Munitionsfabrik zu bombardieren. Er sieht, dass seine Bombardierung gleichzeitig die Schule neben der Munitionsfabrik trifft, und damit den Tod von Kindern bewirken würde. Angenommen er entschließt sich, die Bombardierung durchzuführen. Was ist der Gehalt seiner Intentionen? Die Munitionsfabrik zu bombardieren, ist Teil seiner Absicht. Da er weiß, dass er die Munitionsfabrik nicht bombardieren kann, ohne die Schule zu bombardieren, ist für ihn abzusehen, dass die Schule zerstört wird, wenn er die Munitionsfabrik bombardiert. Das „Choice-Intention Principle“ besagt nicht, dass er nicht beabsichtigen kann, nur die Munitionsfabrik zu bombardieren, auch wenn er weiß, dass er auch die Schule zerstören wird, wenn er es tut [25, S.145]. Da es in seiner Wahl steht, die Fabrik zu bombardieren, ist es auch seine Wahl, die Schule zu zerstören. Wenn etwas in seiner Wahl steht und er weiß, dass er es bewirkt, hat er die Absicht. Würde man aber von ihm verlangen, dass, wenn er beabsichtigt, die Munitionsfabrik zu bombardieren und dabei die Schule zu zerstören, dass er beabsichtigt, die Schule allein zu zerstören? Bratman nennt die Idee hinter dem Schema (M_i) das Prinzip der Intention-Division [25, S.145]. Er kommt zum Schluss, dass (M_i) und das „Choice-Intention Principle“ dazu führen, dass der Bomberpilot, die Schule zu zerstören, beabsichtigen sollte, wenn der Akteur rational hinsichtlich seiner Intentionen sein möchte, oder er müsste die Intention aufgeben, die Munitionsfabrik zu zerstören.⁷⁹ Dies ist nach Bratman unplausibel.

Bratman argumentiert, das Prinzip der Intention-Division aufrechtzuerhalten, und kritisiert das Wahl-Intention-Prinzip. Der Akteur kann eine Kette von Sachverhalten wählen und irgendein Glied der Kette intendieren, ohne die gesamte Kette zu intendieren. Das führt für Bratman nicht dazu, dass der Akteur jedes Glied der Kette intendieren sollte. Dies ist meines Erachtens der falsche Ansatz für Rationalität. Wenn ich

⁷⁹Ich möchte den „Doppeleffekt“ nicht verteidigen oder ablehnen. Es soll nur aufzeigen, dass wir Einstellung zu $\varphi \wedge \psi$ haben können, ohne dieselbe Einstellung zu φ oder ψ zu haben.

4. Überzeugungen, Wünsche und Absichten

einen Sachverhalt intendiere und ich weiß, wenn ich diesen Sachverhalt hervorbringe, dass logisch oder kausal begründete weitere Sachverhalte folgen, dann sollte ich diese ebenfalls intendieren, da ich sie implizit auswähle. Alternativ kann ich die Intention aufgeben, den erstgenannten Sachverhalt hervorzubringen. Es wird möglicherweise vom Akteur gewünscht, dass er nicht beabsichtigen muss, die Schule zu zerstören und die Munitionsfabrik zu bombardieren. Aber es ist nicht rational, wenn er die Munitionsfabrik bombardiert mit dem Wissen, dass er die Schule zerstört. Er hatte immer die Möglichkeit, die Intention aufzugeben, die Fabrik zu bombardieren.

Chisholm dagegen verneint (M_i) [42, S.636]. Er unterscheidet zwischen dem Non-Divisionsprinzip für Intentionen (Negation von (M_i)) und dem Verbreitungsprinzip von Intentionen analog zu dem Choice-Intention-Principle von Bratman. Ein Gegenbeispiel Chisholms gegen (M_i) ist das folgende. Angenommen jemand beabsichtigt, zu einem Zeitpunkt in Paris zu sein, von dem er weiß, dass Charles de Gaulle in Paris ist. Er beabsichtigt also, dass er in Paris ist und dass zu diesem Zeitpunkt, wenn er in Paris ist, auch Charles de Gaulle in Paris ist, $\alpha int:(\varphi \wedge (\varphi \supset \psi))$. Da $(\varphi \wedge (\varphi \supset \psi)) \supset \psi$ müsste nach der Regel der Monotonie (RM_i) (Intention-Division-Prinzip) der Akteur intendieren, dass Charles de Gaulle in Paris ist ($\alpha dint:\psi$). Nehmen wir an, dass es sich bei dem Akteur nicht um Charles de Gaulle selber oder einen möglichen Entführer handelt, dann liegt es kaum im Rahmen seiner Wahlmöglichkeiten und im Rahmen dessen, was er rational beabsichtigen kann. Dass Charles de Gaulle in Paris ist, ist keine mögliche Konsequenz seiner Handlungen.⁸⁰ Demzufolge ist es nicht rational, von einem Akteur zu fordern, dass er jede logische Konsequenz seiner Intentionen intendiert. Möglicherweise ist das Choice-Intention-Principle zurückzuweisen, obwohl Chisholm dies verneint. Was jedoch zurückgewiesen werden sollte, ist das Intention-Division-Prinzip (RM_i) bzw. (M_i) .

Die Alternative, die die Regel (RM_i) bieten würde, wäre die Intention $\alpha dint:(\varphi \wedge (\varphi \supset \psi))$ aufzugeben. Aber sollten wir jede Intention aufgeben, wenn eine der Konsequenzen entweder nicht rational intendiert werden kann, wie im Beispiel Chisholms, oder uns nicht wünschenswert erscheint, wie im Falle des Bomberpiloten? Dies ist

⁸⁰Chisholm spricht sich dafür aus, dass im Falle einer rationalen Intention der Akteur glauben muss, dass es ihm möglich ist, den Sachverhalt hervorzubringen. Wenn er glaubt, dass nicht er es ist, der dafür sorgt, dass Charles de Gaulle nach Paris kommt, kann er auch nicht beabsichtigen, dass Charles de Gaulle nach Paris kommt. Ich werde ähnlich dafür argumentieren, dass eine Intention nur rational ist, wenn der Akteur fähig ist, die beabsichtigte Handlung auch umzusetzen und zwar unabhängig von dem, was er glaubt. In diesem Beispiel weiß der Akteur, um die Unfähigkeit seines Dafür-Sorgens, Charles de Gaulle nach Paris zu bringen, so dass sowohl Chisholms Überzeugung in die Möglichkeit als auch die tatsächliche Möglichkeit als notwendige Bedingung erfüllt ist.

keine rationale Forderung.

Im Falle Chisholms müsste der Akteur die Intention aufgeben, zu einem Zeitpunkt nach Paris zu fahren, an dem Charles de Gaulle in Paris ist, weil er nicht hervorbringen kann, dass Charles de Gaulle in Paris ist, obwohl er weiß, dass er dort ist. Das ist unsinnig.

Im Falle des Bomberpiloten müssen wir uns manchmal entscheiden, Dinge zu tun, die wir allein nicht wollen. Dies gilt sowohl für notwendige Mittel zu bestimmten Zielen als auch zu Folgen, die wir möglicherweise allein nicht beabsichtigen würden. Wenn der Bomberpilot im obigen Beispiel seine Handlung ausführt und um die Folge seiner Handlung weiß, hatte er die Absicht, die Munitionsfabrik zu bombardieren, und die Absicht, die Schule zu zerstören? Es war ein und dieselbe beabsichtigte Handlung, die zu diesen Sachverhalten führte, und er wusste, dass er diese Sachverhalte zusammen hervorbringt. Es ist jedoch nicht rational, von ihm zu verlangen, dass er die Schule ebenfalls zu zerstören beabsichtigte, sondern lediglich in Verbindung mit der Bombardierung der Munitionsfabrik. Deswegen werde ich in Abschnitt 6.4 den deliberativen Intentionsoperator zwar in einer relationalen Semantik interpretieren, so dass (K_i) , (A_i) , (D_i) allgemeingültig sind, jedoch die Montonie durch Hinzufügen einer Negativ-Bedingung aufheben. Die Definition 23 des spezifizierten Intentionsoperators findet sich auf S.296.

Im Falle der Wünsche hatte ich argumentiert, dass ein Agglomerationsprinzip nicht sinnvoll ist. Ebenso ist offensichtlich, dass eine Rationalitätsforderung nicht sinnvoll ist, die darauf beharrt, dass Wünsche frei von Antagonismen sind. Es ist sicherlich ebenso fragwürdig, ob ein Wunschoperator ein monotoner Operator sein sollte oder nicht. Für die meisten scheint (M_d) , ein plausibles Schema zu liefern. Ist es rational, dass wir fordern, dass, wenn ein Akteur wünscht, dass $\varphi \wedge \psi$, er auch wünscht, dass φ , und wünscht, dass ψ , oder den Wunsch, dass $\varphi \wedge \psi$, aufgibt. Obiges Beispiel für den Bomberpiloten legt nahe, darüber kritisch zu sein. Der Akteur hat aufgrund der Motivation für seine Handlung eine konative Einstellung zu dem Bombardieren der Fabrik und dem Zerstören der Schule. Aber würden wir von ihm aus Rationalitätsgründen verlangen zu wünschen, dass er die Schule zerstört. Auch wenn ich in diesem Punkt nicht zustimme, so scheint es mir, dass in den meisten Fällen die Rationalität gebietet, dass, wenn ich mir etwas wünsche, auch jede logische Folge dieses Wunsches ebenfalls wünsche oder aber diesen Wunsch aufgeben sollte. Es ist zu einem Wunsch jederzeit rational möglich, sich zu wünschen, diesen Wunsch nicht zu haben. Daher werde ich den Wunschoperator ebenso wie den Überzeugungsoperator in einer monotonen Nachbarschaftssemantik interpretieren. Der Wunschoperator wird durch (RM_d) und

4. Überzeugungen, Wünsche und Absichten

(M_d) axiomatisiert und, obwohl er antagonistische Wünsche zulässt, ist zu vermeiden, dass der Akteur paradoxe Sachverhalte wünscht. So wird (F_d) $\neg\alpha des:\perp$ ebenfalls als Axiom genommen, siehe Abschnitt 6.1 S.244.

In diesem Abschnitt ist somit eindeutig festgelegt worden, welche Schemata zu erfüllen sind und in welcher Semantik ein Wunsch-, Überzeugungs- und Absichtsoperator demzufolge interpretiert werden sollte. Notwendige Interaktionen zwischen Wunsch- und Überzeugungsoperatoren auf der einen und dem Handlungsoperator auf der anderen Seite hatte ich zurückgewiesen. Auch wenn einer Handlung eine Motivation vorausgeht, die auf den Wünschen und Überzeugungen basiert, so liegt für eine Handlung bzw. für ein Dafür-Sorgen, dass φ , nicht notwendigerweise eine Überzeugung oder ein Wunsch, dass φ , oder zu irgendeinem Sachverhalt vor, der logisch notwendig oder hinreichend für φ ist oder in einer anderen logischen Beziehung zu φ steht. Auf die Frage, warum jemand eine Handlung vollführt bzw. warum er eben diesen Sachverhalt φ hervorbringt, wird selten mit der Antwort geantwortet, weil der Akteur wünscht, dass φ , oder weil er glaubt, dass φ . Es ist nicht rational zu fordern, dass bei jedem 'Dafür-Sorgen, dass' der Akteur einen Wunsch oder eine Überzeugung zu genau diesem Gehalt hat. Vielmehr sorgen wir für eine Menge von Sachverhalten, ohne dass wir eine Einstellung zum Bestehen jedes Sachverhaltes haben. Der Gehalt eines Wunsches bzw. einer Überzeugung ist nicht die Handlung, sondern eine Proposition, die uns einen Sachverhalt entweder im Falle einer konativen Einstellung als wahr zu machend oder im Falle der kognitiven Einstellung als wahr begreifen lässt.

Wenn die Frage ist, welcher mentale Zustand nicht nur auf einen Sachverhalt, sondern auf die Handlung bzw. das darunter liegende Ereignis abzielt, ist die Antwort der Zustand der Intention. Eine Absicht gibt die Antwort auf die Frage, wie der Akteur seine Handlungen vollziehen möchte. Dennoch kann eine Absicht nicht alle Sachverhalte erfassen, die während dem Ereignis der Handlung bestehen. Nach Davidson reicht für eine Handlung aus, dass der Akteur einen der Sachverhalte (unter einer Beschreibung) beabsichtigt, die während des Ereignisses bestehen [48, Essay 3]. So verstehe ich Davidsons Termini der Beschreibung. Aber nicht nur Davidson sieht eine enge Beziehung zwischen dem Dafür-Sorgen von Sachverhalten und dem Intendieren der Sachverhalte, für die der Akteur sorgt. Diesen Zusammenhängen ist der folgende Abschnitt gewidmet, bevor ich mich in Kapitel 5 von anderen Logiken abgrenze und in Kapitel 6 das logische System entwickle, welches die in diesem Kapitel motivierten Schemata erfüllt.

4.3 Absichten und Handlungen

Dieser Abschnitt wird sinnvolle Interaktionsaxiome zwischen Handlungs- und Intentionoperator diskutieren und postulieren. Dafür werden Interaktionen zwischen den Handlungen und Intentionen motiviert, die aufgrund deskriptiver oder normativer Aspekte notwendigerweise in einem rationalen Akteur auftreten sollten. Im vorangegangenen Abschnitt ist festgehalten worden, dass ein Intentionoperator in einer relationalen Semantik interpretiert wird, der durch die angegebenen Schemata axiomatisiert werden sollte. Auf die Motivation, den Handlungsoperator durch einen deliberativen Stit-Operator zu beschreiben, war ich bereits in Abschnitt 1.2 eingegangen. In diesem Abschnitt werde ich die aus dieser Semantik resultierenden Axiome und Schemata kurz benennen. Eine weiterführende Motivation dieser Semantik und der zugehörigen Axiomatisierung findet sich in [19]. Der Schwerpunkt dieses Abschnittes liegt nicht auf der Semantik des Handlungsoperators, sondern auf Schemata, die bestehende Interaktionen zwischen Handlungs- und Intentionoperator in einem logischen System beschreiben und so noch nicht diskutiert wurden.

Sind Absichten Absichten zu handeln?

Unbestritten sind die Begriffe ‘Absicht’ und ‘Handlung’ untrennbar verbunden. Der eine Begriff kann ohne den anderen nicht verstanden werden. Dennoch ist es umstritten, welcher konzeptuell dem anderen vorgelagert ist. So erklärt Chisholm in [42, S.634], dass der Begriff der Handlung nur durch den Rückgriff auf das Konzept der Intention erklärt werden sollte und nicht umgekehrt. Chisholm beruft sich darauf, dass eine Intention beinhaltet, einen bestimmten Sachverhalt hervorzubringen und nicht das Konzept einer Handlung voraussetzt. So kann er eine Handlung begrifflich durch das Konzept ‘Intention’ erklären. Anscombe hingegen sieht die begriffliche Vorgelagertheit umgekehrt. Für sie liegt eine Handlung vor. Dieser kann der Modus ‘absichtlich’ oder ‘unabsichtlich’ zugeschrieben werden [7, § 47]. Searle meint, dass die intentionale Komponente einer Intention das Erlebnis des Handelns ist und dass dieses Erlebnis des Handelns im Falle einer Handlungsabsicht (gegenwärtigen Absicht) die Handlung des Akteurs selbst verursacht. Eine vorausgehende Absicht (zukunftsgerichtete Absicht) hat als intentionale Komponente ebenfalls das Erlebnis des Handelns. Wenn es zu der Handlung kommt, hat diese vorausgehende Absicht die Handlungsabsicht und damit die Handlung verursacht [123, S.118f, S.125].

Welches dieser Konzepte dem anderen vorgelagert ist, möchte ich nicht diskutieren. Ich werde in diesem Abschnitt für notwendige Beziehungen zwischen Handlungen

4. Überzeugungen, Wünsche und Absichten

gen und Absichten plädieren und welche Forderungen sinnvoll sind, wenn Akteure als rational für das Haben einer Intention bezeichnet werden. Für letzteres wird es notwendig sein zu unterstreichen, dass nicht das Haben von bestimmten Überzeugungen eine Intention als rational oder nicht rational kennzeichnet, sondern dass das Haben von Intentionen unabhängig von den Überzeugungen in gewisser Weise selbst rational bzw. nicht rational sind. Um solche Beziehungen zu postulieren, werde ich kurz zusammenfassen, was eine Handlung und was eine Intention ist und wie der jeweilige Begriff in einem logischen System beschrieben wird. Beginnen möchte ich mit dem Begriff der 'Intention'.

Eine Absicht richtet sich auf eine Handlung eines Akteurs, auch wenn die Handlung nicht notwendigerweise explizit beschrieben wird, vgl. [123, S.110]. Wenn ein Akteur beabsichtigt, dass er zu Weihnachten in Paris ist, dann ist explizit keine Handlung erwähnt. Diese Absicht impliziert jedoch eine ganze Reihe von Handlungen bzw. weiteren Intentionen, die für diese Intention notwendig sind. Searle unterscheidet eine vorausgehende Absicht von einer Handlungsabsicht. Eine Handlungsabsicht liegt einer Handlung zugrunde. Die Handlungsabsicht kennzeichnet eine Handlung als absichtlich. Eine vorausgehende Absicht kann eine Handlungsabsicht und somit eine Handlung verursachen. In der vorliegenden Absicht, Weihnachten in Paris zu sein, liegt eine von Searle als vorausgehende Absicht vor. Nicht jeder Handlungsabsicht muss nach Searle eine solche Absicht vorausgehen.

Bratman [25] unterscheidet zwischen gegenwärtigen und zukunftsgerichteten Intentionen. Eine zukunftsgerichtete Intention ist nicht nur ein Wunsch, sondern Teil eines Planes, ein bestimmtes Ziel zu erreichen. Um dieses Ziel zu erreichen, muss der Akteur sich zu bestimmten Handlungen verpflichten. Bratman unterscheidet daher zwischen Ziel bzw. Wunsch, die durch eine Handlung erfüllt werden, und einer zukunftsgerichteten Absicht, die die Handlungen des Akteurs koordiniert und es ermöglichen soll, die Handlungen bzw. das Hervorbringen von Sachverhalten auch zu realisieren. Allerdings ist er nicht der Meinung, dass einer absichtlichen Handlung auch immer eine gegenwärtige Intention zu genau dieser Handlung zugrunde liegen muss. Eine Intention, für Bratman ist dies ein mentaler Zustand, hat ein gewisses motivationales Potential, Sachverhalte als absichtlich hervorgebracht auszuzeichnen, ohne dass die Absicht genau diesen Sachverhalt hervorzubringen tatsächlich vorliegt [25, Abschnitt 8.6]. Z.B. kann jemand absichtlich seine Schuhe abtragen, da er sie mit Absicht trägt und sich bewusst ist, dass das Tragen seiner Schuhe zur Abnutzung führt. Dennoch muss er nicht die Absicht haben, seine Schuhe abzunutzen [25, S.123ff].

Ein Beispiel von Bratman, das zeigen soll, dass einer absichtlichen Handlung nicht

die Absicht zu dieser Handlung zugrunde liegen muss, ist jemand, der ein Videogame spielt. Er schießt absichtlich auf zwei Ziele mit der rechten und der linken Hand. Aber er hat nur die Absicht, eines der beiden Ziele zu treffen, da er gewinnt, wenn er eins von beiden trifft, aber verliert, wenn er beide oder keines trifft. Also schießt er absichtlich auf beide, aber mit der Absicht, nur eines zu treffen [24] bzw. [25, Abschnitt 8.2]. Wenn absichtlich auf beide zu schießen, bedeutet, dass er die Absicht hat, auf Ziel 1 und nicht auf Ziel 2 zu schießen, und die Absicht hat, auf zu Ziel 2 und nicht auf Ziel 1 zu schießen, würde er konfligierende Absichten haben und nach Bratman nicht rational sein. Daher ist absichtlich handeln, nicht als Handeln mit der Absicht, so zu handeln, zu identifizieren. Nach Bratman hat er das Ziel zu gewinnen und die Absicht, das Spiel zu gewinnen, indem er weiterhin beabsichtigt, dass eine oder das andere Ziel zu treffen. Das motivationale Potential der letztgenannten Absicht macht beide Handlungen, die der rechten und die der linken Hand, jedoch absichtlich.

Dies entspricht dem sogenannten „Single Phenomenon View“, ohne dem „Simple View“ verpflichtet zu sein. So sagt Bratman, dass einer Absicht zu handeln, etwas zu beabsichtigen und absichtlich handeln jeweils derselbe mentale Zustandstyp zugrunde liegt („Single Phenomenon View“). Aber im Falle einer absichtlichen Handlung muss die Absicht zu eben dieser Handlung *nicht* vorliegen (Negation des „Simple Views“).⁸¹ Ich werde auf diese Unterscheidung nicht weiter eingehen und stimme Bratmans „Single Phenomenon View“ zu, dass jede dieser drei Kategorien auf demselben mentalen Zustandstyp einer Absicht begründet ist. Eine Absicht kann dabei als eine propositionale Einstellung beschrieben werden, die beide Passrichtungen aufweisen kann, wie in Abschnitt 4.1 dargelegt. Auf der einen Seite enthält sie eine kognitive Einstellung, dass der Akteur es als wahr akzeptiert, dass er einen bestimmten Sachverhalt hervorbringt und in gewisser Weise auch wie er ihn hervorbringt. Auf der anderen Seite ist es eine konative Einstellung zu dem Hervorbringen des beabsichtigten Sachverhaltes.

Ich stimme Bratman weiterhin zu. Wenn wir absichtlich handeln, muss keine Absicht vorliegen, die als Gehalt das Hervorbringen eines Sachverhaltes hat, der durch die Handlung hervorgebracht wird. Einer Handlung liegt ein Ereignis zugrunde, das sich durch mehrere Sachverhalte charakterisieren lässt, die dem Akteur alle, als er sorgt dafür, zugeschrieben werden können. Ist das nicht der Fall, ist dieses Ereignis nicht als Handlung des Akteurs zu beschreiben, siehe Abschnitt 3.1 S.68 die These der Stit-Paraphrasierung von Ereignissen.

⁸¹Diese Unterscheidung ‘beabsichtigen’, ‘handeln mit Absicht’ und ‘absichtlich handeln’ findet sich bei Anscombe [7]. Sie sprach jedoch nur im Falle von ‘beabsichtigen’ von einem Zustand. ‘Absichtlich handeln’ fasste sie als eine bestimmte Art und Weise auf, in der Handlungen vollzogen werden.

4. Überzeugungen, Wünsche und Absichten

$$(AiI) \alpha \text{ dstit}:\varphi \supset \alpha \text{ dint}:\varphi$$

Um ein Ereignis als absichtlich zu bezeichnen, müssen nicht alle Sachverhalte, für die der Akteur sorgt und die somit das Ereignis charakterisieren, Gehalte von Absichten sein, die ein Akteur zu dem Zeitpunkt hat. Wenn Davidsons Beschreibungsbegriff mit Sachverhalten interpretiert wird, muss nach Davidson genau ein solcher Sachverhalt, der Gehalt einer Intention sein.⁸² Nach Bratmans Beispiel muss jedoch kein Sachverhalt, der das Ereignis der Handlung charakterisiert und durch den Akteur hervorgebracht wird, Gehalt einer seiner Absichten sein. Es reicht, wenn er die Handlung mit der Absicht vollführt, einen beliebigen, aber bestimmten Sachverhalt hervorzu- bringen, in dessen motivationales Potential die Handlung bzw. das Hervorbringen des Sachverhaltes fällt. Somit muss einer Handlung nicht notwendigerweise eine Absicht einhergehen, irgendeinen Sachverhalt zu beabsichtigen, der das Ereignis der Handlung charakterisiert. Ich stimme dem „Single Phenomenon View“ und Bratman zu, dass dies nicht notwendigerweise zur Akzeptanz des Simple Views führen muss. Eine Handlung benötigt nicht, dass der Akteur einen die Handlung charakterisierenden Sachverhalt beabsichtigt. Die Verschärfung wäre zu fordern, dass der Akteur jeden Sachverhalt beabsichtigt, für den er sorgt. Ein Prinzip wie (AiI) anzunehmen ist daher sowohl nach Bratman als auch nach Davidson nicht sinnvoll.

Für mich steht fest, dass eine Intention beinhaltet, dass der Akteur einen bestimmten Sachverhalt hervorbringen möchte und er auf eine bestimmte Art akzeptiert, dass er dafür sorgt und wie er dafür sorgt. Im Falle der vorausgehenden Absicht kann es sich dabei um einen partiellen Plan handeln. Der Akteur, wenn er die Intention hat, hat auf gewisse Weise einen Plan bzw. eine Vorstellung, wie er für den Sachverhalt sorgt. Dies ist keine Form einer Überzeugung, dass er bestimmte andere Sachverhalt hervorbringen wird, muss oder kann. Wenn ich intendiere, spazieren zu gehen, dann weiß ich, wie ich gehen werde, ohne davon überzeugt sein zu müssen, wie ich das im Einzelnen tun werde. Ich habe akzeptiert, dass ich dafür sorgen werde und *wie*, d.h. welche Fähigkeiten mir dafür zur Verfügung stehen [11]. Die Aufgabe einer Intention ist nicht nur, dem Akteur einen Sachverhalt zu präsentieren, den er hervorbringen möchte (konative Einstellung). Die Aufgabe einer Intention ist es, dass sich der Akteur die Frage, nach dem ‘wie er für den Sachverhalt sorgen will’, beantwortet. So verstehe ich die Idee hinter Bratmans „Planning Theory“ [25].

⁸²Davidson formuliert, dass ein Ereignis genau dann eine Handlung ist, wenn eine Beschreibung, dessen was er tat, den Satz wahr macht, dass er es absichtlich tat [48, S.46]. ‘Absichtlich tun’ ist nach Davidson ‘tun mit der Absicht, es zu tun’. Er ist somit Vertreter des Simple Views.

4.3. Absichten und Handlungen

Dass die intentionale Komponente einer Intention eine Handlung ist, wenn die Handlung das Ereignis ist, ist sicher nicht möglich. So spricht Searle von dem Erlebnis der Handlung [123, S.118]. Ich würde das Erlebnis der Handlung, als das Akzeptieren bzw. das Erkennen identifizieren, dass und wie der Akteur einen bestimmten Sachverhalt durch eine Handlung hervorbringen wird. Baier stellt richtig, dass dabei das Ereignis mehr an Sachverhalten erfasst, als dem Akteur zugänglich ist [12]. Baier bezeichnet dies als „acceptance of limited access“ [12, S.399]. Notgedrungen ist es der Fall, dass ein Akteur, wenn er eine Handlung vollzieht, für mehr Sachverhalte sorgt, als er beabsichtigt. Im äußersten Fall beabsichtigt er keinen der für die Handlung charakteristischen Sachverhalte, aber sorgt dennoch für das Ereignis.

Aber nicht nur, dass er nicht alle Sachverhalte vorhersehen kann, macht deutlich, dass sich Absichten zwar auf Handlungen richten, es aber nicht möglich ist, jeden Sachverhalt der Handlung auch zu beabsichtigen. Wenn wir eine Absicht haben und selbst, wenn wir wissen, wie wir die Handlung vollziehen bzw. dass wir die Handlungen vollziehen, sind Akteure einer gewissen Unvollständigkeit ausgesetzt. Am Beispiel des Gehens wird es deutlich. Ich beabsichtige, spazieren zu gehen. Ich erkenne, was ich dafür tun muss und wie ich spazieren gehe, ohne dass ich beschreiben könnte, was genau ich dabei tue.

Baier schreibt, dass, wenn jemand, der weiß, wie er eine Pirouette ausführt, und beabsichtigt, eine Pirouette zu drehen, er sicherlich viele verschiedene Beschreibungen seiner Handlung liefern kann. Aber gleichzeitig ist sein Wissen, was er tut, wenn er eine Pirouette dreht, in gewisser Weise unvollständig [12, S.396]. Was er weiß, ist, dass er die Fähigkeit hat, diese Handlung durchzuführen. Was im Detail dieses Ereignis ausmacht, ist ihm nur begrenzt zugänglich.⁸³

„This acceptance of *limited access* to one’s own deeds is recognition of a formal feature of every intention, namely, that its object is an act in the real world, about which there will be truths beyond those the agent has access to.“ [12, S.397]

Baier beruft sich dabei auf Anscombe. In Abschnitt § 48 [7] hält Anscombe fest, dass, auch wenn wir meistens nicht das erreichen, was wir gern würden, ein Fehler bei der Ausführung unserer Intentionen die seltene Ausnahme ist. Es ist nicht notwendig, alle Facetten eines Ereignisses zu erkennen, um als menschliches Wesen mit beschränkten

⁸³Die Diskussion, ob Wissen-Wie eine gewisse Form von Wissen-Dass ist, ist sicherlich weitgefächert. Weder verneine ich noch stimme ich zu, dass Wissen-Wie keine Form des Wissens-Dass ist. Ich betone nur, dass ein Wissen-Wie nicht notwendigerweise zugänglich ist.

4. Überzeugungen, Wünsche und Absichten

Kapazitäten eine Intention erfolgreich umzusetzen.

In der Stit-Paraphrase für Ereignisse Abschnitt 3.1 S.68 habe ich formuliert, dass jedes Ereignis eine Handlung des Akteurs ist, wenn jede Beschreibung des Ereignisses durch eine Stit-Paraphrasierung wiederum eine Beschreibung des Ereignisses liefert. Es ist nicht notwendig, dass dem Akteur jede dieser Beschreibungen zugänglich ist. Es ist ebenfalls fragwürdig, ob sich aus jedem Sachverhalt, der während des Ereignisses besteht, eine Beschreibung des Ereignisses ableiten lässt. Hier sei noch einmal auf die Diskussion verwiesen, dass es nicht möglich ist, Ereignisse anhand ihrer Beschreibungen voneinander abzugrenzen, siehe Abschnitt 2.3 S.60.

Somit ist der Gehalt einer Intention zwar auf eine Handlung gerichtet. Aber der Gehalt einer Intention kann nicht das Ereignis in aller Vollständigkeit erfassen. Dass die Ausführung einer Intention in den seltensten Fällen fehlschlägt, liegt nicht nur daran, dass der Akteur akzeptiert, dass er es hervorbringen wird, sondern dass er ebenso eine Vorstellung hat, wie er es hervorbringen wird. Eine Intention kann sich zwar nur auf das Hervorbringen eines oder mehrere Sachverhalte beziehen, aber ein Erkennen ermöglichen, wie diese Sachverhalte hervorzubringen sind.

Auch wenn wir Intentionen selten im Deutschen mit 'dass'-Sätzen beschreiben, so können wir einem Akteur eine Intention mit einem 'dass'-Satz zuschreiben und sagen damit implizit wesentlich mehr. Der Akteur beabsichtigt, eine Pirouette zu vollführen, beinhaltet, dass der Akteur akzeptiert hat, dass er eine Pirouette vollführen wird, und dass er zudem eine Vorstellung dessen hat, wie er die Pirouette vollführen wird. Im Falle von weiter in der Zukunft liegenden Absichten bzw. Plänen impliziert eine Absicht, meist mehrere andere Absichten zu bilden bzw. Handlungen zu vollführen, vgl. am Ende dieses Abschnittes die Diskussion des sogenannten Mittel-Zweck-Prinzips. Wenn ich beabsichtige, nach Paris zu fliegen, beabsichtige ich, dafür zu sorgen, dass ich nach Paris fliege. Dass ich beabsichtige, dafür zu sorgen, impliziert ebenfalls eine gewisse Vorstellung, wie ich dafür Sorge, bzw. dass ich weitere Pläne fasse, wie ich dieses Dafür-Sorgen umsetzen werde. Z.B. beabsichtige ich, ein Ticket zu kaufen, meine Tasche zu packen, ein Taxi zum Flughafen zu nehmen, usw.

Eine Absicht des Akteurs, etwas zu tun, lässt sich dadurch beschreiben, dass der Akteur beabsichtigt, dass er dafür sorgt, dass er es tut. Dabei sollte ersichtlich bleiben, dass der Akteur nicht nur beabsichtigt, dass er dafür sorgt, sondern dass er sich in gewisser Weise damit verpflichtet, sich eine Vorstellung zu machen, wie er dafür sorgt. 'Wie er etwas tut' muss dem Akteur, wie Baier festhält, nicht in allen Details zugänglich sein. Wir haben Zugang zu unseren Fähigkeiten, sind jedoch nicht notwendigerweise in der Lage unsere Fähigkeiten, bestimmte Handlungen durchzuführen, bis

in jedes Detail in Propositionen und Sachverhalten zu erfassen [11].

In [11] kritisierte Baier die Beschreibung eines Gehaltes einer Intention mit „bringing about that“ wie folgt. Die Aussage ‘er beabsichtigt, den Jungen zur Raison zu bringen’ verliert nach Baier in der Formulierung ‘er beabsichtigt, dafür zu sorgen, ihn zur Raison zu bringen’, dass der Akteur auch weiß bzw. akzeptiert hat, was er tun muss bzw. wie er es tun muss, um die Intention zu erfüllen. Im ersten Beispiel ist es offensichtlich, dass der Akteur eine Vorstellung hat, wie er den Jungen zur Raison bringt. Im letzteren, meint Baier, ist es möglicherweise offen. Dann kann die Antwort auf die Frage, wie der Akteur dafür sorgt, ihn zur Raison zu bringen, mit einem ‘irgendwie’ beantwortet werden.

Dies ist jedoch keine korrekte Stit-Paraphrasierung. Wenn ein Satz eine Handlung ausdrückt, wird durch die Paraphrase des Satzes dieselbe Handlung bezeichnet. Es wird lediglich der Akteur zum Gehalt der Intention hinzugefügt, der im Falle jeder intendierten Handlung ausgezeichnet werden kann. Der Akteur beabsichtigt, dass er den Jungen zur Raison bringt, dann beabsichtigt er eine Handlung, die durch den Sachverhalt charakterisiert wird, dass der Junge von ihm zur Raison gebracht wird. Wenn er beabsichtigt, dass er dafür sorgt, dass er den Jungen zur Raison bringt, besagt dies genau dasselbe. Das Problem tritt sowohl im Englischen als auch im Deutschen auf, wenn Handlungen statt mit ‘dass’-Sätzen mit Infinitivkonstruktionen ausgedrückt werden, da diese den Akteur in gewisser Weise nicht als solchen auszeichnen. Wenn der Akteur dafür sorgt, den Jungen zur Raison zu bringen, dann kann das u.U. bedeuten, dass er dafür sorgt, dass es jemand anderes tut. Wenn er dafür sorgt, dass er selbst den Jungen zur Raison bringt, dann ist offensichtlich, dass er derjenige ist, der den Jungen zur Raison bringt. Daher sollte die Stit-Paraphrase nicht auf Infinitiv-Konstruktionen bauen, sondern mittels ‘dass’-Sätzen gebildet werden. Dann greift auch Baiers Kritik nicht. Wenn der Akteur beabsichtigt, dass er dafür sorgt, dass er den Jungen zur Raison bringt, kann es u.U. sein, dass er keine genaue Vorstellung hat, wie er ihn zur Raison bringt, weil der Moment den Jungen zur Raison zu bringen, noch nicht gekommen ist. Jedoch verpflichtet ihn die Absicht, dass er es ist, der dafür sorgt, dass er sich eine Vorstellung davon machen wird, wie er es tun wird, und dass er alles Nötige dafür ebenfalls tun wird. Im Falle von zukunftsgerichteten Intentionen ist es nicht erforderlich, dass der Plan, wie die Intention umzusetzen ist, bereits vollständig vorliegt.

Demzufolge lässt sich eine Aussage ‘der Akteur beabsichtigt etwas zu tun’ sich durch eine Stit-Paraphrase ‘der Akteur beabsichtigt, dass er dafür sorgt, dass er es tut’ äquivalent beschreiben. Somit ist es durchaus angebracht, einen Intentionoperator als

4. Überzeugungen, Wünsche und Absichten

einen modalen Operator zu beschreiben, der sich auf Propositionen bzw. Sachverhalte anwenden lässt, obwohl der Gehalt einer Intention nicht der Sachverhalt bzw. die Proposition ist, sondern auf das Hervorbringen des Sachverhaltes gerichtet ist. Indem wir den Gehalt jederzeit durch eine Stit-Paraphrase ersetzen können, ist es nicht notwendig, einem Intentionoperator eine besondere Form von Gehalt zuzuschreiben, die sich auf Handlungen richtet. Es ist nicht nötig, eine Ereignisvariable einzuführen, die sich nicht näher beschreiben lässt und somit nicht mehr Informationen liefert, als den Gehalt mittels einer Proposition zu beschreiben, die der Akteur hervorbringen wird. Es ist die Frage, ob bei einer Beschreibung einer Absicht es nicht ausreicht, dass sich der Gehalt auf eine Handlung richtet und dass der Akteur sich mit der Absicht in gewisser Weise verpflichtet, die Fähigkeit zur Handlung zu haben.⁸⁴

$$(TIA) \alpha dint:\varphi \supset \alpha dint:\alpha dstit:\varphi$$

Zusammengefasst wird eine Intention des Akteurs α , einen Sachverhalt φ hervorzu- bringen, durch $\alpha dint:\varphi$ geeignet beschrieben, solange man φ durch die zugehörige Stit-Paraphrase ersetzen kann. Dies zeigt, dass das Interaktionsaxiom (*TIA*) zwischen Handlungs- und Intentionoperator plausibel und notwendig ist.

Interaktionen von Intentions- und Handlungsoperatoren

In den vorangegangenen Abschnitten hatte ich dafür plädiert, dass eine Intention zu fassen, unter der Kontrolle des Akteurs liegt, so dass das Fassen einer Intention ein mentaler Akt ist.⁸⁵ Broome schreibt:

„[F]orming an intention this way is making a decision. Making a decision is as close to acting as reasoning can possibly get you.“ [33, S.407]

Dass der Akteur eine Intention hat, ist wiederum ein Sachverhalt, für den ein Akteur sorgt. Allerdings ist es notwendig, bei mentalen Ereignissen und dessen Sorge-Tragen- Dafür einen entsprechenden Handlungsoperator einzuführen.

In Kapitel 1 sind zwei Handlungsoperatoren diskutiert worden, der deliberative *dstit*-Operator und der Achievement *astit*-Operator. In Abschnitt 1.1 S.20 und in Abschnitt 3.2 S.82 ist festgehalten worden, dass eine Handlungsbeschreibung mittels

⁸⁴Belnap et al. hatten festgestellt, dass ein Komplement eines Satz ‘der Akteur beabsichtigt, dass ...’ eine Stit-Paraphrasierung erlaubt bzw. notwendig macht [19, S.13].

⁸⁵Mit dieser Meinung war ich nicht allein. Wallace [155, 156, 157] meint, dass eine Intention bzw. Volition die Komponente einer Handlungsmotivation ist, die aktiv durch den Akteur hervorgebracht wird. Ich stimme zu, dass sie aktiv hervorgebracht wird, aber nicht, dass die Intention ein notwendiger Teil einer Handlungsmotivation ist.

dstit-Operators der Beschreibung mittels *astit*-Operators vorzuziehen ist. Wenn Handlungen mit *dstit*-Operator beschrieben werden, stellt eine Menge historisch möglicher Situationen die Entscheidung des Akteurs dar.⁸⁶ Dies ist notwendig für die Positiv-Bedingung des *dstit*-Operators. Somit ist der Sachverhalt, für den der Akteur sorgt, auch in der aktualen Geschichte wahr ist. Auf der anderen Seite wird durch eine Negativ-Bedingung festgehalten, dass der Akteur aktiv etwas zu dem Bestehen beiträgt. Denn der Sachverhalt, für den der Akteur sorgt, besteht nicht historisch notwendigerweise, siehe Definition 3 S. 18. Diese Negativ-Bedingung beschreibt, dass der Zustand nach der Handlung hätte auch nicht eintreten können.

Diese Negativ-Bedingung ist für mentale Akte nicht sinnvoll. Am Ende des Abschnittes 3.3 war ich darauf eingegangen. Warum sollte es notwendigerweise einen möglichen Geschichtsverlauf geben, in der der Akteur den mentalen Zustand nicht gebildet hat?⁸⁷ Die Antwort lautet, es ist nicht notwendig. Ein Akteur fasst eine Absicht aufgrund der in ihm vorliegenden mentalen Zustände, wie Überzeugungen, Wünsche und anderen bereits bestehenden Intentionen etc. Es ist nicht nötig zu beschreiben, dass es eine Geschichte gibt, in der er die Intention nicht gehabt hätte. Wenn jemand Kontrolle darüber hat, einen mentalen Zustand in einer Situation zu bilden, so ist dies dieselbe Kontrolle, die er ausübt, wenn er den Zustand bildet bzw. wenn er es unterlässt diesen Zustand zu bilden, vgl. Abschnitt 3.3 88f.

Wenn das Fassen einer Intention ein mentaler Akt ist, liegt eine Intention nur dann vor, wenn der Akteur für das Fassen der Intention gesorgt hat. Die Annahme, dass es historisch gesehen anders hätte kommen können und er sich in einer aktual möglichen Situation anders hätte entschieden, spielt für die Frage, ob er aktual die Intention gefasst, hat keine Rolle. Denn es wird bereits durch das Haben der Intention ausgedrückt, dass er sich entschieden hat, die Intention zu fassen. Es ist nicht möglich, dass jemand anderes oder etwas anderes hätte dafür sorgen können, dass der Akteur die Intention fasst, außer dem Akteur selbst. Es ist somit in derselben Situation der Fall, nachdem er die Kontrolle ausgeübt hat und den Zustand gebildet hat, dass er sich hätte anders entscheiden können. Aber er hat entschieden bzw. dafür gesorgt, dass die Intention vor-

⁸⁶Ich verzichte auf die Bezeichnung als 'Wahl' des Akteurs, weil das Wort 'Wahl' eine gewisse Absichtlichkeit impliziert. In Kapitel 3.3 ist ausgeführt worden, dass für jeden vom Akteur hervorgebrachten bzw. entschiedenen Sachverhalt eine Proeinstellung (Intention, Wahl, etc) vorliegen muss. Eine 'Entscheidung' in diesem Sinne ist keine Form des 'Entscheiden für etwas'. Die unterschiedliche Verwendung soll an folgendem Beispiel deutlich werden. Ödipus entschied sich vorsätzlich, Iokaste zu heiraten. Damit entschied Ödipus ebenfalls, dass er seine Mutter heiratet.

⁸⁷Der Einfachheit halber spreche ich nur von der Bildung von mentalen Zuständen. Ähnliches gilt, für das Beibehalten bzw. Aufgeben von solchen Zuständen, solange es mentale Akte sind.

4. Überzeugungen, Wünsche und Absichten

liegt. Ob der Zustand der Intention in einer anderen möglichen Situation vorgelegen bzw. nicht vorgelegen hätte, spielt keine Rolle. Der mentale Akt des Fassens der Intention bedarf keiner Negativ-Bedingung. Wenn ein Zustand ‘der Akteur beabsichtigt, dass φ ’ vorliegt, hat der Akteur die Positiv-Bedingung des *dstit*-Operators erfüllt. Die Negativ-Bedingung ist irreführend bzw. irrelevant. Die Positiv-Bedingung des *dstit*-Operators wird durch den *cstit*-Operator beschrieben, der ein geeigneter Operator für die Darstellung mentaler Akte ist. Eine Intention in einer Situation zu haben, impliziert, dass dies ein Zustand nach einem mentalen Akt ist, dem Akt des Fassens der Intention.

(TAI)	$\alpha \text{ dint}:\varphi \supset \alpha \text{ cstit}:\alpha \text{ dint}:\varphi$
(\diamond NI)	$\Box \alpha \text{ dint}:\varphi \supset \perp$
(TII)	(i) $\alpha \text{ dint}:\beta \text{ dint}:\varphi \supset \perp$ (ii) $\alpha \text{ dstit}:\beta \text{ dint}:\varphi \supset \perp$ (iii) $\alpha \text{ dint}:\beta \text{ dstit}:\varphi \supset \perp$
(TII _k)	$(\diamond \alpha_1 \text{ dint}:\varphi_1 \wedge \dots \wedge \diamond \alpha_k \text{ dint}:\varphi_k)$ $\supset \diamond (\alpha_1 \text{ dint}:\varphi_1 \wedge \dots \wedge \alpha_k \text{ dint}:\varphi_k)$
(AIA _k)	$(\diamond \alpha_1 \text{ dstit}:\varphi_1 \wedge \dots \wedge \diamond \alpha_k \text{ dstit}:\varphi_k)$ $\supset \diamond (\alpha_1 \text{ dstit}:\varphi_1 \wedge \dots \wedge \alpha_k \text{ dstit}:\varphi_k)$

Dies motiviert weitere Interaktionsaxiome. Wenn der Akteur α und niemand sonst für den Sachverhalt gesorgt hat, ist es nicht möglich, dass ein anderer Akteur β für den Sachverhalt gesorgt hat, dass Akteur α eine Intention hat. Hierbei sollte man sich vor Augen halten, dass die Operatoren instantan sind, also in ein und demselben Moment interpretiert werden, die die Handlungen und Intentionen der Akteure beschreiben. Eine Formel $\alpha \text{ dstit}:\beta \text{ dint}:\varphi$ würde beschreiben, dass der Akteur α in der Situation dafür gesorgt hat, dass β beabsichtigt, dass φ . Das Fassen der Intention von β wäre somit teil einer Handlung, die α ausführt. Dies ist unsinnig. Eine Intention zu fassen, ist ein mentaler Akt, den β vollführt, und nicht α . Es kann keine Handlung α 's sein, dass β eine Intention fasst. Wenn es keine Handlung α 's sein kann, dass β eine Intention fasst, ist es nicht rational von α zu intendieren, im selben Moment dafür zu sorgen, dass β eine Intention fasst. Demzufolge ist auch (TII)(ii) motiviert.⁸⁸

⁸⁸Dies liefert bereits einen Ausblick, wie die Logik sinnvoll zu ergänzen ist. Würde man temporale Operatoren hinzunehmen, könnte man beschreiben, dass $\alpha \text{ dint}:\beta \text{ dint}:\varphi$. Der Akteur sorgt mit seiner Handlung dafür, dass zu einem späteren Zeitpunkt der Akteur β beabsichtigt, für φ zu sorgen. Dies ist hinsichtlich der strategischen Planung von gemeinsamen Handlungen und Interaktionen zwischen

Wenn die Logik mit dem spezifizierten *dint*-Intentionsoperator in Abschnitt 6.4 eingeführt wird, wird deutlich, dass das Theorem (*TII*) auf die Unabhängigkeit der Akteure hinsichtlich ihrer Handlungen zurückgeführt wird. Die Unabhängigkeitsbedingung der BT+AC Struktur führt zu dem Axiom (*AIA_k*), dass *k* Akteure unabhängig hinsichtlich ihrer Handlungen sind, siehe Abschnitt 1.2 S.16. Die Unabhängigkeit wird dabei wie folgt beschrieben. Wenn es für einen Akteur α_1 möglich ist, für einen Sachverhalt φ_1 zu sorgen, und es für einen von α_1 unabhängigen Akteur α_2 möglich ist, für φ_2 zu sorgen, muss es eine mögliche Situation geben, in der beide Akteure den jeweiligen Sachverhalt hervorbringen können. Die Möglichkeiten des einen Akteurs können nicht durch die Möglichkeiten des anderen Akteurs in ein und demselben Moment beschnitten werden, wenn die Akteure als unabhängig voneinander beschrieben werden, vgl. für die Motivation [19, S.217f].

Wenn angenommen wird, dass Akteure hinsichtlich ihrer Handlungen unabhängig sind, so ist es nur folgerichtig, dass sie es ebenfalls hinsichtlich ihrer mentalen Akte sind. Für welche Intention sich ein Akteur entscheidet bzw. für welche Intention er sorgt, schränkt die Möglichkeiten nicht ein, die ein anderer Akteur beim Fassen seiner Intentionen zu diesem Zeitpunkt hat, so dass eine logische Konsequenz des Axiomes (*AIA_k*) das Theorem (*TII_k*) ist, vgl. Abschnitt 6.4 S.302.

Auf einen Unterschied zwischen einem mentalen Akt und einer Handlung hatte ich bereits bei der Motivation von (*TAI*) hingewiesen. Im Falle des *dstit*-Operators ist aufgrund der Negativ-Bedingung ausgeschlossen, dass ein zu einem Zeitpunkt historisch notwendigerweise bestehender Sachverhalt durch den Akteur hervorgebracht werden kann. Dies führt dazu, dass eine Handlung keine historische Notwendigkeit sein kann, vgl. Abschnitt 1.2 S.20. Das Hervorbringen des Sachverhaltes durch α kann keine historische Notwendigkeit sein, siehe Theorem $\Box\alpha\text{dstit}:\varphi \supset \perp$.

Dies ist für mentale Akte und Intentionen unsinnig. Auch wenn ich im Weiteren argumentieren werde, dass es nicht rational ist, notwendigerweise bestehende Sachverhalte zu beabsichtigen, sollte dies nicht dazu führen, dass ein Akteur nicht historisch notwendigerweise eine Absicht formen kann. Dies ist in Schema ($\Diamond NI$) festgehalten. Es kann durchaus Zeitpunkte geben, in denen ein Akteur eine Intention in jeder aktual möglichen Situation fassen würde. Der Grund, dass hierbei zwischen mentalen Akten und Handlungen zu unterscheiden ist, liegt auf der Hand. Im Falle der Handlung muss es möglich sein, dass der nicht notwendigerweise bestehende Sachverhalt, der durch

unabhängigen Akteuren sinnvoll zu betrachten. Aber im selben Moment, wie in (*TII*)(ii) festgehalten, ist es nicht möglich, dass der Zustand der Intention von β ein mentaler Akt oder eine mentale Handlung α 's ist.

4. Überzeugungen, Wünsche und Absichten

die Handlung hervorgebracht wird, nicht jedes zu diesem Zeitpunkt eintretende Ereignis als Handlung des Akteurs beschreibt. Sein Dafür-Sorgen muss einen Beitrag zum Bestehen des Sachverhaltes leisten.

Somit gibt es zu diesem Zeitpunkt mögliche Ereignisse, die der Sachverhalt ebenfalls charakterisiert, die jedoch nicht eine Handlung des Akteurs beschreiben. Beispielsweise charakterisiert der Sachverhalt ‘dass Caesar stirbt’ den Mord, den Brutus begeht, im aktuellen Geschichtsverlauf. Also sorgt Brutus dafür, dass Caesar stirbt. Zu dem Zeitpunkt des Dafür-Sorgens hätte ein anderes Ereignis in einem möglichen Geschichtsverlauf stattfinden können, dass durch den Sachverhalt ‘dass Caesar stirbt’ charakterisiert wird, aber keine Handlung von Brutus ist. So hätte Caesar in diesem Moment an einem Herzinfarkt sterben können. Für mentale Akte zu fordern, dass der Sachverhalt, für den der Akteur sorgt, möglicherweise nicht besteht, ist nicht sinngebend. Ob ein anderes mentales Ereignis hätte stattfinden können als das, für das sich der Akteur entschied, die Intention zu bilden, ist für die Bildung der Intention irrelevant. Im Falle mentaler Akte ist die Zuschreibung hinreichend, dass der Akteur den mentalen Zustand gebildet hat, um den Akt zu beschreiben. Eine Negativ-Bedingung wie im Falle der nicht-mentalenen Handlungen ist nicht erforderlich.

Zwei Dinge sind unplausibel, wenn Handlungen als historische Notwendigkeiten betrachtet werden, die sich nicht auf mentale Zustände beziehen, sondern auf das Hervorbringen von Zuständen bzw. Sachverhalten außerhalb des Akteurs. Zum einen würde jedes mögliche Ereignis, das durch diesen Sachverhalt charakterisiert wird, in jedem möglichen Geschichtsverlauf notwendigerweise als eine Handlung des Akteurs beschrieben werden, diesen Sachverhalt hervorgebracht zu haben. Demzufolge müsste es sich um einen notwendigerweise bestehenden Sachverhalt handeln. Diese hervorbringen zu wollen, ist nicht rational, da sie bereits historisch notwendigerweise bestehen, ohne dass der Akteur dafür sorgen musste. Zum anderen wird dadurch nicht erklärt, was der Akteur beigetragen hat, wenn jedes Ereignis in jedem Geschichtsverlauf zu dem Zeitpunkt, das durch diesen Sachverhalt charakterisiert wird, notwendigerweise zu dem Bestehen des Sachverhaltes in jedem Geschichtsverlauf führt. Worin besteht dann das Zutun des Akteurs? Das Schema $\Box \alpha \text{dstit} : \varphi \supset \perp$ ist plausibel.

Für mentale Akte stellt sich ein anderes Bild dar. Die Frage, ob ein Akteur etwas dazu beigetragen hat, dass er eine Intention hat, stellt sich nicht. Er hat die Intention aktiv hervorgebracht. Er hat dazu beigetragen, sie zu fassen, unabhängig davon, ob man sein Beitragen voluntativ oder kausal fassen möchte. Demzufolge kann der erste Einwand für mentale Akte nicht geltend gemacht werden.

(UND)	$\Box \alpha dstit:\varphi \supset \perp$
(RN_{cstit})	$\vdash \varphi \Rightarrow \vdash \alpha cstit:\varphi$
(M_{dstit})	$\alpha dstit:(\varphi \wedge \psi) \not\supset (\alpha dstit:\varphi \wedge \alpha dstit:\psi)$
(RM_{dstit})	$\vdash \varphi \supset \psi \Rightarrow \vdash \alpha dstit:\varphi \supset \alpha dint:\psi$
(M_{cstit})	$\alpha cstit:(\varphi \wedge \psi) \supset (\alpha cstit:\varphi \wedge \alpha cstit:\psi)$
(RM_{cstit})	$\vdash \varphi \supset \psi \Rightarrow \vdash \alpha cstit:\varphi \supset \alpha cstit:\psi$
(K_{dstit})	$\alpha dstit:(\varphi \supset \psi) \supset (\alpha dstit:\varphi \supset \alpha dstit:\psi)$
(A_{dstit})	$(\alpha dstit:\varphi \wedge \alpha dstit:\psi) \supset \alpha dstit:(\varphi \wedge \psi)$
(K_{cstit})	$(\alpha cstit:(\varphi \supset \psi) \supset (\alpha cstit:\varphi \supset \alpha cstit:\psi))$
(A_{cstit})	$\alpha cstit:\varphi \wedge \alpha cstit:\psi \supset \alpha cstit:(\varphi \wedge \psi)$
(4_{cstit})	$\alpha cstit:\varphi \supset \alpha cstit:\alpha cstit:\varphi$
(4_{dstit})	$\alpha dstit:\varphi \supset \alpha dstit:\alpha dstit:\varphi$
(5_{cstit})	$\neg \alpha cstit:\varphi \supset \alpha cstit:\neg \alpha cstit:\varphi$
(5_{dstit})	$\neg \alpha dstit:\varphi \not\supset \alpha dstit:\neg \alpha dstit:\varphi$
(T_{cstit})	$\alpha cstit:\varphi \supset \varphi$
(T_{dstit})	$\alpha dstit:\varphi \supset (\varphi \wedge \neg \Box \varphi)$

Aus einem Gehalt eines mentalen Aktes lässt sich nicht auf die Wahrheit des Gehaltes schließen, jedenfalls nicht im Falle einer Intention. Aus dem notwendigerweise Hervorbringen der Intention folgt nicht, dass auch der Gehalt notwendigerweise besteht. Somit ist der zweite Grund, der im Falle von Handlungen für ein Axiom $\Box \alpha dstit:\varphi \supset \perp$ sprach, im Falle des mentalen Aktes unplausibel. Es ist daher möglich, dass das Fassen einer Intention ein notwendigerweise bestehender Sachverhalt ist, vgl. Theorem ($\diamond NI$), im Gegensatz zum Dafür-Sorgen-Dass von Sachverhalten $\Box \alpha dstit:\varphi \supset \perp$ (UND), das sich nicht auf mentale Zustände beschränkt. Wenn eine Intention historisch notwendigerweise gefasst wird, handelt es sich beim Fassen dennoch um einen mentalen Akt.

Die Unterscheidung von $cstit$ - und $dstit$ -Operatoren aufgrund der Negativ-Bedingung führt zu ähnlichen gültigen Schemata wie im Falle der int - bzw. $dint$ -Operatoren. Es ergeben sich die auf S.197 notierten Axiome und Theoreme von Handlungsoperatoren. Ich werde in Abschnitt 6.1 darauf zurückkommen. Eine vollständige Axiomatisierung und weitere Theoreme sind in [19, Abschnitt 17B] gelistet. Die Motivation für die jeweiligen Operatoren und die Anwendungen in anderen Gebieten wie der deontischen Logik sind in [19, Abschnitte 11 - 12] gegeben. Anhand der Liste wird deutlich, dass der $cstit$ -Operator keine Alternative ist, wenn man nicht nur mentale Akte, sondern Handlungen beschreibt.

4. Überzeugungen, Wünsche und Absichten

Axiom (5_{cstit}) ist für einen Handlungsoperator ungeeignet. Der *dstit*-Operator ermöglicht folgende Unterscheidung vorzunehmen. Wenn ein Akteur nicht dafür sorgt, dass ein bestimmter Sachverhalt besteht, weil entweder der Sachverhalt notwendigerweise besteht oder aktual nicht besteht, ist es nicht unbedingt eine Handlung des Akteurs, die dieses nicht Sorge-Tragen-Dafür beschreibt $\neg\alpha dstit:\varphi$. Eine Formel $\alpha dstit:\neg\alpha dstit:\varphi$ hingegen beschreiben nach Belnap et al. [19, Abschnitt 2.B.2] eine aktive Unterlassung, einen Sachverhalt hervorzubringen. Der Akteur hat dafür gesorgt, dass er nicht dafür sorgt, dass φ . Diese Unterscheidung von Unterlassung und einem Nicht-Dafür-Sorgen $\neg\alpha dstit:\varphi$ ist einer der Hauptgründe, sich für die Stit-Theorie als Logik zur Beschreibung von Handlungen zu entscheiden. Der *cstit*-Operator liefert diese Möglichkeit nicht. Als *S5*-Operator sind die Formeln im Falle des *cstit*-Operators äquivalent. Somit ist für die Beschreibung von Handlungen der *dstit*-Operator erste Wahl vor dem *cstit*-Operator. In Abschnitt 6.1 werde ich die Axiome mit *cstit*-Operatoren formulieren und die Theoreme für den *dstit*-Operator festhalten.

Ein Unterschied zwischen *cstit*- und *dstit*-Operatoren ist ihr Monotonie-Verhalten. Warum Handlungsoperatoren nicht monoton sind, lässt sich ebenso wie im Fall der Intentionen zeigen. Wenn ein Akteur für zwei Sachverhalte zusammen sorgen will, wobei der eine notwendigerweise besteht, sorgt er nur für den einen. Wenn jemand nach Paris fliegt und Charles de Gaulle dort ist, kann er dafür sorgen, dass er zu dem Zeitpunkt nach Paris fährt, zu dem Charles de Gaulle dort ist. Daraus kann nicht geschlossen werden, dass er dafür gesorgt hat, dass Charles de Gaulle zu dem Zeitpunkt dort ist, da er nicht dafür gesorgt hat, dass Charles de Gaulle dort ist. Er hat nur dafür gesorgt, dass er nach Paris fährt, als Charles de Gaulle dort ist. Die Konjunktion der Sachverhalte ist möglich hervorzubringen, auch wenn der eine Sachverhalt notwendigerweise besteht. Jedoch ist aus dem Hervorbringen der Konjunktion zweier Sachverhalte nicht abzuleiten, dass er für jedes Konjunktionsglied gesorgt hat.

Dies führt zu einem weiteren plausiblen Bild. Ein Akteur kann nicht für notwendigerweise bestehende Sachverhalte sorgen, z.B. dass Charles de Gaulle bereits zu dem Zeitpunkt in Paris verweilt. Wofür er sorgen kann, ist: Wenn ein Sachverhalt notwendigerweise besteht und der Akteur während des Bestehens dieses Sachverhaltes einen anderen Sachverhalt hervorbringt, so kann er die Konjunktion der Sachverhalte hervorbringen, ohne dass er den notwendigerweise bestehenden Sachverhalt einzeln hervorbringen können muss. Es ergibt sich folgendes Theorem für den *dstit*-Operator, der nicht monoton ist, $\alpha dstit:(\varphi \wedge \psi) \supset (\Box\varphi \vee \alpha dstit:\varphi)$. Es ist möglich zu beschreiben, dass der Akteur für eine Konjunktion aus einem notwendigerweise bestehenden und einem nicht notwendigerweise bestehenden Sachverhalt sorgt, ohne dass er ein-

zeln für den notwendigerweise bestehenden Sachverhalt gesorgt hat. Der Akteur kann dafür sorgen, dass er in Paris ist und Charles de Gaulle ebenfalls dort ist. Auch wenn der Sachverhalt, dass Charles de Gaulle in Paris ist, ein notwendigerweise bestehender Sachverhalt ist, den der Akteur nicht selbst hervorgebracht hat. Um dies durch einen Handlungsoperator zu beschreiben, ist es notwendig, dass dieser nicht monoton ist.

Ich hatte bereits erwähnt, dass Handlungsoperatoren sich nur auf Sachverhalte anwenden lassen sollten, die nicht notwendigerweise bestehen. Die obige Konjunktion besteht nicht notwendigerweise. Es gibt dafür zwei Gründe. Zum einen kann ein notwendigerweise bestehender Sachverhalt, der in allen möglichen Geschichten besteht, nicht als ‘durch eine Handlung in diesem Moment hervorgebracht’ zugeschrieben werden. Dass der Akteur in dem Moment dafür sorgt, wenn er historisch notwendigerweise besteht, ist in sich nicht plausibel.

Diesem Einwand kann man kritisch gegenüber stehen, wenn man festhält, dass der Akteur es in diesem Moment nicht kann, aber vielleicht zu einem früheren Zeitpunkt dafür gesorgt hat, dass der Sachverhalt in diesem Moment notwendigerweise besteht. Dies ist zurückzuweisen. In diesem Fall wäre der Zeitpunkt der Handlung ein anderer als der Zeitpunkt, an dem der Sachverhalt eingetreten ist. Wenn der Akteur ‘Hans’ zu einem Zeitpunkt t_0 einem Akteur ‘Peter’ einen tödlichen Stich versetzt und Peter erst drei Tage später an den Folgen dieser Verletzung stirbt, dann hat Hans nicht in dem Moment für den Tod von Peter gesorgt, in dem Peter stirbt, sondern zu dem Zeitpunkt t_0 als Hans Peter den tödlichen Stich zugefügt hat.⁸⁹ Ist der Tod Peters einer Handlung von Hans zuzuschreiben? Die eigentliche Handlung hat in dem Zeitpunkt t_0 stattgefunden. Der Sachverhalt, für den Hans gesorgt hat und der die Handlung als solche charakterisiert, ist aber nicht, dass Peter an den Verletzungen stirbt. Der Sachverhalt, für den Hans in t_0 sorgt, ist, dass Peter erst in drei Tagen an den Verletzungen stirbt. Wenn Hans in t_0 dafür sorgt, dass in allen seinen zu dem Zeitpunkt t_0 wahläquivalenten Geschichten es der Fall ist, dass Peter tatsächlich drei Tage später seinen Verletzungen erliegt und das Erliegen ursächlich dem Stich von Hans zuzuschreiben ist, hat Hans faktisch in dem Moment t_0 , als er den Stich ausführte, dafür gesorgt, dass Peter stirbt. In dem Moment t_0 , in dem er dafür sorgt, ist es bereits erfüllt, dass der Akteur dafür gesorgt hat, dass Peter drei Tage später stirbt. Zum Zeitpunkt t_0 des Stiches ist es jedoch keine historische Notwendigkeit, dass Peter drei Tage später an den Verletzungen stirbt. Nach Ausführung des Stiches steht fest, dass Peter stirbt, wenn Hans mit dem Stich dafür gesorgt hat. Der Geschichtsverlauf lässt sich nach einer Handlung des Ak-

⁸⁹Für diesen Einwurf danke ich Uwe Scheffler.

4. Überzeugungen, Wünsche und Absichten

teurs nicht mehr ändern, wenn der Sachverhalt in allen wahläquivalenten Geschichten eine Tatsache darstellt. Eine Tatsache stellt aber nur dar, dass Hans zugestochen hat.

Das Problem, das dieses Beispiel verdeutlichen soll, ist nicht, dass der Akteur nicht drei Tage vorher dafür gesorgt hat, dass der andere Akteur drei Tage später an den Folgen stirbt. Ein anderes Problem könnte sein, dass eine Zuschreibung eines Beobachters, der dem Akteur den tödlichen Stich als Handlung zuschreibt, erst nach drei Tage und nicht in t_0 erfolgen kann. Es geht jedoch nicht darum, ob und wann jemand einem Akteur das Hervorbringen eines Sachverhaltes als Handlung zuschreiben kann, sondern darum, ob der Akteur dafür gesorgt hat oder nicht. Dies ist unabhängig von der Frage, wann ein Sachverhalt 'Hans hat dafür gesorgt, dass' durch Hans oder durch einen anderen Beobachter als bestehend erkannt werden kann.

Wenn der Akteur für einen Sachverhalt, wie z.B. dass Peter stirbt, durch eine Handlung gesorgt hat, ist es im weiteren Geschichtsverlauf eine Notwendigkeit, dass der Akteur dafür gesorgt hat. In jedem späteren Zeitpunkt einer zur aktuellen Situation wahläquivalenten Situation besteht der Sachverhalt, dass Peter stirbt, notwendigerweise. Wenn es zu einem späteren Zeitpunkt als dem Stich noch eine Möglichkeit gäbe, den Tod Peters zu verursachen, ohne dass er an den Stichverletzungen stirbt, dann hat Hans sicherlich nicht ursächlich zum Tod Peters beigetragen. Und es kann ihm nicht als 'er hat dafür gesorgt' zugeschrieben werden, da es offensichtlich andere Umstände geben kann.⁹⁰

Unmögliches Beabsichtigen

Ausgangspunkt des Beispiels war die Frage, ob ein Akteur für einen bereits notwendigerweise bestehenden Sachverhalt sorgen kann. Die Antwort lautet nein. Möglicherweise hat er zu einem früheren Zeitpunkt dafür gesorgt, aber wenn er bereits notwendigerweise besteht, hat der Akteur keine Kontrolle mehr über das Bestehen des Sachverhaltes. Somit ist es unmöglich, für bereits notwendigerweise bestehende Sachverhalte zu sorgen. Ebenfalls unstrittig ist, dass man nicht für historisch unmögliche Sachverhalte sorgen kann.

⁹⁰Ob Hans für den Tod Peters verantwortlich ist, steht dabei nicht zur Diskussion. Wenn es nach dem Stich in einem wahläquivalenten Geschichtsverlauf noch die Möglichkeit gibt, dass Peter nicht an dem Stich, sondern anders verstirbt, dann hat in allen zu diesem Zeitpunkt möglichen Geschichtsverläufen, Hans nicht mit dem Stich für den Tod von Peter gesorgt. Dass dies möglicherweise unplausibel ist, lässt sich als Motivation verstehen, Operatoren instantan zu behandeln oder eine Handlungslogik mit temporalen Operatoren zu motivieren, um die Unterschiede zwischen dem Tod durch Erstechen und dem anderweitigen Tod von Peter darzustellen.

(PA)	$\alpha \text{ dint}:\varphi \supset \Diamond \alpha \text{ dstit}:\varphi$
	$\alpha \text{ dint}:\varphi \supset \Diamond \varphi$
Negativ-Bedingung	$\alpha \text{ dint}:\varphi \supset \neg \Box \varphi$

Diese Unmöglichkeit für historisch notwendigerweise bestehende bzw. für Sachverhalte, die in keiner aktual möglichen Situation bestehen, zu sorgen, impliziert bereits, dass es nicht rational ist zu intendieren, für solche Sachverhalte zu sorgen. Baier schreibt, dass ein Akteur sich der Grenzen seiner Möglichkeiten bewusst sein sollte. Was er beabsichtigt, sollte innerhalb seiner Handlungsmöglichkeit oder zumindest innerhalb seiner Fähigkeiten liegen [11].

„I cannot intend to stop the sun, nor can I intend to turn the moon [...] proper objects of intending, unlike proper objects of aiming at, seem limited to my actions [...] and to things I can do, though not necessarily to things I know I will succeed in doing on this occasion.“ [11, S.649]

Etwas zu beabsichtigen, was notwendigerweise der Fall ist, ist Verschwendung von Energie und Zeit. Da ein Akteur es nicht hervorbringen kann, wenn es notwendigerweise besteht, sind Sachverhalte, die notwendigerweise bestehen bzw. die unmöglich sind hervorzubringen, keine geeigneten Objekte, die es rational zu beabsichtigen gilt. Bratman schreibt, dass ein Akteur seine Intentionen unter bestimmte Rationalitätskriterien aufgrund dessen stellen sollte, weil ihm nur begrenzte Kapazitäten zur Verfügung stehen. Es ist die Aufgabe seiner Intentionen, die Pläne und Ziele in der ihm aktual möglichen Welt zu realisieren. Daher sollten Intentionen sich nicht auf unmögliche oder bereits notwendigerweise bestehende Sachverhalte richten. Hierbei ist das ‘Sollen’ mit weitem Skopus zu verstehen. Es soll nicht der Fall sein, dass, wenn man etwas intendiert, dies möglich sein soll. Allerdings soll man in der Lage sein, für den Sachverhalt zu sorgen, oder man soll die Intention aufgeben, den Sachverhalt hervorzubringen.

Für Bratman [25, 27], Velleman [147] und Wallace [156] ist nicht die Möglichkeit des intendierten Sachverhaltes ausschlaggebend, ob der Akteur für das Haben der Intention rational ist oder nicht. Sie fordern, dass im Akteur über den intendierten Sachverhalt bestimmte Überzeugungen vorliegen bzw. nicht vorliegen sollen. Dabei ist zwischen sogenannten schwache und starken Forderungen zu unterscheiden. Erstere werden von Davidson [48, Essay 5] und Bratman [25, Abschnitt 3.4], [27] vertreten. Sie postulieren, dass ein Akteur *nicht* glauben sollte, dass die von ihm inten-

4. Überzeugungen, Wünsche und Absichten

dierten Sachverhalte unmöglich sind. Diese Forderungen werden schwach genannt, da sie lediglich die Abwesenheit bestimmter Überzeugungen fordern. Bratman legt sich zudem fest, dass die Gehalte von Intentionen und Überzeugungen konsistent sein sollten.

Sogenannte starke Forderungen werden von Velleman [147] und etwas abgeschwächer von Wallace [156] aufgestellt. Sie verlangen das Vorliegen von bestimmten Überzeugungen, wenn der Akteur eine Absicht hat. Bratman bezeichnet sie als Kognitivisten [27, S.30]. Kognitivisten leiten die Konsistenz und Kohärenz von Intentionen aus der Konsistenz und Kohärenz von involvierten Überzeugungen ab. Vellemans Punkt ist, dass dies notwendig ist, damit die Gehalte von Intentionen erkannt werden können. Nur so können Intentionen konsistent sein. Wenn ein Akteur vollkommen unwissend über seine Intentionen wäre, dann kann er nicht sicherstellen, dass die Gehalte konsistent sind [150]. Wallace gehört für Bratman zu den sogenannten „supplemented cognitivists“. Intentionen sind keine Überzeugungen und werden auch nicht durch sie konsistent bzw. agglomerierend. Aber sie beinhalten die Überzeugung, dass der intendierte Sachverhalt durch den Akteur hervorgebracht werden kann. Nach Wallace muss ein Akteur von sich selbst glauben, dass es möglich für ihn ist, das zu tun, was er beabsichtigt [156, S.20]. Sonst hat er die Absicht nicht. Für Broome [33, 34] ist es möglich, dass der Akteur glaubt, das Intendierte nicht hervorbringen zu können, und er dennoch intendiert, den Sachverhalt hervorzubringen. Der Akteur ist jedoch für sein Halten der Intention oder der Überzeugung irrational.

Im Folgenden werde ich zeigen, dass die Forderung nach einer Überzeugung, die es verbietet, das Intendierte für unmöglich zu halten, nicht zum Ziel der Konsistenz und Kohärenz der Intentionen führt. So werde ich erst die starken und im Anschluss die schwachen Forderungen anfechten. Auch wenn ich Velleman in [147] nicht zustimme, dass eine Intention eine Überzeugung ist, so stimme ich ihm in [150] zu, wenn er von einer Intention als einer kognitiven Einstellung bzw. kognitiven Verpflichtung spricht.

Dagegen, dass eine Intention eine Überzeugung ist, spricht, dass Überzeugungen und Wünsche graduell sind. Diese Graduierungsmöglichkeit zeigt, dass es sich bei einer Intention nicht um einen Wunsch oder eine Überzeugung handeln kann. Wenn ich glaube, dass ich das tun werde, kann diese Überzeugung, mehr oder weniger stark sein. Wenn ich jedoch beabsichtige, das zu tun, werden alle weiteren Handlungen bzw. weiteren Intentionen beim Vervollständigen von Plänen so gefasst, dass diese Absicht als Prämisse fest angenommen wird.⁹¹ Eine Absicht ist nicht gradu-

⁹¹Bratman spricht von „Flat-out“-Überzeugungen, in denen der Akteur einen Sachverhalt als Gewissheit oder als gesetzt erachtet, auch wenn er u.U. nicht auf die Wahrheit des Sachverhaltes wetten würde [25,

ell. Sie kann keine Form von Überzeugung sein, dass ich besonders stark glaube, dass ich es tun werde. Davidson hat an Beispielen festgehalten, dass es durchaus möglich ist, am tatsächlichen Umsetzen der Absicht zu zweifeln und dennoch zu beabsichtigen, die Absicht umzusetzen [48, Essay 5]. Ich beabsichtige, dass ich 10 Kopien mittels Kohledurchschlag erstelle, aber ich glaube, dass ich maximal 8 Kopien damit erstellen kann.

Eine Intention ist eine Einstellung, die nach Velleman [149, Abschnitt 5] kognitiv ist und dem Akteur einen Sachverhalt als wahr zumachend abbildet. Sie ist nicht graduell. Somit ist die Akzeptanz der Proposition, dass ich dies oder jenes tun werde, ein geeigneterer Kandidat die kognitive Komponente einer Intention darzustellen als eine Überzeugung, vgl. Abschnitt 4.1. Diesen Gedanken hat Cohen bereits festgehalten [46, S.128]. Die weiteren Pläne und Handlungen formt ein Akteur nicht unter der Prämisse, dass er glaubt, dass er etwas tun wird, sondern unter der Prämisse, dass er akzeptiert, dass er es tun wird.

Ein Einwand von Cohen gegen die Behauptung, dass eine Intention eine bestimmte Überzeugung beinhalten muss, ist, dass es konzeptuell nicht möglich ist, dass eine Intention eine Überzeugung ist. Sein Grund ist nicht die Graduierung von Überzeugungen und die Nichtgraduierung von Intentionen, sondern dass es nicht möglich ist, dass ein mentaler Zustand, der durch einen mentalen Akt hervorgebracht wird, begrifflich eine Überzeugung implizieren oder beinhalten kann, wenn die Bildung der Überzeugung nicht ebenfalls ein mentaler Akt ist [46, S.127]. Wenn vorausgesetzt wird, dass ich keine Intention fassen kann, wenn ich nicht auch die Überzeugung habe, über die ich jedoch keine Kontrolle habe, um die Bildung als mentalen Akt zu zählen, kann ich nur die schwache Form der Kontrolle über Intentionen haben, die ich über Überzeugungen habe. Auch wenn dies einigen Autoren plausibel scheint, so habe ich doch in Abschnitt 3.3 und Abschnitt 4.1 auf die Unterschiede hingewiesen, zum einen bei der Kontrolle, die wir über das Fassen von Intentionen im Gegensatz zum Bilden von Überzeugungen haben, zum anderen bei der Urheberschaft, die im Falle des Fassens der Intention eher im Akteur und im Falle der Bildung der Überzeugung eher in der Welt liegt. So ist neben der Graduierung und der Aufgabe und Funktion einer Intention die Kontrolle, die wir über das Fassen von Intentionen haben, ein Grund, warum es sich bei Intentionen nicht um Überzeugungen handeln kann und warum eine Intention auch nicht das Vorliegen einer Überzeugung beinhalten kann.

S.36]. Dies ist meines Erachtens keine Form von Überzeugung, sondern von Akzeptanz. Akzeptanz ist wie eine Intention nicht graduell. Ich akzeptiere nicht mehr oder weniger. Ich akzeptiere etwas als wahr oder ich akzeptiere es nicht.

4. Überzeugungen, Wünsche und Absichten

Eine weitere Unplausibilität besteht in den starken Forderungen aufgrund dessen, dass wir uns sehr wohl über unsere Intentionen und Überzeugungen täuschen können. Angenommen ich glaube, dass ich die Überzeugung, dass φ habe. Nehme man an, für die Intention, dass ψ , zu fassen, ist es notwendig, die Überzeugung, dass φ , zu haben. Ich täusche mich jedoch darin, die Überzeugung, dass φ , zu haben. Tatsächlich bin ich nicht davon überzeugt, dass φ , auch wenn ich glaube, dass ich überzeugt bin, dass φ . Ich würde bei jedem Versuch, die Intention, dass ψ , zu fassen, immer wieder scheitern. Dies scheint mir unplausibel. Wenn ich eine Intention fasse, bilde ich den mentalen Zustand aus. Ich übe eine Art von Kontrolle aus. Wenn die Überzeugung 1. Stufe für die Intention notwendig ist und nicht vorliegt, kann ich den mentalen Akt, die Intention zu fassen, jedoch im Falle der starken Forderungen nicht vollführen. Ich wüsste nicht, ob ich die Intention gefasst habe oder nicht. Auch wenn es sicherlich möglich ist, sich hinsichtlich seiner Intentionen zu täuschen, so scheint es doch unplausibel, dass wir beim Fassen von Intentionen, die unsere Pläne umsetzen sollen, darauf warten müssen, die richtigen Überzeugungen zuvor zu bekommen. Dies macht Ansätze wie Velleman in [147] und Wallace in [155] problematisch, die Intentionen als Überzeugungen auffassen bzw. die Intentionen unterstellen begrifflich notwendigerweise Überzeugungen zu beinhalten.

Wenn die Überzeugung in den Forderungen von Wallace und Velleman durch Akzeptanz ersetzt werden würde, ist dies ein mentaler Akt, den der Akteur vollführen kann, und Cohens Gegenargument wäre nicht wirksam. Eine Überzeugung als notwendige Bedingung für das Bilden einer Intention begrifflich festzuhalten, ist jedoch zurückzuweisen.

Eine Intention ist nicht die Überzeugung, dass man etwas tun wird, wie Velleman es z.B. in [147, S.51] unter der Bezugnahme auf Harman [73] vertrat. Harman fasste eine Intention als eine Überzeugung auf, dass man etwas tun wird, die dazu führt, dass man es tatsächlich tut, so dass eine Intention eine bestimmte Art selbsterfüllender Überzeugung ist.⁹² Searle schlägt vor, dass der Gehalt einer Intention (Erlebnis des Handelns) das ist, was durch die Intention verursacht wird (die Handlung), diese sich somit quasi selbst verursacht [123, S.126, Schema S.128]. Aber deswegen ist eine Intention nach Searle keine Überzeugung.

Vellemans Grund in [147] eine Intention als Überzeugung aufzufassen, ist, dass er theoretische Rationalität auf praktische Rationalität zurückführen möchte.⁹³ Wel-

⁹²Dieser Aufsatz ist ursprünglich 1976 erschienen. Harman ist später in [72] von dieser Auffassung abgekommen.

⁹³In [74] argumentiert Harman am Beispiel des Wunschenkens dagegen. Es gibt Dinge, die einem Ak-

che Intentionen ein Akteur haben sollte, hängt an dem, was er glaubt, was nötig ist zu tun, bzw. was er erwartet, dass er als nächstes tut [147]. Da Velleman selbst von 'Erwarten' statt 'Überzeugt-sein' spricht, erinnert dies an eine Form der 'Akzeptanz'. Davidson spricht auch nicht von Überzeugung, sondern von „assume“ bzw. von einer Art „judgement“, dass man etwas tun wird, [48, Essay 5 S.100]. Es ist nach Davidson möglich, etwas zu beabsichtigen und gleichzeitig zu zweifeln, ob man es erreicht, bzw. zu glauben, dass man scheitert.⁹⁴ In [150] deutet Velleman ebenfalls an, dass er nicht notwendigerweise auf dem Begriff der Überzeugung bestehen würde [150, Fußnote 10]. Der Schluss, den ich daher aufgrund der Unterschiede zwischen Überzeugung und Intention ziehe, ist gerechtfertigt. Bei Intentionen handelt es sich nicht um Überzeugungen und auch nicht um mentale Zustände, die Überzeugungen beinhalten oder voraussetzen.

Ein anderes Argument ist es, dass aus Rationalitätsgründen der Akteur, wenn er etwas intendiert, eine bestimmte Überzeugung hat bzw. eben nicht hat. Wenn man allerdings konfligierende Überzeugungen zulässt, ist es auch zugelassen, dass der Akteur zweifelt, dass er die Handlung vollzieht, und gleichzeitig glauben, dass er die Handlung vollzieht. Nun habe ich nicht dafür plädiert, dass Überzeugungen konfligieren sollten, sondern dass es u.U. möglich und sogar rational möglich ist, konfligierende Überzeugungen zu haben. Diese Überzeugungen sollten aber nicht derart sein, dass die

teur praktische Gründe liefern zu handeln, aber durch theoretische Überlegungen von der Rationalität ausgeschlossen werden müssten. Er stellt umgekehrt die Frage, ob es möglich ist, dass Gründe, die einen Akteur als praktisch rational auszeichnen, es erlauben, zumindest von theoretischer Irrationalität im Falle des Wunschdenkens abzusehen. Theoretisches Schlußfolgern („theoretical reasoning“) beschäftigt sich nach Harman in erster Linie mit dem, was man glauben sollte. Praktisches Schlußfolgern („practical reasoning“) beschäftigt sich in erster Linie mit dem, was man tun, intendieren und planen sollte [74, S.49f]. Für Harman steht jedoch fest, dass sich dies nur *in erster Linie* darauf richtet, weil praktische Überlegungen Auswirkungen auf Überzeugungen haben und umgekehrt. Harman argumentiert dafür, dass es sich bei Zielen in der theoretischen Rationalität, wie z.B. Einfachheit, auch um ein praktisches Ziel handeln kann. Es soll damit nicht gezeigt werden, dass theoretische Rationalität durch praktische Rationalität begründet wird, sondern lediglich dass eine Rückführung in der anderen Richtung möglicherweise ebenfalls nicht adäquat ist. Ich unterscheide nicht zwischen praktischer und theoretischer Rationalität, sondern wann ein Akteur für eine Überzeugung, Intention bzw. sonstige Einstellung als rational betrachtet werden kann, unabhängig ob praktische oder theoretische Überlegungen dazu führten. Auch wenn mir bewusst ist, dass das eine möglicherweise problematische Vereinfachung ist, da jemand für eine Überzeugung als theoretisch rational und für das Hervorbringen der Überzeugung als praktisch nicht rational bewertet werden kann. Das Beispiel von Engel [54] könnte ein solches Szenario sein. Ein Akteur lernt ein Telefonbuch auswendig, um möglichst viele wahre Überzeugungen zu sammeln. Er hat jedoch keine praktische Gründe dafür.

⁹⁴Vgl. Davidsons Versuch zehn Durchschriften anzufertigen [48, Essay 5 S.92].

4. Überzeugungen, Wünsche und Absichten

Einstellungen des Akteurs und seine Handlungen die Gehalte der rational konfligierenden Überzeugungen darstellten. Auch wenn man sich über seine Überzeugungen täuschen kann, ist es nicht rational, konfligierende Überzeugungen über die eigenen Handlungen zu haben. 'Ich glaube, dass ich x tun werde, und ich glaube, dass ich x nicht tun werde' drückt eine Unentschlossenheit aus, die eher darauf zurückzuführen ist, dass der Akteur nicht intendiert, dass eine oder das andere zu tun. Somit sind konfligierende Überzeugungen rational haben zu können, kein hinreichendes Kriterium um zu widerlegen, dass das Haben und Nichthaben von Überzeugungen ein Rationalitätskriterium für Intentionen ist. Ich werde im Folgenden jedoch genau diese Ansicht in Zweifel ziehen und Rationalitätskriterien für Intentionen begründen, die unabhängig von den Überzeugungen sind, die im Akteur vorliegen.

Es kann rational sein, etwas zu beabsichtigen und zu glauben, dass man es nicht tun wird, nämlich genau dann, wenn die Überzeugung, es nicht zu tun, dazu führt, über Mittel und weitere Handlungen zu rasonieren, die für den Erfolg der Handlung notwendig sind. 'Ich beabsichtige, dass ich x tue, und glaube, dass ich x nicht tue' klingt nicht irrational oder paradox. Ich beabsichtige, die Prüfung zu bestehen. Jedoch glaube ich, dass ich die Prüfung nicht bestehe, bzw. bezweifle ich, dass ich die Prüfung bestehe. Paradox klingt hingegen: 'Ich beabsichtige, dass ich die Prüfung bestehe, und akzeptiere, dass ich die Prüfung nicht bestehe.' Ersterer Gegensatz lässt sich möglicherweise darauf zurückführen, dass der Akteur beabsichtigt, die Prüfung zu bestehen, jedoch noch nicht sicher ist, ob er dies mit Erfolg umsetzen kann. Die Zweifel in Form der Überzeugung, dass er die Prüfung nicht besteht, sind möglicherweise rational, weil er im Moment noch schlecht vorbereitet ist.

Somit kann ein Grund für die Überzeugung sein, dass er x nicht tun wird, dass ihm notwendige Mittel nicht bzw. noch nicht zur Verfügung stehen. Er akzeptiert zwar, dass er es tun wird, aber er ist überzeugt, dass es unter den gegebenen Umständen nicht möglich ist, es zu tun. Wenn er rational ist, dann sollte er möglicherweise weitere Handlungen akzeptieren bzw. beabsichtigen, dass er mehr lernen sollte, sich mehr Zeit nehmen sollte usw. Somit ist die Überzeugung, die Prüfung zu bestehen, nicht notwendig. Im Gegenteil es kann sogar rational sein, zu glauben, dass man es nicht tun wird, da dies zur Reflektion über weitere notwendige Mittel und zur Intention führt, diese Mittel hervorzubringen.

Das Zweifeln am Gelingen bzw. die Überzeugung, dass es vielleicht nicht möglich ist, es hervorzubringen, führt eher zur Kohärenz und Vollständigkeit hinsichtlich der Intentionen, als die Überzeugung, dass man x tun wird. Wenn der Akteur überzeugt ist, die Prüfung zu bestehen, beabsichtigt er möglicherweise, die Prüfung zu bestehen.

Aber es hilft nicht beim erfolgreichen Bestehen der Prüfung. Daher kann es bei zukunftsgerichteten Intentionen rational sein, eine kognitive Einstellung wie Akzeptanz zu haben, dass er etwas tun wird, obwohl er überzeugt ist, dass er es möglicherweise nicht tun wird. Man besteht eine Prüfung nur, wenn man tatsächlich alles Erforderliche getan hat. Dies impliziert nicht, dass der Akteur glauben soll, dass er die Möglichkeit hat, sondern dass es tatsächlich möglich sein sollte, dass er die Möglichkeit hat. Erst dann ist ein Akteur meines Erachtens rational für seine Intentionen. Eine sinnvolle Rationalitätsforderung ist somit, dass ein Akteur eine Intention, etwas zu tun, nur rational fassen kann, wenn er die Möglichkeit hat, die intendierte Handlung umzusetzen, vgl. Schema (*PA*), bzw. wenn er bereit ist, die notwendigen Mittel zu intendieren, vgl. Schema (*MEC*) S.215. Falls die Handlung nicht möglich ist oder er die Mittel nicht intendieren möchte, sollte er aus Rationalitätsgründen die Intention aufgeben. Mir ist bewusst, dass diese Rationalitätsforderungen externe Bedingungen aufstellen. Da eine Intention eine kognitive Verpflichtung bereits in Form der Akzeptanz beinhaltet, sind diese externalen Forderungen nicht abwegiger als Forderungen, die an Überzeugungen von manchen Autoren gestellt werden. Z.B. es soll nur geglaubt werden, was wahr ist [149, Abschnitt 11], bzw. es soll genau das geglaubt werden, was wahr ist [160]. Im Gegenteil, da wir im Falle von Intentionen und Akzeptanz Kontrolle über die Bildung dieser mentalen Zustände ausüben können, sind solche externalen Forderungen für Intentionen nicht den Anfechtungen ausgesetzt wie im Falle der Überzeugungen, vgl. Abschnitt 4.1 S.145f die Diskussion zu Schema (*A_b*).

Zusammengefasst wird Folgendes festgehalten. Aufgrund der Unmöglichkeit, Absichten graduell zu fassen bzw. aufgrund der Kontrolle, die wir über Absichten aber nicht über Überzeugungen haben, kann die kognitive Komponente einer Intention keine Form der Überzeugung sein.⁹⁵ Eine Form der Akzeptanz, dass ein bestimmter Sachverhalt als hervorzubringen erachtet wird, bzw. die Erwartung, dass eine Handlung zu vollziehen ist, ist eher geeignet, die kognitive Einstellung im Falle einer Intention zu erfassen, zumal es sich dabei ebenso wie im Falle des Fassens einer Intention um einen mentalen Akt handelt.⁹⁶

⁹⁵Die Inklusionsbeziehung zu Wünschen war nicht unter deskriptiven Aspekten gefasst. Auch die konative Komponente einer Intention kann kein Wunsch sein. Ich würde (*DI*) jedoch unter der Maßgabe aufrecht erhalten, dass eine Absicht eine Proeinstellung im Sinne von positiv konativer Einstellung zum Hervorbringen des Sachverhaltes ist. Auch wenn ein Wunsch zu stark ist, so impliziert eine Intention, dass der Akteur dem Hervorbringen des Sachverhaltes positiv gegenüber eingestellt sein sollte, um motiviert zu sein, die Handlung auszuführen. Dies sollte ihn dazu bringen, den Wunsch zu dem intendierten Sachverhalt zu bilden.

⁹⁶In [25, S.36] formuliert Bratman zwar, dass es bei Plänen bzw. Intentionen zu sogenannten „Flat-Out-

4. Überzeugungen, Wünsche und Absichten

Eine Intention ist keine und impliziert keine Überzeugung und setzt auch keine voraus. Somit ist ein Axiomenschema (*IB*) zurückzuweisen, wenn es an deskriptiven Aspekten begrifflicher Zusammenhänge zwischen Intention und Überzeugung festgemacht werden soll. Dass es auch nicht rational ist zu fordern, dass ein Akteur glauben sollte, dass er den intendierten Sachverhalt hervorbringt, ist ebenfalls offensichtlich an Davidsons Beispiel. Dass man zweifelt, ob man zehn Kopien mit einer Durchschrift machen kann, die man zwar intendiert, aber nicht glaubt, dass dies erfolgreich sein wird, ist u.U. rational.

$$(IB) \alpha \text{ dint}:\varphi \supset \alpha \text{ bel}:\alpha \text{ dstit}:\varphi$$

Die Lesart von (*IB*) mit weitem Skopus ist ebenfalls keine Rationalitätsforderung. Wenn der Akteur nicht überzeugt ist, dass er den Sachverhalt hervorbringen wird, ist es keine rationale Forderung, die Intention sofort aufzugeben. Dies würde bedeuten, der Akteur sollte glauben, dass er die Prüfung besteht, oder die Intention aufgeben, die Prüfung zu bestehen. Es ist nicht rational zu verlangen, die Intention aufzugeben, die Prüfung zu bestehen, wenn er nicht glaubt, dass er die Prüfung besteht. Die Zweifel an einem Hervorbringen des Sachverhaltes motivieren ihn, dass notwendige Mittel intendiert werden, um den Sachverhalt doch noch erfolgreich hervorzubringen. Darauf sollte Rationalität bei Intentionen ausgerichtet sein. Eine Intention hat nicht die Aufgabe, die Motivation für die Handlung zu liefern. Eine Intention hat die Aufgabe, die Frage nach dem ‘wie kann ich meine Ziele erreichen’ zu beantworten. Sie liefert keine Antwort auf die Frage, was ich denke, was möglich ist, sondern wie ich erreichen kann, worauf ich abziele bzw. was ich außerdem beabsichtigen sollte. Im Falle von Plänen kann es sein, dass eine Intention gefasst wird, um einer anderen Intention den Erfolg zu sichern. Darauf laufen Bratmans Forderungen noch konsistenten und kohärenten Plänen hinaus [25, S.31]. Dies kann dazu führen, dass ich von Sachverhalten glaube, dass ich sie nicht hervorbringen kann. Aber es ist diese Überzeugung, die mich dann

Beliefs“ kommt. Er sagt jedoch auch, dass es eine Form der Akzeptanz ist. Wenn ich glaube, dass ich nur ein Auto habe, sollte ich meine Pläne nicht so fassen, dass ich zwei Autos brauche. Ich sollte es als gesetzt ansehen, dass ich ein Auto habe und mit einem Auto meine Pläne schmiede, auch wenn ich möglicherweise nicht auf die Überzeugung wetten würde, dass ich nur ein Auto habe. Bratman spricht davon, dass dieser Flat-out-Belief nicht mit der Gewissheit (m.E. Grad der Überzeugung) zusammenhängt, die ich in die Proposition habe, sondern mit der Kohärenz, die meinen Plänen zugrunde liegen sollte (m.E. Akzeptanz dessen, was ich hervorbringen werden). „What makes my attitude toward my having just one car one of my flat-out-belief, [...] is, at least in part, its distinctive role in the background of my further planning – in particular its role in providing a screen of admissibility of my options“ [25, S.37].

motiviert, weitere Intentionen zu fassen. In Schema (*BMEC*), vgl. [27, S.53f] bzw. (*MEC*) S.215, ist festgehalten, dass der Akteur zu einer Intention notwendige Mittel ebenfalls intendieren sollte. Wenn er jedoch glaubt, dass er die Handlung ausführen wird, was sollte ihn motivieren, weitere Mittel zu intendieren? Ein Zweifel an dem Hervorbringen kann ihn eher motivieren, die notwendigen Mittel zu intendieren, als die Überzeugung, dass er es hervorbringen wird. Grundvoraussetzung ist, dass eine Intention eine kognitive Einstellung ist, die den Akteur als wahr akzeptieren lässt, dass er den Sachverhalt hervorbringen wird. Hiermit widerspreche ich Bratman. Aber eine Überzeugung, dass er es tun wird, ist weder hilfreich noch ist sie rational zu fordern. Somit ist (*IB*), d.h. die Forderung, dass eine Intention eine Überzeugung beinhaltet, dass der Akteur das Intendierte tun wird, weder anhand von Rationalitätsforderungen noch aufgrund begrifflicher Zusammenhänge von ‘Intention’ und ‘Überzeugung’ zu begründen.

Ebenfalls sind rationale Forderungen nicht adäquat, die postulieren, dass es nur rational ist, eine Intention zu haben, wenn man glaubt, dass dies möglich ist, vgl. Schema (*IB*◇), oder man nicht glaubt, dass es unmöglich ist, Schema (*I*→*B*◇). Aufgrund der Möglichkeit, an der eigenen Unfehlbarkeit sinnvoll und rational zu zweifeln, sind beide Schemata keine sinnvollen Rationalitätsforderungen.

Wenn der Akteur beabsichtigt, dass φ , sollte er glauben, dass es möglich ist, den Sachverhalt hervorzubringen oder die Absicht aufzugeben [34]. Gegen dieses Schema ist nichts einzuwenden, wenn man konfligierende Überzeugungen für Intentionen zulässt. So kann der Akteur glauben, dass es möglich ist, und glauben, dass es nicht möglich ist. Auf der anderen Seite bestand der Sinn einer Rationalitätsforderungen nach Bratmans Definition von Kognitivismus, auf die Konsistenz und Kohärenz von Intentionen hinzuwirken. Dafür wäre dieses Schema ungeeignet. Es sichert nicht, dass die Gehalte der Intentionen realisierbar in der aktuellen Welt sind. Für eine solche Realisierbarkeit der Gehalte spricht eine Rationalitätsforderung wie (*PA*), die an die Möglichkeiten bzw. die Fähigkeiten geknüpft ist, die der Akteur tatsächlich hat. Erst wenn es tatsächlich möglich ist, dass der Akteur seine intendierte Handlung durchführen kann, ist sichergestellt, dass die Handlung realisiert werden kann. Da, wie Anscombe festhält, die Ausführung von Intentionen in den seltensten Fällen fehlschlägt, ist dies keine utopische Forderung. Anscombe spricht von praktischem Wissen und nicht von praktischen Überzeugungen. Wir wissen, wie unsere Intentionen umgesetzt werden können, aber sind nicht notwendigerweise überzeugt, dass wir es erfolgreich tun werden [7, § 48].

$$(IB\Diamond) \alpha \text{ dint}:\varphi \supset \alpha \text{ bel}:\Diamond\alpha \text{ dstit}:\varphi$$

4. Überzeugungen, Wünsche und Absichten

Ein Schema ($IB\Diamond$) ist nicht unsinnig. Es gibt jedoch keinen Grund, es als eine sinnvolle Rationalitätsforderung zu betrachten. Es gibt aber auch keinen zwingenden Grund, es als nicht haltbar zu titulieren. Da (PA) eine wesentlich sinnvollere Forderung im Hinblick auf Konsistenz und Kohärenz von Intentionen ist, werde ich mich auf (PA) berufen. Es sollte so sein, dass ein Akteur eine Intention nur hat, wenn es ihm möglich ist, den intendierten Sachverhalt hervorzubringen. Hampshire bringt ein Beispiel, dass es nicht möglich ist zu beabsichtigen, hundert Yards in acht Sekunden zu laufen. Man kann es nur versuchen, und beabsichtigen, so schnell, wie man halt kann, zu laufen. Wer behauptet, er würde es beabsichtigen, irrt sich [71]. Nach Hampshire [71] ist es paradox zu formulieren, 'ich beabsichtige es zu tun, aber ich werde es nicht tun, weil ich glaube, dass es unmöglich ist'. Doch was den Satz paradox macht, ist nicht die Aussage, dass er glaubt, dass es unmöglich ist, sondern die Aussage, dass er das, was er beabsichtigt, nicht tun wird. Eine Absichtserklärung ist eine Aussage, darüber was ich tun werde, siehe z.B. [7, § 2] „I intend to go for a walk but shall not go for a walk' does sound in some way contradictory“. Warum wird er es im obigen Beispiel nicht tun? Er kann hundert Yards nicht in acht Sekunden laufen. Er weiß, dass er es nicht kann, demzufolge wäre zu beabsichtigen, es zu tun, nicht rational. Dass es psychologisch unmöglich ist, möchte ich nicht behaupten. Möglicherweise beabsichtigen manche Akteure, Sachverhalte hervorzubringen, von denen sie wissen, dass sie unmöglich sind. Sie sind jedoch nicht mehr rational, wenn sie diese Intention nicht aufgeben.

Eine schwächere Forderung als (PA) und von Bratman als nicht-kognitivistische Annahme bezeichnet ist ($I\rightarrow B\Diamond$). Wenn der Akteur beabsichtigt, so sollte der Akteur nicht glauben, dass es nicht möglich ist, den Sachverhalt hervorzubringen, bzw. wenn er glaubt, dass es nicht möglich ist, dann soll er die Intention aufgeben [25, 27]. Davidson formuliert dies ähnlich, dass die Überzeugungen eines Akteurs seinen Intentionen nicht widersprechen sollten [48, S.100f]. Dass diese Forderung nicht zielführend ist, ist bereits am Beispiel der Prüfung deutlich geworden. An dieser Stelle gehe ich nur kurz auf die Probleme eines Nicht-Kognitivisten ein, der eine Intention frei von jeder kognitiven Einstellung sprechen möchte und gleichzeitig ($I\rightarrow B\Diamond$) vertritt.

$$\boxed{(I\rightarrow B\Diamond) \alpha \text{ dint}:\varphi \supset \neg\alpha \text{ bel}:\neg\alpha \text{ dstit}:\varphi}$$

Bratman sieht die Notwendigkeit, konsistent und in gewisser Weise kohärent mit seinen Überzeugungen zu sein, weil Intentionen ermöglichen sollen, Pläne zu fassen, wie bestimmte Ziele in Zukunft erreicht werden [25, S.31f]. Um die Effektivität und Koordination von Handlungsabläufen zu gewährleisten, die bei der Umsetzung solcher Pläne notwendig ist, fordert Bratman, dass Pläne intern konsistent sind. „Internal

consistent“ versteht Bratman in seinen Betrachtungen als konsistent mit den Überzeugungen, die ein Akteur hat. Hier ist jedoch der kritische Punkt, der meines Erachtens Bratman selbst wieder als Kognitivisten auszeichnet. Bratman meint. Da Überzeugungen, wenn wir rational sind, konsistent sein sollten, sind die einzigen Rationalitätsforderungen an Intentionen, dass ein Akteur von allem, was er intendiert, es nicht für unmöglich hält, es hervorzubringen, und dass eine gewisse Kohärenz hinsichtlich der Intentionen und ihren notwendigen Mittel besteht. Letzteres soll durch das sogenannte Mittel-Zweck-Prinzip, siehe [25, S.38], verdeutlicht werden.

Die erste Forderung hält Bratman als asymmetrische These fest [25, S.37f]. Eine Intention, dass φ , wird unter normalen Umständen die Überzeugung unterstützen, dass man φ hervorbringen wird. Es liegt jedoch keine Form von Irrationalität vor, wenn diese Überzeugung nicht vorliegt. Irrationalität liegt laut der These Bratmans erst vor, wenn man intendiert, dass φ , aber glaubt, dass man φ nicht hervorbringen kann. Daher ist nach Bratman die Inkonsistenz der Gehalte von Intentionen und Überzeugungen ($I \neg B \diamond$) näher an einer kritikwürdigen Irrationalität als eine Unvollständigkeit zwischen diesen Gehalten von Intentionen und Überzeugungen, wie sie durch ($IB \diamond$) ausgeschlossen wäre.

Aber erreicht Bratman mit dieser Rationalitätsforderung ($I \neg B \diamond$), dass ein Akteur seine Pläne effektiv gestaltet und koordiniert. Bratman schreibt:

„It should be possible for my entire plan to be successfully executed. Further, a good coordinating plan is a plan for the world I find myself in.“
[25, S.31]

Die Funktion eines Planes und somit einer Intention, als Teil eines Planes, ist die Koordinierung und erfolgreiche Durchführung von Handlungen. Um diese Rolle zu erfüllen, ist es angebracht, Rationalitätsforderungen an Intentionen zu stellen, die an Wünsche und Bewertungen nicht sinnvoll gestellt werden können [25, S.32]. Hierin stimme ich mit Bratman überein.

Neben der Konsistenz eines solchen Planes ist es die Kohärenz vom Ziel und den dafür für notwendig erachteten Mitteln, die das Haben bzw. die Abwesenheit bestimmter Intentionen als rational oder irrational auszeichnen, vgl. [25, S.31], [27, S.29,53f] bzw. Schema (*BMEC*). Kritik wurde von Velleman an Bratman geübt. Der Akteur hat keinen Zugang zu seinen Intentionen, wenn er nicht in irgendeiner Weise eine kognitive Einstellung zu den Gehalten seiner Intentionen hat, da bei Bratman Intentionen selbst keine kognitive Einstellung sind [150, S.206]. Eine Überzeugung, dass man etwas nicht tun wird, nicht zu haben, macht nach Velleman eine Intention, dies zu tun, nicht mehr oder weniger rational.

4. Überzeugungen, Wünsche und Absichten

Die Frage ist, ob Bratman mit der Forderung, dass, wenn die Intention, dass φ , vorliegt, die Überzeugung, dass der Akteur φ nicht bewirken wird, tatsächlich dazu führt, dass die Konsistenz und somit die Möglichkeit der Umsetzung des vollständigen Planes gesichert ist. Insbesondere ist das dann nicht der Fall, wenn die Überzeugungen selbst rational konfligierend sein können. Dies lehnt Bratman ab.

Bratman lehnt jedoch auch den Kognitivismus ab. Laut seiner Definition des Kognitivismus sind es die Forderungen nach Konsistenz und Kohärenz von in Intentionen involvierten Überzeugungen die Konsistenz und Kohärenz der Intentionen bestimmen. Wenn Bratman sich nicht selbst unter die Kognitivisten zählen möchte, dann sollte die Konsistenz und Kohärenz von Intentionen nicht auf die Konsistenz und Kohärenz von als abwesend geforderten Überzeugungen zurückzuführen sein. Die Abwesenheit von bestimmten Überzeugungen mit den Überzeugungen, die sonst im Akteur vorliegen, machen den Plan nach Bratman konsistent und vollständig aufgrund seiner Mittel-Zweck-Forderung (*BMEC*), die ebenfalls wieder auf Überzeugungen basiert.

Wenn man den Skopus des Sollens sowohl in (*BMEC*) als auch in ($I \rightarrow B \diamond$) weit lesen würde, dies tut Bratman, impliziert die Forderung, dass, wenn die Gehalte von Intentionen und Überzeugungen nicht konsistent sind, dass die Intention aufzugeben ist. Dies würde dafür sorgen, dass die Überzeugung, dass es nicht möglich ist, etwas zu tun, zum Aufgeben der Intention und somit wieder zur Konsistenz von Intentionen und Überzeugungen führt. D.h. es ist wieder die Konsistenz der involvierten Überzeugungen, aus der sich die Konsistenz der Intentionen ableitet. Ohne involvierte Überzeugungen sind die Intentionen nicht konsistent. Dass Überzeugungen im Fall der Intentionen vorliegen müssen, ist nach Bratman jedoch eine Form von Kognitivismus. D.h. entweder Bratman muss die Forderung nach der Konfliktfreiheit der Gehalte von Überzeugungen oder die Forderungen (*BMEC*) oder ($I \rightarrow B \diamond$) aufgeben oder sich selbst unter die Kognitivisten zählen.

Bratman in [27] versucht sich gegen die Argumente, dass er selbst Kognitivist ist, zu wahren. Er stellt die Bedingung auf, dass Intentionen agglomerieren, aber nicht aufgrund von Überzeugungen, die die Intentionen beinhalten, sondern weil die Intentionen a) an eine einheitliche Welt angepasst sein sollten und b) weil eine Intention darauf abzielt, dass sie ihren Teil zu einer Realisierung eines vollständigen Planes in der vom Akteur für real erachteten Welt beiträgt [27, S.52]. Nach Bratman „a projected unity of agency helps explain the rational pressure toward agglomeration of intention“ [27, S.52]. Hier ist jedoch Bratman entweder versteckter Kognitivist oder er muss (*PA*) und (*MEC*) zustimmen. In den Schemata (*PA*) und (*MEC*) sind keine Überzeugungen involviert. Wenn die projizierten Intentionen agglomerieren, weil sie

4.3. Absichten und Handlungen

eine Einheit der tatsächlichen Welt abgeben, dann sollte Bratman meine Forderung unterstützen, dass es rational ist, nur das zu intendieren, was möglich ist, oder die Intention anderenfalls aufzugeben. Dies besagt (*PA*).

Bratman projiziert jedoch die Intentionen und Handlungen eines Akteurs in eine mögliche Welt, die kompatibel mit dem ist, was der Akteur glaubt,⁹⁷. Diese Welt ist jedoch nur konsistent und kohärent, wenn die Überzeugungen ebenfalls konsistent und kohärent sind. Wenn Bratman nun die Konsistenz und Kohärenz der Intentionen aus der Konsistenz der projizierten Welt ableitet, leitet Bratman die Forderungen der Konsistenz und Kohärenz für Intentionen aus der Konsistenz und Kohärenz von Überzeugungen ab, auch wenn er von einer Projektion auf eine Welt spricht. Er müsste sich selbst wieder als Kognitivist bezeichnen, da diese Intentionen nur konsistent bestehen, weil sie eine aus den Gehalten von Überzeugungen projizierte Welt beinhalten.

Ein Nicht-Kognitivist in Bratmans Sinn sollte demzufolge die Konsistenz und Kohärenz nicht aus den Überzeugungen ableiten. Die Frage, die im Raum stehen bleibt, ist: Woher leitet sich die Forderung der Konsistenz und Kohärenz von Intentionen ab? Die einzig mögliche Antwort ist, dass die reale Welt, in der wir unsere Intentionen erfüllen, konsistent ist und dass zur effektiven Umsetzung von Zielen auch eine Kohärenz hinsichtlich der Bereitstellung der dafür notwendigen Mittel erfolgen muss. So schreibt Baier:

„Human intentions are for particular actions, which fit in, more or less, with our larger plans, and which must be made to fit in with a context only part of which is provided by our own past and future intentional acts.“ [12, S.398]

Dieses Zitat verlangt eine Kohärenz zwischen den einzelnen beabsichtigten Handlungen eines Akteurs. Dieses Passend-Machen im Kontext größerer Pläne ist die Aufgabe der Intention. Warum dies notwendig ist, unterstreicht Baier dann im darauffolgenden zweiten Abschnitt.

„[E]very intention, [...], is an intention to act in the real world, a world in which other intentions are also realized, possibly realized by the very event which realizes this intention.“ [12, S.398]

Weiter unten fasst Baier zusammen:

„For intentions, we require recognition of the real world.“ [12, S.398]

⁹⁷Bratman formuliert, wie der Akteur glaubt, dass die Welt sei [27, S.52].

4. Überzeugungen, Wünsche und Absichten

Dieses Erkennen umfasst nach Baier nicht nur, dass wir erkennen, wie wir den intendierten Sachverhalt hervorbringen, sondern dass wir ebenfalls erkennen, dass es eine Möglichkeit gibt, dass dieses Hervorbringen des Sachverhaltes auch nicht wird stattgefunden haben können. D.h. wir erkennen gleichfalls, dass die Möglichkeit besteht, dass die intendierte Handlung nicht ausgeführt werden könnte. Dies liegt nach Baier daran, dass wir nur einen begrenzten Zugang zu allen Sachverhalten haben, die bei einer Handlung bestehen, die also das Ereignis der Handlung charakterisieren.

„[O]ne would have to put, into every self conscious intention, not merely recognition and acceptance of limited access [...], but a recognition of possible non-performance.“ [12, S.399]

Dies sind keine deskriptiven Aspekte, die das Vorliegen einer Intention beschreiben, sondern rationale Forderungen. Jemand sollte eine Intention nur dann haben, wenn er erkennt, dass das Intendierte auch möglich ist. Er sollte nichts intendieren, was notwendigerweise eintritt, oder er sollte seine Intention aufgeben. Dies spiegeln die Schemata (*PA*) bzw. $\alpha \text{ dint} : \varphi \supset \neg \Box \varphi$ wider.

Dies macht eine weitere Forderung der Konsistenz und Kohärenz von Überzeugungen, um die Konsistenz und Kohärenz von Intentionen fordern zu können, nicht erforderlich, zumal diese, wie gezeigt, nicht zielführend wäre. Es gibt eine kognitive Einstellung ‘Akzeptanz’, die sich wesentlich eher eignet als die Überzeugung, um den kognitiven Zugang eines Akteurs zu den Gehalten seiner Intentionen zu erklären. Cohen hatte festgestellt, dass an die Akzeptanz von Sachverhalten eher als an die Überzeugungen die Forderungen nach Konsistenz und Agglomeration zu stellen ist, als an Überzeugungen [45]. Cohen fügt hinzu:

„[T]he knowledge or certainty that an intender has about his own future action should be said to have acceptance, not belief, at its core. Belief may be present in addition but its presence is not conceptually necessitated.“ [46, S.128]

Er greift das Beispiel von Davidson auf. Es mag sein, dass man nicht glaubt, dass man tatsächlich zehn Kopien mit einem Durchschlag schafft. Aber wenn man beabsichtigt, es zu tun, hat man akzeptiert, dass man es tun wird. Es gibt keine Art von Verpflichtung im Gegensatz zur Akzeptanz, die eine Intention an eine Überzeugung koppelt. Wenn jemand tatsächlich beabsichtigt, dass er einen bestimmten Sachverhalt hervorbringt, ist dies nur rational, wenn er sich auch verpflichtet, diesen Sachverhalt hervorzubringen [46, S.128]. Diese Verpflichtung beinhaltet, dass er akzeptiert, dass er

diesen Sachverhalt hervorbringen wird. Ich hatte bereits festgehalten, dass ‘ich beabsichtige, dass φ , und ich akzeptiere nicht, dass ich φ hervorbringe’ in gewisser Weise paradox ist.

Demzufolge ist die Akzeptanz der kognitive Zugang zu den Gehalten der Intention. An diese Akzeptanz sind die Forderungen (*PA*) und (*MEC*) gerichtet, nur das als hervorzubringend zu akzeptieren, was auch historisch möglich ist. Bzw. soll das, was für die Intention notwendig zu intendieren ist, intendiert werden. Wenn es nicht möglich ist, dies als hervorzubringend zu akzeptieren, sollte die Intention aufgegeben werden, siehe (*MEC*).

$ \begin{aligned} (BMEC) \quad & (\alpha \text{ bel}:\Box(\varphi \supset \psi) \wedge \alpha \text{ bel}:\Box(\psi \supset \alpha \text{ dint}:\psi)) \\ & \supset (\alpha \text{ dint}:\varphi \supset \alpha \text{ dint}:\psi) \\ (MEC) \quad & (\Box(\varphi \supset \psi) \wedge \Box(\psi \supset \alpha \text{ dint}:\psi)) \\ & \supset (\alpha \text{ dint}:\varphi \supset \alpha \text{ dint}:\psi) \end{aligned} $
--

Die Mittel-Zweck-Forderung in Schema (*BMEC*) basiert auf den Überzeugungen des Akteurs und wurde von Bratman in [25, 27] aufgestellt. Ich habe argumentiert, dass es nicht sinnvoll ist, dass sich Rationalitätsforderungen für Intentionen auf Überzeugungen stützen sollten, sondern auf die Akzeptanz, dass bestimmte Sachverhalte tatsächlich bestehen. Dann ist es eine logische Folgerung, dass das von Bratman mit Überzeugungen formulierte Prinzip auf ein Prinzip reduziert wird, dass die Kohärenz von Intentionen tatsächlich sichern kann. Daher wird (*MEC*) als gültiges Schema in *bdi-stit* Logik verankert.

Dass dieses Prinzip (*MEC*) selbst rational sinnvoll ist, ist selbsterklärend. Wenn meine Intentionen als partielle Pläne verstanden werden, die gefasst werden, um bestimmte Ziele zu erreichen, dann sollte ich, wenn es notwendig ist, diese Pläne mit weiteren Intentionen vervollständigen. Ich intendiere, nach Paris zu fahren. Wenn ich mir dafür ein Ticket kaufen muss, und es für den Kauf des Tickets notwendig ist, den Kauf zu intendieren, dann sollte ich diesen Kauf intendieren. Dass der Antezedens nicht nur aus der Forderung, dass $\Box(\varphi \supset \psi)$ bestehen kann, sondern auch aus $\Box(\psi \supset \alpha \text{ dint}:\psi)$ ist leicht einzusehen. Ein notwendiger Sachverhalt für meine Reise nach Paris ist, wenn ich bereits mit gekauftem Ticket im Flugzeug sitze, dass dieses Flugzeug auch abhebt. Nun würde niemand von mir verlangen, dass ich intendiere, diese Maschine dazu zu bringen, dass sie abhebt, solange ich nicht der Pilot. Um Intentionen kohärent zu fassen, ist es notwendig, dass die Mittel, die dafür notwendig sind, auch intendiert werden. Es ist jedoch aus rationalen Gründen nicht erforderlich, jedes notwendige Mittel zu intendieren. Die Forderung beinhaltet zu erkennen, wel-

4. Überzeugungen, Wünsche und Absichten

ches notwendige Mittel intendiert werden muss und welches nicht. Beides geht aus (*MEC*) hervor.

Dieses Mittel-Zweck-Prinzip bildet den Abschluss der Diskussion, welche Interaktionsschemata welcher Operator erfüllen sollte. Ich habe in diesem vierten Kapitel motiviert, in welcher Semantik ein Überzeugungs-, Wunsch- und Intentionoperator interpretiert werden sollte und welchen Axiomen bzw. Theoremen der jeweilige Operator genügen sollte. Bevor im nächsten Kapitel jedoch die Umsetzung dieser Operatoren in einem logischen System vorgenommen wird, werde ich einen kurzen Überblick auf bereits in der Literatur vorhandene logische Systeme geben, die sich ebenfalls mit der Beschreibung von Überzeugungen, Wünschen und Intentionen eines Akteurs beschäftigen.

5 Von Stit und *BDI* zu *bdi-stit*

Der bisherige Teil der Arbeit hat sich in den vier vorangegangenen Kapiteln der Aufgabe gewidmet, die Begriffe zu klären, die einer Logik zugrunde liegen, die Handlungen, Überzeugungen, aber auch Intentionen und Wünsche beschreibt. In Kapitel 2 und 3 ist motiviert worden, warum eine Stit-Logik auch dann eine geeignete Form der Beschreibung von Handlungen ist, wenn man in Betracht zieht, dass Handlungen ein Ereignis unterliegt. Ebenso ist in Kapitel 2 und 3 auf die Unterschiede und Gemeinsamkeiten von Ereignissen und Zuständen eingegangen worden. Gemein ist ihnen, dass sie durch Sachverhalte beschrieben werden und dass es sowohl Typen von Ereignissen als auch Typen von Zuständen gibt, die durch Mengen von Sachverhalten charakterisiert werden. Der gravierende Unterschied zwischen einem Zustand und einem Ereignis ist das jeweilige Verhältnis zur Zeit.

Im letzten Teil von Kapitel 3 wurden mentale Zustände und mentale Ereignisse betrachtet. Dabei stand die Frage im Vordergrund, inwieweit mentale Ereignisse bzw. mentale Aktivitäten als mentale Akte eines Subjektes zählen. Solange der Akteur Kontrolle über eine solche Aktivität in dem Sinne hat, dass er entscheiden kann, einen bestimmten mentalen Zustand mit konkretem Gehalt, herbeizuführen oder aufzugeben, zähle ich diese Aktivität zu den mentalen Akten. Ein mentaler Zustand ist unter der Kontrolle eines Akteurs, wenn er ein mentales Ereignis herbeiführen kann, dass den Zustand bildet oder zum Fortbestehen bzw. zum Aufgeben des Zustandes führt.

In Kapitel 4 wurden die propositionalen Einstellungen, Überzeugung, Wunsch und Intention,¹ die als mentale Zustände im Akteur vorliegen, genauer untersucht. Vor allem fand dabei die Kontrolle Berücksichtigung, die der Akteur über den jeweiligen Zustand hat. Dies hatte weitreichende Folgen bei der Beantwortung der Fragen, welche mentalen Zustände rational für einen Akteur einzunehmen ist und in welchem Verhältnis die mentalen Zustände zueinander stehen. Aus diesen Überlegungen resultieren Eigenschaften und Interaktionen, die bei der Beschreibung dieser mentalen

¹Die aussagenlogische Formel hinter dem jeweiligen Operator *bel*, *des*, oder *int* korrespondiert mit einem Sachverhalt. Die Situation (Moment-Geschichtsverlauf), in der die Formel bewertet wird, schafft somit einen nicht-transienten Sachverhalt dem eine Proposition korrespondiert.

5. Von Stit und BDI zu *bdi-stit*

Zustände im Zusammenspiel mit Handlungsoperatoren wieder zu finden sein sollten.

Der Schwerpunkt des zweiten Teiles der Arbeit liegt nun in der Beschreibung dieser Zustände und Handlungen in einem gemeinsamen logischen System. In den folgenden zwei Kapiteln wird dafür die *bdi-stit* Logik entwickelt und dargestellt. Handlungen von Akteuren werden mittels deliberativen Stit-Operatoren beschrieben. Diese deliberative Stit-Logik wird um Operatoren für Überzeugungen, Wünsche und Intentionen, wie sie in BDI-Logiken diskutiert werden, sinnvoll ergänzt. Hierfür werde ich die BDI-Logik *BDICTL** vorstellen und die Unterschiede bei der Handhabung der Operatoren für mentale Zustände in der *BDICTL** und in der *bdi-stit* Logik herausstellen. Diese unterschiedliche Handhabung wurde in Kapitel 4 für die einzelnen propositionalen Einstellungen motiviert.

In Abschnitt 5.1 grenze ich meine Betrachtungen von einigen anderen Vorschlägen zur Handhabung von mentalen Zustandsoperatoren ab. Ich beschränke mich auf eine Auswahl von in der Literatur diskutierten logischen Systemen, die sich mit mentalen Operatoren für Wünsche, Intentionen und Überzeugungen beschäftigen. Eine der meist diskutierten dieser Logiken ist sicherlich *BDICTL**, die in den 90ziger Jahren von Rao und Georgeff entwickelt wurde [64, 65, 66, 114]. Im darauffolgenden Abschnitt 5.2 wird die *bdi-stit* Logik eingeführt und Unterschiede und Gemeinsamkeiten zu *BDICTL** benannt. In Abschnitt 5.3 zeige ich, dass die *bdi-stit* Logik eine Generalisierung von deliberativer Stit-Logik und einer *BDICTL** ohne temporale Operatoren ist. Die Berücksichtigung der temporalen Operatoren erfolgt in Abschnitt 6.3.

5.1 Von BDI zu *bdi-stit*

Wenn man eine Logik konstruiert, die sich mit der Modellierung von mehreren Subjekten und deren mentalen Zuständen wie Überzeugungen, Wünsche und Intentionen beschäftigt, so ist ein grundlegender Ansatz in den *BDI* Logiken zu finden, wie sie von Rao und Georgeff in den neunziger Jahren entwickelt wurden [64, 65, 66]. Diese Logiken basieren auf der in Kapitel 1 bereits erwähnten *CTL** Logik, einer verzweigenden Zeitlogik, auf der die Modaloperatoren, die Überzeugungen (*Beliefs*), Wünsche (*Desires*) und Absichten (*Intentions*) beschreiben, durch eine relationale Semantik interpretiert werden. Für diese Zustandsoperatoren werden verschiedene Axiomenschemata angenommen, die unterschiedliche Beziehungen der mentalen Zustände untereinander wiedergeben. Rao und Georgeff berufen sich bei der Motivation der Interaktionsaxiome auf Bratman [24, 25]. Aufgrund der relationalen Semantik und dem Ausschluss nicht-leerer Zugänglichkeitsrelationen sind Überzeugungen, Wünsche und

Intentionen in einem Akteur als nicht-konfigurierend festgelegt. In Abschnitt 4.1 wurde begründet, dass dies für einen rationalen Akteur möglich ist. Es kann durchaus rational sein, antagonistische Wünsche und konfligierende Überzeugungen zu haben und für das Haben dieser Einstellung rational zu sein, obwohl es möglicherweise nicht rational ist, paradoxe Wünsche und inkonsistente Überzeugungen zu haben. Eine relationale Semantik wie in den *BDI* Logiken schließt sich dann für die Operatoren aus. Bei der Modellierung von Wünschen und Überzeugungen in *bdi-stit* Logik wird eine Form der *Nachbarschaftssemantik* statt einer relationalen Semantik verwendet.²

Aber nicht nur Rao und Georgeff haben sich mit Logiken beschäftigt, die mentale Zustände modellieren. Es gibt in der Literatur immer wieder Vorschläge für Logiken mit Überzeugungs-, Wunsch- und Intentionoperatoren. Dabei sind bei der Modellierung von mentalen Zuständen häufig Alternativen, u.a. Nachbarschaftssemantiken, zu einer relationalen Semantik vorgeschlagen worden. Beispielsweise entstand in der Diskussion um das Problem der logischen Allwissenheit der Vorschlag, Wissens- und Überzeugungsoperatoren in einer Nachbarschaftssemantik zu interpretieren. Stalnaker schlug dies für den Wissensoperator vor [134]. In [55] liefern Fagin und Halpern eine Übersicht von verschiedenen Definitionen von expliziten und impliziten Überzeugungen zur Vermeidung der logischen Allwissenheit, vgl. für die Diskussion Abschnitt 4.1 S.152. In [55, Abschnitt 6] betrachten sie eine „Logic of local reasoning“. Sie enthält einen klassischen, monotonen Operator für die Modellierung von expliziten Überzeugungen und einen anderen Operator für implizite Überzeugungen. Die erste Definition einer solchen „Logic of local reasoning“ schreiben sie Levesque zu.³ Während es in einer „Logic of local reasoning“ möglich ist, explizit („local“) konfligierende Überzeugungen zu haben, ist dies implizit (global über alle Nachbarschaften) nicht möglich. Unmöglich ist es für eine explizite und erst recht für eine implizite Überzeugung, einen inkonsistenten Gehalt zu haben. Dass dies sinnvoll ist, habe ich in Abschnitt 4.1 unterstrichen.

Es ist in einer Nachbarschaftssemantik möglich, Zustände mit konfligierenden Gehalten auszuschließen. In [139] schlagen Su et al. eine besondere Form der Nachbarschaftssemantik nicht für Überzeugungen, sondern für Proeinstellungen vor. Im Gegensatz von den in [55, 56, Kapitel 9] vorgestellten monotonen Nachbarschaften

²Diese Form der Definition von Wahrheitsbedingungen in einem Modell gehen zurück auf D. Scott und R. Montague. Daher werden solche Modelle auch Scott-Montague Modelle genannt [100, 121].

³In [55] wird erwähnt, dass diese Logik von expliziten und impliziten Überzeugungen auf Levesque [92] zurückgeht. Aber nicht nur Levesque schlug Nachbarschaftssemantiken für die Modellierung von Überzeugungen vor. Auch in [146] wird von Vardi eine Nachbarschaftssemantik verwendet.

5. Von *Stit* und *BDI* zu *bdi-stit*

beschreiben diese von ihnen als linear bezeichneten Nachbarschaftssysteme die Proeinstellung eines Akteurs als nicht-konfligierend.⁴ Eine lineare Nachbarschaftssemantik sorgt dafür, dass weder Tautologien noch Kontradiktionen gewünscht werden. Ein Subjekt kann in linearen Nachbarschaftssystemen keine konfligierenden Wünsche haben. Als Motivation für eine lineare Nachbarschaft postulieren Su et al. ein sogenanntes „Side-Effect-Free“ Prinzip ($\alpha des:\varphi$ und $\alpha bel:(\varphi \supset \psi)$ und nicht $\alpha des:\psi$). Dies führt zu der meines Erachtens nicht gut motivierten Folge, dass, wenn der Akteur wünscht, dass φ und dass ψ , es nicht notwendig ist, dass er wünscht, dass φ oder ψ . Da zudem konfligierende Proeinstellungen ausgeschlossen sind, scheint mir dies keine gute Alternative zu einer monotonen Nachbarschaftssemantik für Wunschoperatoren zu sein.

Mit der Vermeidung der Modellierung von Nebeneffekten bei Intentionen beschäftigen sich auch Cohen und Levesque in [47]. Sie berufen sich auf Bratman und setzen sich sieben Eigenschaften zur Aufgabe, die ein Intentionsoperator im Wechselspiel mit den anderen Operatoren beachten soll, vgl. [47, S.217f]:

- (1) Intentionen regen den Akteur an, für den intendierten Sachverhalt (Ereignis bei Cohen und Levesque) zu sorgen.
- (2) Rationale Intentionen sind nur möglich zu fassen, wenn sie zulässig sind.
- (3) Ein Akteur kann etwas nur intendieren, wenn es nicht durch andere Intentionen ausgeschlossen wird. Akteure suchen die Umsetzung ihrer Intentionen. Sie verpflichten sich dem intendierten Sachverhalt.
- (4) Dafür ist es notwendig nach Cohen und Levesque, dass sie glauben, dass dieser möglich ist.⁵
- (5) Akteure glauben nicht, dass sie den intendierten Sachverhalt nicht hervorbringen. Dies ist in Analogie zu der asymmetrischen These Bratmans aus [25, S.38].
- (6) Ein Akteur glaubt unter bestimmten Bedingungen, dass er den von ihm intendierten Sachverhalt hervorbringt.

⁴Auch wenn die Operatoren für Proeinstellungen in einer Nachbarschaftssemantik interpretiert werden, werden die Überzeugungsoperatoren durch relationale Semantiken interpretiert [139].

⁵Dieser 4. Punkt in [47, S.218] ist Bratman nicht zuschreibbar. Er spricht sich in [27] gegen einen solchen involvierten Glauben in einer Intention aus.

- (7) Der letzte der von Cohen und Levesques genannten Punkte ist, dass es nicht der Fall sein muss, dass Akteure alle erwarteten Nebeneffekte ihrer intendierten Handlungen selbst intendieren.

Im letzten Kapitel hatte ich erläutert, warum (1-3) und (7) sinnvoll und nötig sind. Punkt (1) ist in Übereinstimmung mit der Motivation des Axioms (IA) S.192 zu begreifen. Punkt (2) spiegelt sich in Axiom (PA) S.200 wider. Punkt (3) ist erfüllt, wenn Akteure keine konfligierende Intentionen haben, sich die Intentionen nicht gegenseitig ausschließen und Intentionen agglomerieren. Dies spiegelt sich in den Axiomen (D_i) und (A_i) wider S.174f. Punkt (7) beschreibt eine Motivation, warum die Formeln (M_i) bzw. (RM_i) keine gültigen Rationalitätsforderungen wiedergeben, vgl. die Diskussion von (M_i) und (RM_i) S.180). Von den Punkten (4) und (5) hatte ich mich distanziert, siehe die Diskussion von (IB), ($IB\Diamond$) und ($I\rightarrow B\Diamond$) S.208ff. In Abschnitt 6.4 werde ich daher eine modale propositionale Logik vorstellen, die (1-3) und (7) für einen spezifizierten Intentionsoperator postuliert und die Punkte (4) und (5) nicht als notwendige Gesetzmäßigkeiten festlegt. Da Cohen und Levesque die Bedingungen offenlassen, unter denen er glaubt, dass er den intendierten Sachverhalt hervorbringt, würde ich Punkt (6) nicht ablehnen. Es ist immer möglich, dass ein Akteur glaubt, dass er den intendierten Sachverhalt hervorbringt. Aber es ist nicht notwendig für die Rationalität des mentalen Zustandes der Intention, tatsächlich zu glauben, den intendierten Sachverhalt hervorzubringen.

Zur Konstruktion ihrer Logik erweitern Cohen und Levesque eine dynamische Logik um Modaloperatoren für Überzeugungen *bel*, Operatoren für Ausdrücke der Form ‘welches Ereignis ist grad passiert’ (*done*) bzw. ‘passiert als nächstes’ (*happens*) und einen Modaloperator *goal* für Ziele. Ihre Logik enthält Variablen sowohl für Handlungen als auch für Akteure, so dass es möglich ist, einem Akteur eine Handlung zuzuschreiben. Eine Intention wird dabei als ein nicht-primitiver Term verstanden, der über die genannten Operatoren und Prädikate für Handlungen definiert ist. Eine Intention wird mittels Überzeugungs- und Zieloperatoren eines Akteurs dargestellt, die als Gehalte haben, was er glaubt, was gerade passiert ist, was er als nächstes tun möchte, und was er glaubt, was dieser Handlung folgt. Dies widerspräche Bratmans Auffassung von Intention als nicht zurückführbar auf andere mentale Zustände.

Ein interessanter Punkt in ihrer Ausführung ist die Unterscheidung zweier Arten von Intentionen. Entweder jemand beabsichtigt eine bestimmte Handlung, [47, S. 245], oder jemand intendiert einen Sachverhalt p , wobei diese Absicht ausdrückt, dass der Akteur intendiert, den Sachverhalt p zu bewirken [47, S. 248]. Im ersteren Fall wird der Intentionsoperator auf ein Akteur-Handlung-Paar angewandt, im letzteren

5. Von *Stit* und *BDI* zu *bdi-stit*

Fall auf einen Akteur und eine Formel.

Beide Definitionen implizieren anhand der anderen Operatoren nach Cohen und Levesque eine Verpflichtung, eine bestimmte Handlung hervorzubringen. Eine Verpflichtung ist hierbei ein Ziel, von dem er glaubt, dass es im Moment nicht der Fall ist, und es erst aufgibt, wenn er glaubt, dass es erfüllt ist bzw. dass es unmöglich ist. Im Fall der Intention einer Handlung ist es die Verpflichtung, die durch die Handlungsvariable beschriebene Handlung im kommenden Moment auszuführen und die Überzeugung zu haben, dass dies passiert. Im zweiten Fall des intendierten Sachverhaltes ist es die Verpflichtung, dass der Akteur eine bestimmte Folge von Ereignissen herbeiführt, von denen er glaubt, dass nach dessen Abfolge der angestrebte Sachverhalt vorliegt. Er ist zudem verpflichtet, dass der Sachverhalt durch die Ereignisse nicht herbeigeführt worden wäre, wenn es nicht sein Ziel gewesen wäre.

Als Probleme für den Ansatz von Cohen und Levesque kann man anführen, dass gerade Bratman, auf den sie sich berufen, versucht hat zu zeigen, dass Intentionen eben nicht auf Überzeugungen und Wünsche des Akteurs zurückzuführen sind. Die Verwendung des Begriffes „Goal“ anstelle von Wunsch hilft dabei wenig, zumal ein „Persistence-Goal“ (Verpflichtung) eine Kombination von Ziel und Überzeugungen darstellt, welche eher durch ein Belief-Desire Modell von Handlungsmotiven, vgl. [48, Essay 1], als durch Bratmans Begriff einer Intention motiviert ist, vgl. [25]. Bratmans Intensionsbegriff, dem inhärent ist, dass die Umsetzung von Plänen ein Grund ist, warum sich Intentionen nicht auf Überzeugungen und Wünsche zurückführen lassen, wird erwähnt, aber nicht weiter berücksichtigt. Sie gehen auf viele Vorteile ihrer Definition von Intentionen ein [47, S.251], u.a. folgt aus ihren Definitionen nicht, dass jeder Akteur alle Nebeneffekte seiner Absichten beabsichtigen muss. Jedoch glaubt ein Akteur im Falle einer Absicht bei Cohen und Levesque, dass es möglich ist, die beabsichtigte Handlung zu vollziehen, und er glaubt auch, dass er es tatsächlich im nächsten Moment tun wird. Diesen beiden kognitivistischen Implikationen steht Bratman kritisch gegenüber, vgl. [27]. Bratman fordert von rationalen Akteuren lediglich, dass die Gehalte von Intentionen und Überzeugungen eines rationalen Akteurs konsistent und kohärent sind. Dies impliziert, dass ein Akteur nichts beabsichtigt, von dem er glaubt, er werde es nicht tun (asymmetrische These). Dies impliziert nicht, dass er, wenn er etwas beabsichtigt, auch glaubt, es tun zu werden.⁶

⁶Eine weitere propositionale, dynamische Logik ist das KARO Framework (Knowledge, Abilities, Results, Opportunities). Hierbei wurde eine dynamische Logik für Handlungen und Ereignisse um modale Operatoren ergänzt. Diese Logik ist expressiv, aber hinsichtlich Vollständigkeit und Entscheidbarkeit schwer zu handhaben, vgl. [166]. Ein Einblick in das KARO Framework findet sich in [144] und für

Ein modaler Ansatz zur Beschreibung von Intentionen mit Handlungsvariablen wurde von Longin und Herzig vorgeschlagen. In [78] diskutieren Longin und Herzig Interaktionsaxiome zwischen Intentions- und Überzeugungsoperatoren. Hierbei verwenden sie ‘Intention’ als einen primitiven Term, der sich nicht aus anderen Operatoren zusammensetzt. Beide Operatoren werden in einer monotonen Nachbarschaftssemantik definiert. Da dies ermöglicht, dass ein rationaler Akteur konfligierende Absichten hat, scheint mir dies für eine Logik, die ideelle und demzufolge rationale Akteure betrachtet, als zu wenig restriktiv. Obwohl sie Ereignisoperatoren einer dynamischen Logik und Handlungsvariablen in die Logik aufnehmen, betrachten sie keine Interaktionen zwischen den Handlungen und den Operatoren für mentale Zustände wie z.B. Intentionen.

In [97] betrachten Herzig und Lorini im \mathcal{LIA} System, „Logic of Intention and Attempt“, Intentionen und Versuche, etwas zu tun, wobei sie normale Modaloperatoren für Ziele („goal“), für Überzeugungen und für Versuche diskutieren, eine Basishandlung auszuführen. Sie unterscheiden zwischen zukunftsgerichteten und gegenwärtigen Intentionen⁷ und definieren letztere, indem sie Intentionen für Basishandlungen auf das Ziel zurückführen, etwas zu tun. Ein Versuch, eine Basishandlung zu vollziehen, befindet sich auf dem sogenannten „executive level“ [97, p.65]. Dementgegen setzen sie Intentionen, die als „mental counterparts“ [97, p.50] bezeichnet werden.

Jemand beabsichtigt gegenwärtig eine basale Handlung genau dann, wenn er das Ziel hat, diese zu versuchen, wobei das Ziel zu haben bewirkt, dass er sie effektiv versucht, vgl. Prinzip 1 und 2 [97, S.50]. Lorini und Herzig beschränken sich auf Basishandlungen, d.h. eine Handlung, „which is not performed by way of another action“ [97, S.48], da „[o]nly basic action types of an agent (i.e. bodily movements in an agent’s repertoire) can be under the agent’s voluntary control, that is, only basic action types of an agent can be the object of his volitions“ [97, S.49]. Dem stimme ich nicht zu, wie aus den vorangegangenen Kapiteln hervorgeht. Es ist durchaus möglich, dass ein Akteur direkte voluntative Kontrolle auf Umstände außerhalb seines Körpers hat, z.B. indem er Pläne über einen längeren Zeitraum schmiedet und diese dann umsetzt. Demzufolge können nicht nur Basalhandlungen, sondern auch solche Pläne Objekte seiner Volitionen bzw. Intentionen sein.

den Kern des KARO Frameworks findet sich ein Vollständigkeitsresultat in [145].

⁷Die Unterscheidung geht zurück auf Anscombe [7] und Searle [123] und wurde bereits im letzten Abschnitt S.185 diskutiert. Die Bezeichnungen „future-directed“ und „present-directed“ stammen von Bratman [25]. Searle sprach von „distal“ und „proximal intentions“ bzw. vorausgehender Absicht und Handlungsabsicht.

5. Von Stit und BDI zu bdi-stit

Herzig und Lorini nehmen mehrere Interaktionsaxiome für Versuch, Überzeugung und Ziel eines Akteurs an. Überzeugungen und Ziele eines Akteurs werden jeweils durch eine relationale Semantik beschrieben. Der Überzeugungsoperator wird durch die Schemata *KD45* und der Zieloperator durch *KD* axiomatisiert [97, S.57]. Als Interaktionen zwischen diesen Operatoren werden genannt, dass ein Akteur von jedem Ziel glaubt, dass es ein Ziel ist, bzw. glaubt, dass es keines ist, wenn es keines ist. Zudem ist etwas nur ein Ziel, wenn man nicht glaubt, dass es nicht der Fall ist. Es ist somit nicht möglich, etwas als Ziel zu haben, von dessen Bestehen man bereits überzeugt ist. Verwendet man Wunsch statt Ziel hatte ich in Abschnitt 4.1 gezeigt, dass es durchaus möglich ist, sich etwas zu wünschen, auch wenn man glaubt, dass der Wunsch bereits erfüllt ist. Jemand kann sich wünschen, gesund zu sein, auch wenn er es ist. Jemand kann sich wünschen, den Herd ausgeschaltet zu haben, auch wenn er glaubt, ihn ausgeschaltet zu haben. Durch die Verwendung von Ziel anstelle von Wunsch scheint dies zwar plausibler, aber nicht bindend. Ein Akteur kann sich als Ziel etwas stecken, von dem er glaubt, es in demselben Moment zu erreichen. Um die Idee hinter dieser Interaktion zu beschreiben, benötigt man zumindest temporale Operatoren. Im Falle von instantanen Operatoren sind diese Interaktionen jedoch sowohl für Wünsche als auch für Ziele unplausibel.

Interessant ist die Definition des Versuchsoperators, der beschreibt, dass ein Akteur eine bestimmte Körperbewegung versucht auszuführen, so dass ein bestimmter Sachverhalt besteht. Für jede Körperbewegung und jeden Akteur wird eine Relation eingeführt, so dass die Menge an eingeführten Operatoren und Relationen auf den ersten Blick sehr umfangreich erscheint. Aus den Interaktionsaxiomen ergibt sich jedoch, dass für beliebige Akteure i, j und Handlungsversuche α, β diese aufgrund der Determiniertheit in einem linearen Ablauf zusammenfallen.⁸ Wenn ein Akteur i in Welt w versucht, die Basishandlung α auszuführen, so dass φ , dann ist nur eine mögliche Welt w' zu betrachten. Jeder Versuch vom Akteur j in Welt w, β zu tun, führt zu demselben Zustand, vgl. Axiom 6a und Erklärung [97, S.57f]. Da i und j nicht verschieden sein müssen, ist es quasi unerheblich, welche Basishandlung α der Akteur i versucht auszuführen. Sie haben denselben Zustand zur Folge. Welche Basishandlung versucht wird, spielt für den nachher bestehenden Sachverhalt keine Rolle. Die Motivation für verschiedene Variablen für Basishandlungen bleibt unklar. Eine Stit-Formel, die ausdrückt, der Akteur sorgt dafür, dass φ , ohne die Erwähnung einer Basishandlung durch eine Handlungsvariable, liefert somit nicht weniger Informationen.

⁸Lorini und Herzig bezeichnen Akteursvariablen mit i, j und Handlungsvariablen mit α, β .

Obwohl Lorini und Herzig viele interessante Interaktionen von Überzeugung, Ziel und Versuch einer Handlung diskutieren, sind sie meines Erachtens zu restriktiv für Überzeugungen. Warum die Axiome *D45* für den Überzeugungsoperator abgelehnt werden sollte, habe ich bereits im Abschnitt 4.1 argumentiert. Sie vertreten jedoch Interaktionsaxiome von Versuch und Ziel, die ich in ähnlicher Weise für Handlungen fordere. Jeder Versuch einer Basishandlung wird intendiert, siehe Prinzip 2 [97, S.50]. Ein Akteur versucht eine Handlung nur, wenn er das Ziel hat, die Handlung zu versuchen, vgl. „discussions“ [97, S.63]. Dies ist analog motiviert wie das von mir vorgeschlagene Interaktionsaxiom (*DI*), dass eine Intention, eine Handlung zu vollziehen, nur besteht, wenn man den Wunsch nach dem Bestehen des Sachverhaltes hat, der durch das Ausführen der intendierten Handlung hervorgebracht wird.

Die Kritik, die Lorini und Herzig an „bringing about“ und „seeing to it that“-Logiken vorbringen,⁹ ist berechtigt. Allerdings wird das Problem ihrerseits nicht behoben. Sie kritisieren, dass es in solchen Logiken nicht möglich ist, zwischen basalen und nicht-basalen Handlungen zu unterscheiden, da sie eine Handlung daran bewerten, welcher Sachverhalt nach der Handlung besteht [97, S.55]. Dass dies kein Kritikpunkt sein kann, sondern sinnvoll ist, habe ich in Kapitel 3 gezeigt. Bei der Beschreibung der Trennung von Basishandlungen und des durch die Handlung hervorgebrachten Sachverhaltes sind Herzig und Lorini ihrerseits an Grenzen gestoßen. Da die Wahrheitsbedingung für einen Versuchsoperator, eine bestimmte Basishandlung durchzuführen, durch die Sachverhalte definiert ist, die nach dem Durchführen der Handlung in einem determinierten nächsten Zustand vorliegen, ist das Einführen einer Handlungsvariable ein scheinbarer Beweis für die Unterscheidung von basalen und nicht-basalen Handlungen. Dies bleibt jedoch nur ein scheinbarer. Unabhängig welche Basishandlung (bzw. welche möglicherweise Nicht-Basishandlung) durch eine Handlungsvariable ausgedrückt wird, wird bei Lorini und Herzig die Bewertung, welche Handlung stattgefunden hat, nur an dem durch die Handlung erreichten Folgezustand gemessen wurde. Dies ist ein Grund, warum ich auf Handlungsvariablen für verschiedene Körperbewegungen, was der Akteur im Einzelnen getan hat, verzichte und eine Handlung dadurch beschreibe, für welchen Sachverhalt der Akteur gesorgt hat. Von besonderem Interesse ist dann, welche Einstellungen ein Akteur im Falle einer Handlung zu dem Sachverhalt hat, der durch die Handlung hervorgebracht wird. Daher ist es sinnvoll nicht nur Handlungsoperatoren, sondern auch Intentionaloperatoren zu betrachten.

⁹Eine Übersicht der Ansätze zu „bringing about“ und „see to it that“ Logiken bis 1991 ist in [125] zu finden.

5. Von Stit und BDI zu bdi-stit

Dafür schlage ich eine *bdi-stit* Logik vor, Erweiterung der deliberativen Stit-Logik um die sogenannten BDI-Operatoren. Um Unterschiede zwischen *BDICTL** und *bdi-stit* aufzeigen zu können, wird die *BDICTL** kurz vorgestellt, bevor die Syntax und Semantik von *bdi-stit* eingeführt wird. Dabei vernachlässige ich vorerst die temporalen Operatoren in *BDICTL**, die in der darunter liegenden *CTL** Logik vorkommen. Für den Moment stehen die mentalen Zustandsoperatoren im Vordergrund. Im folgenden Abschnitt 5.2 wird die allgemeine *bdi-stit* Logik definiert, wobei ich auf Interaktionen zwischen den Operatoren verzichte. In Kapitel 6 werde ich sowohl auf die Erweiterungen um Interaktionsaxiome als auch auf die Vollständigkeit und die Frage nach der Entscheidbarkeit der allgemeinen *bdi-stit* Logik eingehen. In Abschnitt 6.3 finden dann die temporalen Operatoren ebenfalls Berücksichtigung.

Eine Logik für Überzeugungen, Wünsche und Intentionen - *BDICTL**

Um zu verstehen, warum ich im Folgenden von der *bdi-stit* Logik als einer Generalisierung der *BDICTL** spreche, vgl. Abschnitt 5.3, ist es nötig, sich mit dieser Logik kurz zu befassen. Die Sprache der *BDICTL** ist wie folgt aufgebaut, vgl. [66].

Definition 5 Die Formeln für *BDICTL** sind induktiv aufgebaut gemäß:

- (S1) Jeder Aussagebuchstabe p_1, p_2, \dots ist eine Zustandsformel.
- (S2) Falls φ, ψ Zustandsformeln sind, dann sind $\neg\varphi, (\varphi \wedge \psi)$ Zustandsformeln.
- (S3) Falls φ eine Pfadformel ist, dann sind $A\varphi$ und $E\varphi$ Zustandsformeln.
- (S4) Falls φ eine Zustandsformel und α eine Akteursvariable ist, dann sind $\alpha \text{ bel} : \varphi, \alpha \text{ des} : \varphi$ und $\alpha \text{ int} : \varphi$ Zustandsformeln.
- (P1) Jede Zustandsformel ist auch eine Pfadformel.
- (P2) Falls φ, ψ Pfadformeln sind, dann sind $\neg\varphi, (\varphi \wedge \psi)$ Pfadformeln.
- (P3) Falls φ, ψ Pfadformeln sind, dann sind $X\varphi, (\varphi U \psi)$ ebenfalls Pfadformeln.

Wooldridge [144, 166], Rao & Georgeff [64, 66] und andere Autoren verwenden anstelle der hier genannten Notationen die Notationen $[Bel \alpha \varphi]$ oder $Bel(\varphi)$. In Analogie zu den Stit-Operatoren gebrauche ich für die Zustandsoperatoren die Infix-Notation.

Falls lediglich *BDICTL* statt *BDICTL** betrachtet werden soll, d.h. die darunter liegende Zeitlogik *CTL* statt *CTL** ist,¹⁰ werden (P1) – (P3) durch (P0) ersetzt:

¹⁰Zur Unterscheidung der beiden Logiken siehe Abschnitt 1.1 S.12.

(P0) Falls φ, ψ Zustandsformeln sind, dann sind $X\varphi$ und $\varphi U\psi$ Pfadformeln.

Der Sinn der Unterscheidung in Pfad- und Zustandsformeln zeigt sich, wenn man die Modelle betrachtet. Dafür sind folgende Definitionen vonnöten, vgl. [66, 302f].

Definition 6 Eine Kripke-Struktur ist $\mathcal{M}_e = (W, \{S_w\}_{w \in W}, \{R_w\}_{w \in W}, \mathcal{A}, B, D, I, v)$, wobei W eine Menge von Welten, S_w eine Menge von Zuständen für jedes $w \in W$, $R_w \subseteq S_w \times S_w$ eine Zugänglichkeitsrelation über den Zuständen in w ist und $B^\alpha, D^\alpha, I^\alpha \subseteq W \times S \times W$ mit $S = \{s \in S_w \mid w \in W\}$ Zugänglichkeitsrelationen für die mentalen Operatoren für jeden Akteur $\alpha \in \mathcal{A}$ sind. Die Bewertungsfunktion v bildet die propositionalen Variablen in die Potenzmenge der Welt-Zustands-Paare ab.

Die Menge S ist die Menge aller Zustände in dem Modell. Zu jeder Welt w wird die Menge der Pfade in w mit P_w angegeben. Ein Pfad $(w, s), (w, s_1) \dots \in P_w$ ist hierbei eine Folge von Zuständen aus S_w , der in (w, s) mit dem Zustand $s \in S_w$ beginnt. Ein Pfad heißt maximal, wenn jeder Zustand des Pfades einen Folgezustand hat. Insbesondere sind maximale Pfade unendlich. Die Menge R_w ist die Zugänglichkeitsrelation der Zustände in einer Welt und bestimmt die Konstruktion der Pfade. Für zwei aufeinander folgende Zustände $(w, s_i), (w, s_{i+1})$ in einem Pfad gilt, dass $(s_i, s_{i+1}) \in R_w$. Eine Welt $w \in W$ unterscheidet von einem Geschichtsverlauf, dass die Menge der Zustände in der Welt nicht linear geordnet ist. Die Relationen R_w sind total, so dass jeder Zustand mindestens einen Folgezustand in der Welt w hat. Für die Relationen B^α, D^α und I^α können Zustände fixiert werden, so dass diese Relationen binäre Zugänglichkeitsrelationen über Welten definieren: $B_s^\alpha \subseteq W \times W$ und $(w, w') \in B_s^\alpha$ gdw. $(w, s, w') \in B^\alpha$.

Es wird zwischen der Erfüllbarkeit von Zustandsformeln und von Pfadformeln unterschieden [66, S.303]. Zustandsformeln werden in Paaren von (w, s) erfüllt. Ihre Erfüllbarkeit ist wie folgt definiert:

- $\mathcal{M}, (w, s) \models \varphi$ gdw. $(w, s) \in v(\varphi)$ für atomare Formeln.
- $\mathcal{M}, (w, s) \models \neg\varphi$ gdw. $\mathcal{M}, (w, s) \not\models \varphi$.
- $\mathcal{M}, (w, s) \models (\varphi \wedge \psi)$ gdw. $\mathcal{M}, (w, s) \models \varphi$ und $\mathcal{M}, (w, s) \models \psi$.
- $\mathcal{M}, (w, s) \models A\varphi$ gdw. für alle maximalen Pfade $(w, s), (w, s_1) \dots \in P_w$ gilt:
 $\mathcal{M}, (w, s), (w, s_1), \dots \models \varphi$.
- $\mathcal{M}, (w, s) \models \alpha bel: \varphi$ gdw. für alle $w' \in B_s^\alpha(w)$ gilt, dass $\mathcal{M}, (w', s) \models \varphi$.
- $\mathcal{M}, (w, s) \models \alpha des: \varphi$ gdw. für alle $w' \in D_s^\alpha(w)$ gilt, dass $\mathcal{M}, (w', s) \models \varphi$.
- $\mathcal{M}, (w, s) \models \alpha int: \varphi$ gdw. für alle $w' \in I_s^\alpha(w)$ gilt, dass $\mathcal{M}, (w', s) \models \varphi$.

Für die Pfadformeln ist die Erfüllbarkeit dann definiert, wie folgt:

5. Von Stit und BDI zu bdi-stit

$\mathcal{M}, (w, s_0), (w, s_1), \dots \models \varphi$	gdw.	$(w, s_0) \in v(\varphi)$ für atomare Formeln.
$\mathcal{M}, (w, s_0), (w, s_1), \dots \models \neg\varphi$	gdw.	$\mathcal{M}, (w, s_0), (w, s_1), \dots \not\models \varphi$.
$\mathcal{M}, (w, s_0), (w, s_1), \dots \models (\varphi \wedge \psi)$	gdw.	$\mathcal{M}, (w, s_0), (w, s_1), \dots \models \varphi$ und $\mathcal{M}, (w, s_0), (w, s_1), \dots \models \psi$.
$\mathcal{M}, (w, s_0), (w, s_1), \dots \models X\varphi$	gdw.	$\mathcal{M}, (w, s_1), \dots \models \varphi$.
$\mathcal{M}, (w, s_0), \dots \models (\psi U \varphi)$	gdw. ¹¹	es gibt $k \geq 0$: $\mathcal{M}, (w, s_k), \dots \models \varphi$ und für $0 \leq j < k$: $\mathcal{M}, (w, s_j), \dots \models \psi$.

Der Operator E ist der zu A duale Operator. Es gilt $E\varphi = \neg A\neg\varphi$. Betrachtet man lediglich Zustandsformeln und lässt Pfade und A -, E - und temporale Operatoren außer Acht, wird anstatt von Welten und Zuständen eine Menge M von Welt-Zuständen (w, s) betrachtet, auf denen die Zugänglichkeitsrelationen B, D, I gewöhnliche Zugänglichkeitsrelationen für normale Modaloperatoren darstellen. Für zwei $(w, s), (w', s') \in M$ gilt dann, dass (w, s) und (w', s') in der Zugänglichkeitsrelation für Überzeugungen genau dann stehen, wenn $s = s'$ und $(w, s, w') \in B$ ist. Für D und I gilt dies analog.

5.2 Die bdi-stit Logik

In diesem Abschnitt möchte ich die Basis für die *bdi-stit* Logik legen, die in Kapitel 6 erweitert und vollständig axiomatisiert wird. Bei dieser Logik handelt es sich um eine Aussagenlogik, die um weitere modale und temporale Operatoren ergänzt wird. Insbesondere wird verworfen, Teile einer Prädikatenlogik bzw. einer dynamischen Logik aufzunehmen, da die Untersuchung der Relationen zwischen Handlungen, Intentionen und Überzeugungen im Mittelpunkt steht. Wenn ein logisches System betrachtet wird, das Überzeugungs-, Wunsch- und Intentionoperatoren enthält, scheint es aus der Historie der Disziplin nur natürlich zu sein, sich insbesondere von der *BDICTL** abzugrenzen. In diesem Abschnitt wird nicht nur die Syntax und Semantik von *bdi-stit* definiert, sondern bereits auf die Unterschiede zu *BDICTL** hingewiesen.

¹¹Es ist zu beachten, dass z.B. in [64, 65, 66] die Wahrheitsbedingung für den Operator $(\psi \tilde{U} \varphi)$ auf einem Pfad als erfüllt definiert wurde, falls es entweder einen späteren Zustand gibt, in dem φ gilt, und in allen dazwischen liegenden Zuständen ψ gilt oder in jedem Zustand des Pfades ψ gilt. Da sich \tilde{U} mit dem hier definierten Operator U darstellen lässt, $(\psi \tilde{U} \varphi) = ((\psi U \varphi) \vee \neg(\top U \neg\psi))$, verzichte ich auf die zweite Bedingung der Disjunktion aus [64, 65, 66].

Welche Semantik den Operatoren für Überzeugungen und Intentionen zukommen sollte, welche Rationalitätsforderungen Interaktionen zwischen diesen Operatoren plausibilisieren und wie diese in Axiome gefasst werden können, habe ich in Kapitel 4 motiviert. Wie diese nun in der Logik modelliert sind, welche Eigenschaften die Operatoren besitzen und in welchen Relationen sie zueinander stehen, wird in diesem Abschnitt und in Kapitel 6 festgehalten.

Bevor ein Modell einer Akteur-Handlung-Überzeugungslogik vorgestellt wird, wird der syntaktische Aufbau der *bdi-stit* Sprache definiert. Das Alphabet besteht aus den logischen Konstanten \perp , \top , einer abzählbar endlich oder unendlichen Menge von Aussagebuchstaben p, q, r, p_1, p_2, \dots , einer abzählbaren Menge von Variablen $\alpha, \beta, \gamma, \alpha_1, \alpha_2, \dots$, die Platzhalter für eine Menge von Akteuren darstellen, und einer Menge von Operatoren $\{bel, des, int, cstit, \square, \diamond\}$ sowie den Junktoren \neg, \wedge . Über diesem Alphabet werden induktiv *bdi-stit* Formeln definiert.

- Definition 7**
1. Jeder Aussagebuchstabe p, q, r, p_1, p_2, \dots sowie die Konstanten \perp, \top sind *bdi-stit* Formeln.
 2. Falls α, β Variablen für Akteure sind, dann ist $(\alpha = \beta)$ eine *bdi-stit* Formel.
 3. Falls φ, ψ *bdi-stit* Formeln sind und α eine Akteursvariable ist, dann sind $\neg\varphi, (\varphi \wedge \psi), \square\varphi, \diamond\varphi, \alpha cstit:\varphi, \alpha bel:\varphi, \alpha des:\varphi$ und $\alpha int:\varphi$ *bdi-stit* Formeln.
 4. Nichts anderes ist eine *bdi-stit* Formel.

Formeln, die sich anhand von Teil 1 und 2 der Definition 7 bilden lassen, werden als atomare Formeln bezeichnet. Eine Formel der Form $\alpha op : \varphi$ für einen Operator op aus der Menge $\{bel, des, int, cstit\}$ steht für eine Aussage der Form: der Akteur, der durch die Variable α benannt wird, glaubt, wünscht, beabsichtigt oder bewirkt, dass φ . D.h., er hat die Überzeugung, den Wunsch, die Absicht oder bewirkt, dass die Aussage, die durch φ symbolisiert wird, wahr ist. Somit ist die Formel $\alpha op : \varphi$ selbst eine Aussage.

Der *dstit*-Operator wird hierbei als abgeleiteter Operator von *cstit* und Notwendigkeitsoperator definiert, da sich die vollständige Axiomatisierung einer deliberativen Stit-Logik mittels des *cstit*-Operators übersichtlich formulieren lässt, vgl. [168]. Ein Operator, der wie der von Horty benannte *cstit*-Operator fungiert, wurde erstmalig von Chellas in [38] eingeführt. Daher schuf Horty das Akronym *cstit* aus Chellas-Stit. Heutzutage wird vor allem in der Stit-Theorie, die sich mit Gruppen befasst, der *cstit*-Operator betrachtet, um Gruppenhandlungen zu beschreiben. Die Operatoren *dstit* und *astit* treten in den Hintergrund. In Kapitel 1 und Abschnitt 3.3 hatte ich

5. Von Stit und BDI zu bdi-stit

begründet, dass ein *dstit*-Operator nichtsdestotrotz am ehesten geeignet ist, um Handlungen eines Akteurs zu beschreiben. In Abschnitt 3.3 S.106 wurde erläutert, dass sich zur Beschreibung mentaler Akte der *cstit*-Operator eignet, da er keine Negativ-Bedingung an die äußeren Umstände des Akteurs stellt. Die Ausführung eines mentalen Aktes resultiert zudem immer in dem angestrebten mentalen Zustand. In Abschnitt 6.4, wenn der spezifizierte Intentionoperator *dint* eingeführt wird, wird dies deutlich. Das Theorem (*TAI*) beschreibt, dass eine Intention $\alpha \text{dint} : \varphi$ immer ein Zustand nach einem mentalen Akt $\alpha \text{cstit} : \alpha \text{dint} : \varphi$ ist.

Der Operator $\Box\varphi$ wird wie üblich als historische Notwendigkeit gelesen, wohingegen der Operator $\Diamond\varphi$ als ‘es ist kognitiv möglich’ oder auch ‘mental zugänglich, dass φ ’ interpretiert wird.¹²

Wenn eine logische Sprache Konzepte unterschiedlicher Sprachen wie der Stit-Theorie und der *BDI*-Logiken vereinigen soll, stellt sich spätestens bei der Modellkonstruktion und damit bei der Rahmenkonstruktion die Frage, welche der zur Vereinheitlichung gedachten Strukturen der neu gebildeten Sprache zugrunde gelegt werden sollten. Da sowohl die Stit-Theorie als auch die *BDICTL** auf einer sich verzweigenden Zeitlogik basiert, ist eine solche auch unstrittig Grundlage eines *bdi-stit*-Rahmen. Allerdings enden die Gemeinsamkeiten hier. Schon bei dem Begriff der Situation über einer BT Struktur, vgl. Definition 1, und einer Kripke-Struktur, vgl. Definition 6, unterscheidet sich das Konzept in der Stit-Theorie fundamental von demjenigen in der *BDICTL**. In der Stit-Theorie ergibt sich die Menge der Situationen bereits aus der BT Struktur. Die Situationen sind eine Teilmenge des direkten Produktes der Geschichten und Momente.

In der *BDICTL** werden weitere explizite Benennungen notwendig. So existiert keine zeitliche Ordnung auf den Zuständen, sondern eine Zugänglichkeitsrelation, die für jeden Zustand die möglichen Nachfolgerzustände in einer Welt angibt. Jeder Welt werden eine Teilmenge von Zuständen und eine Zugänglichkeitsrelation auf diesen Zuständen zugeordnet, so dass Pfade (Folge von Zuständen) in dieser Welt definiert werden. Dies hat zur Folge, dass ein und derselbe Zustand im Pfad an anderer Stelle erneut eintritt. Demzufolge sollte eher von Zustandstypen anstatt von Zuständen

¹²Die Existenz des \Diamond -Operators resultiert aus den verschiedenen Formen der Modalitäten. Eine Aussage, die historisch möglich oder sogar notwendig ist, wird unter Umständen nie als Tatsache in dem Sinne erkannt, dass irgendein Akteur eine kognitive Einstellung zu der durch die Aussage φ ausgedrückten Proposition einnimmt, d.h., dass φ weder gewünscht noch geglaubt noch anderweitig Gehalt einer propositionalen Einstellung wird. Umgekehrt heißt es nicht, nur weil etwas mental zugänglich ist, dass es tatsächlich historisch möglich ist.

gespröchen werden.

Ein Welt-Zustands-Paar kann analog zum Begriff der Situation aufgefasst werden, wobei ein Pfad einen Geschichtsverlauf darstellt und die möglichen Geschichtsverläufe zu Welten zusammengefasst werden. Welt-Zustands-Paare treten dann entlang eines Pfades mehrfach auf. Man geht dann von wiederkehrenden Situationen (Typen) aus. Die Angabe der Welten, der Zustände, die zu jeder Welt gehören, und die Menge der Zugänglichkeitsrelationen in jeder Welt lassen den Rahmen eines *BDICTL** Modells sehr unübersichtlich erscheinen. Es ist somit nötig, nicht nur eine Menge von Welten anzugeben, sondern auch eine Menge von Zuständen zu jeder Welt und eine Relation auf der zugehörigen Menge von Zuständen zu spezifizieren, die die möglichen Nachfolgerzustände ergeben.

Um einen Rahmen sowohl für *BDICTL**-Operatoren als auch für Stit-Operatoren zu konstruieren, ist es nur natürlich, sich für eine der beiden Konzepte von Situationen zu entscheiden. Ein wichtiges Kriterium sollte dabei sein, dass der Rahmen möglichst einfach definiert ist. In *BDICTL** ist das Konstrukt der auf der BT definierten Welten überflüssig. Es vergrößert unnötig, wie in Abschnitt 6.3 gezeigt wird, die Anzahl der in einem Rahmen zu spezifizierenden Mengen und enthält keine wirklich neuen Informationen. Umständlich zu handhaben, ist auch die Unterscheidung in Pfad- und Zustandsformeln in *BDICTL**. Nicht nur der Begriff der Situation ist schwer zu beschreiben, da es sich dabei sowohl um Zustände als auch um Pfade handeln kann, die selbst Aneinanderreihungen von Zustandstypen sind. Es ist zudem erforderlich, bei der Erfüllbarkeit zwischen Pfad- und Zustandsformeln zu unterscheiden. Diese Unterscheidung dient dazu, dass in *BDICTL** Zustandsformeln notwendigerweise erfüllt sind, während Pfadformeln über verschiedene Pfade mit demselben Ausgangszustand ihren Wahrheitswert ändern. Bei näherem Betrachten, insbesondere für aussagenlogische Formeln, mutet dies jedoch entbehrlich an. In Abschnitt 1.3 habe ich diskutiert, dass es in Stit-Theorie möglich ist, Formeln in gemeinsamen vergangenen Geschichtsverläufen (ausgenommen dem aktuellen Moment) denselben Wahrheitswert zuzuschreiben. Daher scheint eine Restriktion, dass alle Zustandsformeln (einschließlich des aktuellen Momentes) notwendigerweise wahr bzw. falsch sein müssen, zu stark und unnötig.

Ein Situationskonzept mit Moment-Geschichte-Paaren macht eine solche Differenzierung zwischen Zustands- und Pfadformeln (zwischen Moment- und Geschichteformel) nicht mehr zwingend. Das Konzept der überlappenden Welten in der *BDICTL**-Semantik wird vereinfacht und auf lineare Geschichten projiziert, ohne an Aussagekraft zu verlieren. Somit ließ die Nichtunterscheidung von Zustands- und

5. Von Stit und BDI zu bdi-stit

Pfadformeln sowie die Forderung, den Rahmen möglichst einfach zu halten, die Wahl auf die Moment-Geschichte-Paare als Situationen für *bdi-stit* fallen. Dies führt jedoch zu dem Verlust einer endlichen vollständigen Axiomatisierung.¹³

Die Grundlage eines *bdi-stit* Rahmen ist somit eine BT Struktur (M, \leq) . Wie in der Stit-Theorie ergibt sich die Menge der Geschichten H als die Menge aller maximalen, linearen Teilmengen von M , siehe Abschnitt 1.1. Eine Situation ist jedes Moment-Geschichte-Paar (m, h) , wobei $m \in h$. Die Menge aller Situationen wird durch S symbolisiert. Auf dieser zeitlichen BT Struktur werden die Interpretationsfunktionen der einzelnen Operatoren definiert. Da die Interpretation der Handlungsoperatoren aus der Stit-Theorie übernommen wird, ist die Interpretationsfunktion für *cstit* und somit auch für den *dstit*-Operator in *bdi-stit* ebenso wie in der Stit-Theorie definiert, vgl. Abschnitt 1.2. Die Funktion C in Stit-Theorie ordnet die Menge der Geschichten, die durch einen Moment verlaufen, für jeden Akteur in disjunkte Teilmengen ein, wobei je zwei Geschichten, die zu einer Teilmenge gehören, für den Akteur ununterscheidbar und wahläquivalent sind. Eine solche Teilmenge wird daher als Wahlzelle des Akteurs in diesem Moment bezeichnet.

In *bdi-stit* weist die Funktion C jedem Akteur-Situations-Paar $(\alpha, (m, h))$ die Wahlzelle zu, die die Situationen enthält, die der Akteur in dem Moment m von der Geschichte h nicht von den Situationen (m, h') mit $h' \in H_m$ unterscheiden kann. Dies ist sinnvoll, da es für die Interpretation des Ausdrucks $\alpha \text{ cstit} : \varphi$ in einer gegebenen Situation irrelevant ist, was α in einer anderen Geschichte nicht von dieser hätte unterscheiden können. Formal ergibt sich folgendes Bild

$$C : \mathcal{A} \times S \rightarrow \mathcal{P}(H), \quad (\alpha, s) \mapsto C(\alpha, s) = C_s^\alpha.$$

Für eine gegebene Situation $s = (m, h)$ benennt C_s^α die Wahlzelle, die aus den Geschichten besteht, die Akteur α nicht von h durch seine Handlungen im Moment m differenzieren kann. Wie in Stit-Theorie üblich, lässt sich zu gegebenem Moment eine Äquivalenzrelation definieren, wobei die Wahlzellen die einzelnen Äquivalenzklassen darstellen. Daher muss für C gelten, dass $H_m = \cup \{ C_{(m,h)}^\alpha \mid h \in H_m \}$ und für je-

¹³Dies jedoch nur, wenn temporale Operatoren hinzugenommen werden und die ungebündelte BT Struktur betrachtet wird. Jede auf einer gebündelten BT Struktur gültige Formel ist ebenfalls auf der vollständigen Struktur gültig, vgl. [63, S.68]. Jedoch ist auf einer gebündelten Struktur nicht jede über der vollständigen ockhamistischen Logik gültige Formel ebenfalls gültig. Die gebündelte Struktur stellt somit eine Generalisierung der ungebündelten Struktur dar, vgl. Abschnitt 5.3. Da ich in diesem Abschnitt keine temporalen Operatoren betrachte, ist eine gebündelte Struktur umständlicher zu handhaben, als eine nicht gebündelte Struktur. Für die ungebündelte CTL^* ist von Reynolds [115] eine Axiomatisierung angegeben worden. Für eine ungebündelte BT Struktur steht diese noch aus.

des $m \in M$, $h \in H_m$ und jede Situation $s' = (m, h')$ mit $h' \in H_m$, dass entweder $C_{(m,h)}^\alpha = C_{(m,h')}^\alpha$ oder $C_{(m,h)}^\alpha \cap C_{(m,h')}^\alpha = \emptyset$.

Um konfligierende Überzeugungen und antagonistische Wünsche zu modellieren, ist eine relationale Semantik inadäquat. Daher wird in der *bdi-stit* Logik eine Form der Nachbarschaftssemantik verwandt, vgl. [39, 55, 100, 121]. Diese macht es möglich, dass Formeln vom Typ $(\alpha des : \varphi \wedge \alpha des : \neg\varphi)$ und $(\alpha bel : \varphi \wedge \alpha bel : \neg\varphi)$ in *bdi-stit* erfüllbar sind, wobei Formeln der Form $\alpha des : (\varphi \wedge \neg\varphi)$ und $\alpha bel : (\varphi \wedge \neg\varphi)$ kontradiktorisch sind.¹⁴

In einer Nachbarschaftssemantik wird jeder Situation s eine Menge von Welten¹⁵ zugeordnet. Jede Welt besteht aus einer Menge von Situationen, die *Umgebung* oder *Nachbarschaft* von s genannt wird, da sie Situationen darstellt, die von s aus erreichbar sind. Die Menge aller Umgebungen von s wird als *Nachbarschafts-* bzw. *Umgebungssystem* von s bezeichnet.

Die Menge der Welten, aus der die Umgebungssysteme der einzelnen Situationen konstruiert werden, macht den Operator der kognitiven Möglichkeiten sinnvoll. Wenn jemand in einer Situation etwas wünscht, glaubt oder beabsichtigt, ist dies dem Akteur auf eine gewisse Art mental zugänglich. Die Formel $\diamond\varphi$ drückt aus, dass der Sachverhalt φ auf irgendeine Art Gehalt eines mentalen Zustandes eines Akteurs sein kann. Während bei einer historisch möglichen Aussage es der Fall sein muss, dass diese zum aktuellen Zeitpunkt tatsächlich in irgendeinem Geschichtsverlauf realisiert ist, ist eine kognitive Möglichkeit $\diamond\varphi$ auch ohne temporalen Bezug zur aktuellen Situation erfüllbar, wenn die Aussage φ in einer zugänglichen Menge von zeitlich beliebigen Situationen (beliebiger Moment und Geschichtsverlauf) realisiert ist.

Der Inhalt eines Wunsches oder einer Überzeugung kann sogar unmöglich in der aktuellen Situation des Wünschens oder des Glaubens sein. Z.B. kann er an einem Regentag wünschen, dass die Sonne scheint. Da es sich um rationale Akteure handelt, ist es ihnen zwar möglich, konfligierende Wünsche zu haben, dass die Sonne scheint, weil er baden gehen möchte, und dass es regnet, da der Akteur nicht selbst Blumen gießen möchte. Es ist ihm auch möglich, rational Wünsche zu haben, die der aktuellen Situation widersprechen. Jedoch ist es rational unmöglich, sich Widersprüche zu wünschen, z.B. dass er wünscht, dass die Sonne scheint und nicht scheint, vgl. Ab-

¹⁴Für die Motivation siehe Abschnitt 4.1 die Diskussion von antagonistischen Wünschen im Gegensatz zu paradoxen Wünschen und konfligierenden Überzeugungen im Gegensatz zu inkonsistenten Gehalten von Überzeugungen.

¹⁵Man beachte an dieser Stelle, dass der Terminus 'Welt' in *BDICTL**-Logik und *bdi-stit*-Logik verschieden gebraucht wird.

5. Von Stit und BDI zu bdi-stit

schnitt 4.2.

Eine durch φ ausgedrückte Proposition ist dann mental zugänglich in einer Situation s , wenn es eine zugängliche Umgebung (Welt) von s gibt, in der die Aussage φ in jeder Situation wahr ist.¹⁶ Einzige Bedingung an eine Welt ist, dass es mindestens eine Situation geben muss, in der φ erfüllt ist, um kontradiktorische Aussagen auszuschließen. Eine Welt bzw. eine Umgebung ist eine nicht-leere Menge von Situationen.

In der Stit-Theorie ist jede historische Unmöglichkeit eine Form der Notwendigkeit $\neg\Diamond\varphi \equiv \Box\neg\varphi$ und umgekehrt. Was heißt, dass eine durch φ ausgedrückte Proposition mental notwendigerweise zugänglich bzw. kognitiv unmöglich ist? Für ersteres sollte sie in irgendeiner Form, denkbar, wünschbar, vorstellbar oder wahrnehmbar in jeder zugänglichen Situation sein. D.h. der Sachverhalt φ sollte in allen mental zugänglichen Situationen in allen Welten bestehen. Folglich ist es nicht mental zugänglich, dass der Sachverhalt φ nicht besteht. In allen Umgebungen von s muss die Aussage φ demnach wahr sein. Aus der kognitiven Notwendigkeit folgt somit, dass die Negation von φ kognitiv unmöglich ist. Umgekehrt, wenn $\neg\varphi$ kognitiv unmöglich ist, kann sich niemand vorstellen, denken, wahrnehmen, wünschen, usw., dass $\neg\varphi$. Sagt das etwas über die kognitive Modalität von φ aus? Das tut es nicht. Eine Aussage der Form, dass sowohl $\neg\varphi$ mental unzugänglich als auch φ nicht notwendigerweise mental zugänglich ist, ist nicht zwangsläufig inkonsistent. Es gibt Aussagen, zu denen niemand eine mentale Einstellung hat. Es ist also nicht möglich, einen Begriff der Notwendigkeit von mentaler Zugänglichkeit als Unmöglichkeit der mentalen Zugänglichkeit der Negation zu definieren. Daher wird im Folgenden lediglich der Fall der mentalen Zugänglichkeit (kognitiven Möglichkeit) betrachtet.

Der Operator der mentalen Zugänglichkeit wird durch \Diamond ausgedrückt. Für jede Situation s wird ein Umgebungssystem N_s eingeführt. Dies wird im Rahmen eines *bdi-stit* Modells durch eine Funktion N ,

$$N : S \rightarrow \mathcal{P}(\mathcal{P}(S) \setminus \emptyset), \quad s \mapsto N(s) = N_s$$

beschrieben. Die Funktion bildet jede Situation auf eine Teilmenge N_s der Potenzmenge aller Situationen ab. Jede Menge U in N_s stellt eine Welt dar. In einem gegebenen Rahmen mit einer Nachbarschaftsfunktion N wird unter N zudem die Vereinigung aller Nachbarschaftssysteme $N = \cup\{N_s \mid s \in S\}$ verstanden. Aus dem Kontext geht eindeutig hervor, ob N die Menge aller Welten oder die Funktion bezeichnet.

¹⁶Wenn die Aussage wahr in jeder Situation einer Welt ist, wird auch abkürzend von wahr in der Welt gesprochen.

Für die Interpretation von Überzeugungs-, Wunsch- und Intentionsooperatoren werden ebenfalls Interpretationsfunktionen im Rahmen benötigt. Um konfligierende Überzeugungen bzw. antagonistische Wünsche zu modellieren, werden Nachbarschaftsfunktionen verwendet. Die Funktion B ordnet jedem Akteur-Situations-Paar (α, s) die Menge der Welten zu, die kompatibel sind, mit dem was α in der Situation s glaubt. Es ist dabei zu beachten, dass die Überzeugungen kompatibel sein müssen, die in einer Welt repräsentiert werden. Lediglich durch die Zuordnung mehrerer Welten wird erreicht, dass ein Akteur konfligierende Überzeugungen haben kann. Analog wird die Funktion D definiert, um einem Akteur die Menge der Welten zuzuordnen, die kompatibel mit dem sind, was er in der Situation wünscht. Es ergibt sich formal

$$D : \mathcal{A} \times S \rightarrow \mathcal{P}(N), \quad (\alpha, s) \mapsto D(\alpha, s) = D_s^\alpha,$$

$$B : \mathcal{A} \times S \rightarrow \mathcal{P}(N), \quad (\alpha, s) \mapsto B(\alpha, s) = B_s^\alpha.$$

Auch wenn ein rationaler Akteur antagonistische Wünsche und Überzeugungen besitzt, so trifft dies nicht für seine Intentionen zu, siehe Abschnitt 4.2. Intentionen sind Vorstufen von Handlungen. Es ist sicher nicht möglich, dass in ein und derselben Situation ein Akteur sowohl φ als auch $\neg\varphi$ bewirken kann. Somit ist es ebenfalls nicht möglich, dass der Akteur rational sowohl φ als auch $\neg\varphi$ in ein und derselben Situation beabsichtigen kann. In dem obigen Fall, wo er wünscht, dass die Sonne scheint, und wünscht, dass es regnet, ist dies kein Grund anzunehmen, dass der Akteur rational intendiert, dafür zu sorgen, dass die Sonne scheint und dass es regnet. Aber auch wenn wir etwas wünschen, auf das wir einwirken können, würden wir, sobald wir von Absichten reden, Antagonismen ausschließen. In dem Moment, wo der Akteur die Absicht hat, seinem Bruder eine Niere zu spenden, kann er noch den Wunsch haben, dies nicht zu tun. Aber er wird nicht die Absicht haben, die Niere nicht zu spenden. Die Intention, dass φ , zu haben, schließt aus zu intendieren, dass $\neg\varphi$. Es ist immer möglich, weder φ noch $\neg\varphi$ zu beabsichtigen. Die Verwendung einer Nachbarschaftssemantik ist hierfür nicht nötig, so dass wie in *BDICTL** für den Intentionsooperator eine relationale Semantik definiert wird,

$$I : \mathcal{A} \times S \rightarrow N, \quad (\alpha, s) \mapsto I(\alpha, s) = I_s^\alpha.$$

Die Interpretationsfunktion I des Intentionsooperators ordnet jedem Akteur-Situations-Paar eine Menge von Situationen (Welt) zu, die kompatibel sind, mit dem was der Akteur in der Situation intendiert. Da eine Welt nicht leer ist, ist es den Akteuren unmöglich, paradoxe Intentionen zu hegen. Die weiteren ebenfalls motivierten Interaktionsaxiome aus den Abschnitten 4.2 und 4.3 zwischen den Handlungsoperato-

5. Von Stit und BDI zu bdi-stit

ren und Wunschoperatoren für rationale Akteure werden in Abschnitt 6.4 modelliert. Dafür ist es erforderlich, einen spezifizierten Intentionoperator zu betrachten.

Ein *bdi-stit* Modell \mathcal{M} , in dem eine Formel nach Definition 7 bewertet wird, besteht aus einem Rahmen $\mathcal{F} = (M, \leq, \mathcal{A}, N, C, B, D, I)$ und einer Bewertungsfunktion v . Die Bewertungsfunktion v bildet die Aussagebuchstaben in die Potenzmenge der Menge aller Situationen ab. Die Menge $v(p)$ entspricht der Menge der Situationen, in denen die Aussage p wahr ist. Für die Konstanten \perp , \top und jede Bewertungsfunktion gilt $v(\perp) = \emptyset$ und $v(\top) = S$. Den Akteursvariablen ordnet die Bewertungsfunktion v die Akteure \mathcal{A} aus dem Rahmen zu. Sodann lässt sich Erfüllbarkeit in einem *bdi-stit* Modell $\mathcal{M} = (\mathcal{F}, v)$ definieren.

Definition 8 (*bdi-stit Semantik*) Sei $s = (m, h)$ eine Situation, seien α, β Akteursvariablen und seien φ, ψ Formeln gemäß Definition 7, dann gelte:

$\mathcal{M}, s \models \varphi$	gdw.	$s \in v(\varphi)$, falls φ eine Formel gemäß 7.1 ist.
$\mathcal{M}, s \models \alpha = \beta$	gdw.	$v(\alpha) = v(\beta)$.
$\mathcal{M}, s \models \neg\varphi$	gdw.	$\mathcal{M}, s \not\models \varphi$.
$\mathcal{M}, s \models \varphi \wedge \psi$	gdw.	$\mathcal{M}, s \models \varphi$ und $\mathcal{M}, s \models \psi$.
$\mathcal{M}, s \models \Box\varphi$	gdw.	$\mathcal{M}, (m, h') \models \varphi$ für alle $h' \in H_{(m)}$.
$\mathcal{M}, s \models \Diamond\varphi$	gdw.	es ein $U \in N_s$ mit $U \subseteq \{s' \mid \mathcal{M}, s' \models \varphi\}$ gibt.
$\mathcal{M}, s \models \alpha \text{ cstit} : \varphi$	gdw.	$\{(m, h') \mid h' \in C_s^\alpha\} \subseteq \{(m, h') \mid \mathcal{M}, (m, h') \models \varphi\}$,
$\mathcal{M}, s \models \alpha \text{ bel} : \varphi$	gdw.	es ein $U \in B_s^\alpha$ mit $U \subseteq \{s' \mid \mathcal{M}, s' \models \varphi\}$ gibt.
$\mathcal{M}, s \models \alpha \text{ des} : \varphi$	gdw.	es ein $U \in D_s^\alpha$ mit $U \subseteq \{s' \mid \mathcal{M}, s' \models \varphi\}$ gibt.
$\mathcal{M}, s \models \alpha \text{ int} : \varphi$	gdw.	$I_s^\alpha \subseteq \{s' \mid \mathcal{M}, s' \models \varphi\}$

Als abkürzende Schreibweise wird $s \models \varphi$ verwandt, falls ersichtlich wird, welches Modell gemeint ist. Für die Menge $\{s' \mid s' \models \varphi\}$ aller Situationen eines bekannten Modells \mathcal{M} , die eine gegebene Formel φ erfüllen, wird kurz $\|\varphi\|$ geschrieben. Diese Menge wird als sogenannte Wahrheitsmenge („truth set“) von φ bezeichnet, vgl. z.B. [39, S.38]. Offensichtlich ist es in einem *bdi-stit* Modell möglich, dass die historische Notwendigkeit \Box und die kognitive Möglichkeit \Diamond zusammenfallen, nämlich genau dann, wenn jede Welt eines beliebigen Nachbarschaftssystem $N_{(m,h)}$ jedes Moment-Geschichte-Paar (m, h') mit $m \in h'$ und $h' \in H_m$ enthält.

Bemerkung Sei $\mathcal{M} = (M, \leq, \mathcal{A}, N, C, B, D, I, v)$ ein *bdi-stit* Modell. Falls für die Nachbarschaftsfunktion gilt,

$$N : (m, h) \mapsto \{U \subseteq S \mid \forall h' \in H_{(m)} : (m, h') \in U\},$$

5.2. Die *bdi-stit* Logik

	BDI_{CTL}^K	$B^{KD45}D^{KD}I_{CTL}^{KD}$	<i>bdi-stit</i>
inkonsistente Überzeugungen $\alpha \text{ bel} : (\varphi \wedge \neg\varphi)$	erfüllbar	nicht erfüllbar	nicht erfüllbar
konfligierende Überzeugungen I $\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : \neg\varphi$	erfüllbar	nicht erfüllbar	erfüllbar
konfligierende Überzeugungen II $\alpha \text{ bel} : \varphi \supset \neg\alpha \text{ bel} : \neg\varphi$	widerlegbar	gültig	widerlegbar
abgeschlossen unter geglaubter Implikation (K_b) $(\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : (\varphi \supset \psi)) \supset \alpha \text{ bel} : \psi$	gültig	gültig	erfüllbar
Abgeschlossenheit (C_b) $(\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : \psi) \supset \alpha \text{ bel} : (\varphi \wedge \psi)$	gültig	gültig	erfüllbar
Monotonie (M_b, RM_b) $\alpha \text{ bel} : (\varphi \wedge \psi) \supset (\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : \psi)$	gültig	gültig	gültig
pos. Introspektion (4_b) $\alpha \text{ bel} : \varphi \supset \alpha \text{ bel} : \alpha \text{ bel} : \varphi$	erfüllbar	gültig	erfüllbar
neg. Introspektion (5_b) $\neg\alpha \text{ bel} : \varphi \supset \alpha \text{ bel} : \neg\alpha \text{ bel} : \varphi$	erfüllbar	gültig	erfüllbar

Tabelle 5.1: Axiome des Überzeugungsoperators in BDI_{CTL}^K , in $B^{KD45}D^{KD}I_{CTL}^{KD}$ und in *bdi-stit* Logik

gilt für jede Formel φ

$$\mathcal{M}, (m, h) \models \Box\varphi \text{ gdw. } \mathcal{M}, (m, h) \models \Diamond\varphi.$$

Bevor ich für einen bestimmten Teil der Sprache von BDI_{CTL}^* zeige, dass sie sich in *bdi-stit* wiederfinden lässt, möchte ich mit den Tabellen 5.1-5.2. die unterschiedlichen Axiome und Theoreme in relationaler und Nachbarschaftssemantik kurz deutlich machen. Die Motivation für diese Änderung der Axiomatisierung von BDI_{CTL}^* zu *bdi-stit* ist in Abschnitt 4.1 dargelegt worden. Dafür werden zwei Standardaxiomatisierung von Wunsch-, Überzeugungs- und Intensionsoperator betrachtet, wie sie in den BDI Logiken von Rao und Georgeff diskutiert wurden [66, S.305], BDI_{CTL}^K und $B^{KD45}D^{KD}I_{CTL}^{KD}$ der *bdi-stit* Axiomatisierung gegenüber gestellt. In Tabelle 5.1 sind die Eigenschaften des Überzeugungsoperators in den jeweiligen Logiken festgehalten. Analog ist die Tabelle 5.2 für den Wunschoperator aufgestellt.

5. Von Stit und BDI zu bdi-stit

	BDI_{CTL}^K	$B^{KD45}D^{KD}I_{CTL}^{KD}$	$bdi-stit$
kontradiktorische Wünsche $\alpha des: (\varphi \wedge \neg\varphi)$	erfüllbar	nicht erfüllbar	nicht erfüllbar
antagonistische Wünsche I $\alpha des: \varphi \wedge \alpha des: \neg\varphi$	erfüllbar	nicht erfüllbar	erfüllbar
antagonistische Wünsche II $\alpha des: \varphi \supset \neg\alpha des: \neg\varphi$	widerlegbar	gültig	widerlegbar
abgeschlossen unter gewünschter Implikation (K_d) $(\alpha des: \varphi \wedge \alpha des: (\varphi \supset \psi)) \supset \alpha des: \psi$	gültig	gültig	widerlegbar
Abgeschlossenheit (C_d) $(\alpha des: \varphi \wedge \alpha des: \psi) \supset \alpha des: (\varphi \wedge \psi)$	gültig	gültig	widerlegbar
Monotonie (M_d, RM_d) $\alpha des: (\varphi \wedge \psi) \supset (\alpha des: \varphi \wedge \alpha des: \psi)$	gültig	gültig	gültig
pos. und neg. Introspektion ($4_d, 5_d$)	erfüllbar	erfüllbar	erfüllbar

Tabelle 5.2: Axiome des Wunschoperators in BDI_{CTL}^K , in $B^{KD45}D^{KD}I_{CTL}^{KD}$ und in $bdi-stit$ Logik

5.3 Eine Generalisierung von Stit-Theorie und

$$B^{KD45}D^{KD}I_{CTL}^{KD}$$

In diesem Abschnitt soll gezeigt werden, dass es sich bei der $bdi-stit$ Logik tatsächlich um eine Verallgemeinerung von der deliberativen Stit-Logik und einer $BDICTL^*$ Logik handelt. Wird Logik als ein Paar $\mathcal{L} = (L, T)$ bestehend aus einer formalen Sprache L und einer Menge von Theoremen T verstanden, wird eine Logik als Verallgemeinerung einer anderen Logik aufgefasst, wenn folgendes gilt: Die formale Sprache der zu verallgemeinernden Logik ist in der Sprache der verallgemeinernden Logik enthalten.¹⁷ Für die Theoreme gilt, falls die generalisierende (verallgemeinernde) Logik auf die formale Sprache der Ausgangslogik (der generalisierten oder zu verallgemeinernden Logik) beschränkt wird, dass die Menge der Theoreme der generalisierten

¹⁷Die Ideen und Beweise aus diesem Abschnitt und die Motivation der konfligierenden Überzeugungen bei rationalen Akteuren aus Abschnitt 4.1 sind in [126] bereits veröffentlicht.

5.3. Eine Generalisierung von Stit-Theorie und $B^{KD45}D^{KD}I_{CTL}^{KD}$

Logik die Menge der Theoreme der generalisierenden Logik einschließt. Formal ist $\mathcal{L}^G = (L^G, T^G)$ eine Generalisierung (Verallgemeinerung) von \mathcal{L} genau dann, wenn $L \subseteq L^G$ und $T^G|_L \subseteq T$.

Es ist ebenfalls möglich, die Generalisierung über die Erfüllbarkeit in Modellen zu definieren. Sei $L \subseteq L^G$. Gibt es eine Klasse von Modellen der Logik \mathcal{L}^G , so dass die Menge der Theoreme bezüglich dieser Klasse die Menge der Theoreme der verallgemeinernden Logiken darstellt, solange sie auf die Sprache von \mathcal{L} beschränkt sind, generalisiert \mathcal{L}^G die Logik \mathcal{L} . Wird die Generalisierung mittels der Erfüllbarkeit definiert, gilt das Folgende. Eine Logik \mathcal{L} wird durch eine Logik \mathcal{L}^G generalisiert, falls für jedes Modell von \mathcal{L} gilt, dass es eine Abbildung der Situationen in die Situationen eines Modells der Logik \mathcal{L}^G gibt, so dass in der abgebildeten Situation genau dieselben Formeln wie in der abzubildenden Situation erfüllt werden. Somit ist jede Formel über \mathcal{L} erfüllbar in \mathcal{L}^G . Die Umkehrung gilt im Allgemeinen nicht. Des Weiteren ist jedes Theorem der Logik \mathcal{L}^G , insoweit es in der Sprache von \mathcal{L} ist, ein Theorem in \mathcal{L} .

Für ein beliebiges Modell der $B^{KD45}D^{KD}I_{CTL}^{KD}$ oder Stit-Logik gibt es jeweils ein *bdi-stit* Modell und eine Abbildung der Situationen in die Situationen des *bdi-stit* Modells, so dass dieselben Formeln, die im Urbild erfüllt sind, ebenfalls im Bild erfüllt sind. Somit generalisiert die *bdi-stit* Logik die $B^{KD45}D^{KD}I_{CTL}^{KD}$ der *BDICTL** Logiken und eine deliberative Stit-Logik.¹⁸ Die Modelle der zu verallgemeinernden Logik werden punktweise in Modelle der Generalisierung eingebettet. Es muss jedoch für ein beliebiges Modell keine Abbildung der Situationen der Generalisierung auf die zu verallgemeinernde Logik geben, die die Erfüllbarkeit beibehält. So kann die formale Sprache der Generalisierung Formeln umfassen, die nicht bewertbar in der generalisierten Logik sind. Falls die formale Sprache konstant bleibt und es eine solche Abbildung gibt, wird nicht mehr von punktweiser Einbettung von Modellen gesprochen, sondern von punktweiser Äquivalenz zweier Modelle, vgl. [39, S.36].

Falls eine Formelmenge, die zu der formalen Sprache der Ausgangslogik gehört, in dieser Logik erfüllt wird, existiert ein Modell der Generalisierung, welches die Formelmenge ebenfalls erfüllt. Daraufhin ist es möglich, die Klasse der Modelle der verallgemeinernden Logik festzuhalten, in der die Theoreme der zu verallgemeinernden Logik ebenfalls gültige Schemata sind.

Für die Stit-Theorie ist es offensichtlich, dass *bdi-stit* eine Generalisierung in diesem Sinne darstellt. Es gilt für eine Einschränkung der Sprache auf das Vokabular einer deliberativen Stit-Logik, dass die Stit-Modelle punktweise äquivalente Modelle

¹⁸Die *bdi-stit* Logik generalisiert ebenfalls eine BDI_{CTL}^{KD} . Die Erfüllbarkeit paradoxer Wünsche schließt eine Generalisierung von BDI_{CTL}^K durch *bdi-stit* Logik allerdings aus.

5. Von Stit und BDI zu bdi-stit

in *bdi-stit* besitzen.

Bemerkung Sei φ eine Formel nach Definition 2. Dann existiert für jedes Stit-Modell $\mathcal{M}_s = (M, \leq, \mathcal{A}, \tilde{C}, \nu)$ ein *bdi-stit* Modell $\mathcal{M}_{ep} = (M, \leq, \mathcal{A}, N, C, B, D, I, \nu)$, so dass für jede Situation (m, h) das Folgende gilt:

$$\mathcal{M}_s, (m, h) \models \varphi \quad \text{gdw.} \quad \mathcal{M}_{ep}, (m, h) \models \varphi.$$

BEWEIS: Das Modell \mathcal{M}_{ep} übernimmt alle Komponenten des Modells \mathcal{M}_d wie M, \leq, \mathcal{A} sowie die Bewertungsfunktion ν . Die Funktion C von \mathcal{M}_{ep} lässt sich definieren durch $C_{(m,h)}^\alpha = \{h' \in H_m \mid h' \in \tilde{C}_m^\alpha(h)\}$. Da die Definition von N, B, D und I irrelevant ist, weil φ keine Operatoren aus $\{\diamond, bel, des, int\}$ enthält, werden die Nachbarschaftssysteme (nicht die Welten) für jede Situation als leer angenommen. Die Wahl der Funktion I ist ebenfalls irrelevant, z.B. $I_s^\alpha = \{s\}$ für alle $s \in S$. Offensichtlich erfüllt ein so generiertes *bdi-stit* Modell \mathcal{M}_{ep} die Behauptung. \square

Eine ähnliche aber weniger offensichtliche Behauptung wird für $B^{KD45}D^{KD}I_{CTL}^{KD}$ aufgestellt. Ein Problem hierbei ist die Unterscheidung in Pfad- und Zustandsformeln und die unterschiedliche Behandlung der Zustände und Pfade, in denen die Formeln bewertet werden. Daher werde ich mich vorerst auf folgende Behauptung beschränken, die sich lediglich auf eine bestimmte Menge von Zustandsformeln bezieht, und diese später geeignet für Pfadformeln erweitern, siehe Abschnitt 6.3. An dieser Stelle werden *bdi* Formeln betrachtet, die wie folgt aufgebaut sind, vgl. Definition 5 bzw. [64, 65, 66]:

- (S1) Jeder Aussagebuchstabe p_1, p_2, \dots ist eine *bdi* Formel.
- (S2) Falls φ, ψ *bdi* Formeln sind, dann sind $\neg\varphi, (\varphi \wedge \psi)$ *bdi* Formeln.
- (S4) Falls φ eine *bdi* Formel und α eine Akteursvariable ist, dann sind $\alpha bel : \varphi, \alpha des : \varphi$ und $\alpha int : \varphi$ *bdi* Formeln.

Dann gilt folgende Behauptung.

Theorem 9 Sei φ eine solche *bdi* Formel. Dann existiert für jedes Modell $\mathcal{M}_e = (W, \{S_w\}_{w \in W}, \{R_w\}_{w \in W}, \mathcal{A}, B, D, I, \nu)$ ein *bdi-stit* Modell $\mathcal{M}_h = (M', \leq, \mathcal{A}', N', C', B', D', I', \nu')$, so dass für jede (w, m) in \mathcal{M}_e eine Situation (m, h) in \mathcal{M}_h gibt mit:

$$\mathcal{M}_e, (w, s) \models \varphi \quad \text{gdw.} \quad \mathcal{M}_h, (m, h) \models \varphi. \quad (*)$$

5.3. Eine Generalisierung von Stit-Theorie und $B^{KD45}D^{KD}I_{CTL}^{KD}$

BEWEIS: Es wird ein Modell $M_h = (M', \leq, \mathcal{A}', C', N', B', D', I', v')$ konstruiert. Sei $M_e = (W, \{S_w\}_{w \in W}, \{R_w\}_{w \in W}, \mathcal{A}, B, D, I, v)$ ein $BDICTL^*$ Modell, vgl. Definition 6. Die Menge der Welten W wird um ein neues Element w_0 erweitert, welches in der Momentmenge einen neuen Zustand s_0 und eine BT Struktur beinhalten soll, die nicht näher spezifiziert werden muss. Dann sei $M' = \{z = (w, s) \mid s \in S_w\} \cup \{(w_0, s_0)\}$ und für beliebige $z = (w, s)$ und $z' = (w', s')$ aus M' sei die Relation \leq gegeben mit:

$$z \leq z' \quad \text{gdw.} \quad w = w_0.$$

Bei (M', \leq) handelt es sich um eine BT Struktur, d.h. für alle z, z', z'' mit $z' \leq z$ und $z'' \leq z$ gilt, dass $z' \leq z''$ oder $z'' \leq z'$. Dies ist offensichtlich, falls $z = (w_0, s_0)$, ebenfalls offensichtlich, wenn z' oder z'' gleich (w_0, s_0) . Falls alle verschieden von (w_0, s_0) sind, dann sind sie alle samt untereinander nicht vergleichbar.

Die Geschichten H_z , die durch gegebenes $z = (w, s)$ verlaufen, sind entweder alle Geschichten, falls $z = z_0 = (w_0, s_0)$, bzw. ist es genau eine für $z \neq z_0$. Dann kann eine Abbildung ι von der Menge der Zustände in M_e in die Menge der Situationen in M_h angegeben werden,

$$\iota : (w, s) \mapsto ((w, s), h),$$

wobei $h \in H_{(w,s)}$ beliebig gewählt werden kann, kurz $\iota : z \mapsto (z, h)$. Da an dieser Stelle die Sprache auf *bdi* Formeln beschränkt wird, d.h. lediglich Zustandsformeln ausgewertet werden, werden die Zustände in die Situationen von M_h abgebildet. Eine analoge Aussage zu Theorem 9 für Pfade und Pfadformeln erfolgt in Abschnitt 6.3. Dort werden die komplexere Struktur der Pfade ebenfalls in *bdi-stit* Situationen eingebettet. Für beliebigen Akteur $\alpha \in \mathcal{A}$, $z = (w, s) \in M'$ und jedes $h \in H_z$ sowie für gegebenes, atomares φ liegen folgende Definitionen dem Modell M_h zugrunde:

$$\begin{aligned} \mathcal{A}' &= \mathcal{A}, \\ N'(z, h) &= \mathcal{P}(\mathcal{P}(\{(z, h) \mid (z, h) \text{ Situation in } M_h\}) \setminus \emptyset), \\ B'(\alpha, ((w, s), h)) &= \{\{\iota(w', s) \mid (w, s, w') \in B(\alpha)\}\}, \\ D'(\alpha, ((w, s), h)) &= \{\{\iota(w', s) \mid (w, s, w') \in D(\alpha)\}\}, \\ I'(\alpha, ((w, s), h)) &= \{\iota(w', s) \mid (w, s, w') \in I(\alpha)\}, \\ v'(\varphi) &= \{(z, h) \mid h \in H_z, z \in v(\varphi)\}. \end{aligned}$$

Somit gilt für gegebenes $z \in M'$, dass für $h, h' \in H_z$ gilt: $T'(\alpha, (z, h)) = T'(\alpha, (z, h'))$ und für jede Funktion $T' \in \{B', D', I'\}$ und auch $N'(z, h) = N'(z, h')$.

Da die Funktionen B, D, I die Akteure und Situationen auf nicht-leere Mengen abbildet, sind auch die Mengen $B'(\alpha, (z, h)), D'(\alpha, (z, h)), I'(\alpha, (z, h))$ nicht leer. Die

5. Von Stit und BDI zu bdi-stit

Behauptung (*) wird wiederum induktiv gezeigt,

$$\mathcal{M}_e, (w, s) \models \varphi \text{ gdw. } \mathcal{M}_h, \iota(w, s) \models \varphi, \quad (*)$$

wobei ι eine oben beschriebene Abbildung der Welt-Zustands-Paare des *bdi* Modells in die Menge der Situationen in \mathcal{M}_h ist. Sei φ eine atomare Formel.

$$\begin{aligned} \mathcal{M}_e, (w, s) \models \varphi & \text{ gdw. } (w, s) \in v(\varphi) \\ & \text{ gdw. } \iota(w, s) \in v'(\varphi) \\ & \text{ gdw. } \mathcal{M}_h, \iota(w, s) \models \varphi. \end{aligned}$$

Im Falle einer Negation oder Konjunktion ist der Induktionsschritt trivial.

Sei $\varphi = \alpha \text{ bel} : \psi$, dieselbe Argumentation trifft auf $\varphi = \alpha \text{ des} : \psi$ zu, und $z = (w, s)$:

$$\begin{aligned} \mathcal{M}_e, z \models \alpha \text{ bel} : \psi & \text{ gdw. für alle } w' \in W \text{ mit } w' \in B_s^w(\alpha) : \mathcal{M}_e, (w', s) \models \psi \\ & \text{ gdw. für alle } z' = (w', s) \in U \text{ mit } U \in B'(\alpha, \iota(w, s)) : \\ & \quad \mathcal{M}_h, \iota(w', s) \models \psi \\ & \text{ gdw. } \mathcal{M}_h, \iota(z) \models \alpha \text{ bel} : \psi \end{aligned}$$

Da in einem *bdi-stit* Modell der Operator *int* anders interpretiert wird, wird der Fall $\varphi = \alpha \text{ int} : \psi$ getrennt betrachtet.

$$\begin{aligned} \mathcal{M}_e, (w, s) \models \alpha \text{ int} : \psi & \text{ gdw. für alle } (w', s) \text{ mit } (w, s, w') \in I(\alpha) : \\ & \quad \mathcal{M}_e, (w', s) \models \psi \\ & \text{ gdw. für alle } (w', s) \text{ mit } \iota(w', s) \in I'(\alpha, \iota(w, s)) : \\ & \quad \mathcal{M}_h, \iota(w', s) \models \psi \\ & \text{ gdw. } \mathcal{M}_h, \iota(w, s) \models \alpha \text{ int} : \psi \end{aligned}$$

□

Nachdem in diesem Kapitel die *bdi-stit* Logik definiert wurde und diese verschiedene bekannte und oft diskutierte Logiken generalisiert, werde ich im folgenden Kapitel die *bdi-stit* Logik näher untersuchen und spezifizieren. Zum einen wird gezeigt werden, dass diese eben eingeführte *bdi-stit* Logik vollständig und entscheidbar ist. Zum anderen wird die bereits begonnene Generalisierung der *BDICTL** durch Einführung von temporalen Operatoren vervollständigt. Und zum dritten steht die Umsetzung der in Kapitel 4 motivierten Interaktionsaxiome in einer formalen Sprache aus. Diesen Themen ist Kapitel 6 gewidmet.

6 Vollständigkeit und Erweiterungen von *bdi-stit*

Von der deliberativen Stit-Logik mit endlich vielen einzelnen Akteuren ist gezeigt worden, dass sie endlich vollständig axiomatisierbar und entscheidbar ist. Desgleichen gilt für die *BDICTL**. Da es sich bei der *bdi-stit* Logik um eine Erweiterung der Stit-Theorie um klassische Modaloperatoren und normale Operatoren handelt, lässt sich ebenfalls eine endliche, vollständige Axiomatisierung angeben und die Entscheidbarkeit nachweisen.

Eine vollständige Axiomatisierung für die *bdi-stit* Logik wird in Abschnitt 6.1 angegeben. Es wird gezeigt, dass sie zudem die endliche Modelleigenschaft besitzt und somit entscheidbar ist. In Abschnitt 6.2 wird zur besseren Handhabung der Logik ein Tableaurekalkül angegeben. Die Vollständigkeit des Kalküls wird nachgewiesen.

In Abschnitt 6.3 werden temporale Operatoren in *bdi-stit* eingeführt und kurz diskutiert. Es wird darauf eingegangen, dass eine Erweiterung um temporale Operatoren existiert, so dass *bdi-stit* Logik auf einer geeigneten BT Struktur mit abzählbar unendlicher Momentmenge eine Generalisierung der *BDICTL** darstellt. Somit ist dies die Verallgemeinerung zu Theorem 9 aus Abschnitt 5.3. Dort wurde bereits festgehalten, dass *bdi-stit* Logik eine Generalisierung der *BDICTL** für Zustandsformeln darstellt. In Abschnitt 6.3 wird diese Behauptung erweitert auf die Sprache, die zudem temporale und Pfadoperatoren enthält.

Im Anschluss daran wird in Abschnitt 6.4 auf die Umsetzung der in den Abschnitten 4.2 und 4.3 motivierten Interaktionen zwischen dem Wunsch-, Intensions- und Handlungsoperator eingegangen. Dazu wird ein spezifizierter Intensionsoperator eingeführt, der die Interaktionen mit den anderen mentalen Zuständen und Handlungen eines Akteurs geeignet repräsentiert. Die entstandene Logik wird durch eine Erweiterung der Axiome aus Abschnitt 6.1 vollständig axiomatisiert. Für die Konstruktion des kanonischen Modells wird das kanonische Modell der *bdi-stit* Logik ohne Interaktionsaxiome herangezogen und angepasst, so dass gezeigt wird, dass die entstandene Logik *bdi-stit_{dint}* mit adäquatem Intensionsoperator vollständig axiomatisiert ist.

6.1 Vollständige Axiomatisierung der *bdi-stit*

Die Operatoren \diamond , *des*, und *bel* werden in einer monotonen Nachbarschaftssemantik interpretiert, vgl. [39]. In dieser Semantik ist es möglich, die Erfüllbarkeit von Formeln der folgenden Gestalt zu definieren. Sei *op* ein solcher Operator, dann ist $op\varphi \wedge op\neg\varphi$ erfüllbar.¹ Weil lediglich Nachbarschaften zugelassen sind, die nicht leer sind, gilt dies nur, falls φ erfüllbar ist.

Die allgemeine Nachbarschaftssemantik für einen Operator ist wie folgt definiert. Es ist ein Rahmen $\mathcal{F} = (S, N)$ gegeben, wobei S eine Menge von Situationen und N eine Funktion $N : S \rightarrow \mathcal{P}(\mathcal{P}(S))$ sei. Ein Modell besteht aus einem Paar $\mathcal{M} = (\mathcal{F}, v)$, wobei \mathcal{F} der Rahmen und v eine Bewertungsfunktion der atomaren Formeln und der Akteursvariablen ist. Die Erfüllbarkeit einer Formel $op\varphi$ in einer Situation eines Modells \mathcal{M} ist durch

$$\mathcal{M}, s \models op\varphi \text{ gdw. } \|\varphi\| \in N(s)$$

definiert. Diese Definition einer Semantik geht zurück auf D. Scott [121] und R. Montague [100], die diese sogenannten klassischen Operatoren unabhängig voneinander vorstellten. Solche Modelle werden daher Scott-Montague Modelle aber auch minimale Modelle genannt. Eine mögliche Axiomatisierung für verschiedene Logiken, die auf einer solchen Nachbarschaftssemantik aufbauen, findet sich in [39].

In der *bdi-stit* Logik wird eine bestimmte Form dieser Nachbarschaftssemantik verwendet. Fordert man, dass es sich bei den verwendeten Modellen, um *supplementierte* Modelle bzw. Rahmen handelt [39, S.215], d.h. dass für alle Situationen und Nachbarschaften $s \in S$ und $U \in N(s)$ gilt, dass jede Obermenge V von U in $N(s)$ ist, so kann gezeigt werden, dass diese *supplementierten* Modelle mit obiger Wahrheitsbedingung genau dieselben Formeln erfüllen, wie ein beliebiges Scott-Montague Modell \mathcal{M} mit der Wahrheitsbedingung $\mathcal{M}, s \models op\varphi$ gdw. ein $U \in N(s)$ mit $U \subseteq \|\varphi\|$. Die erhaltene Logik wird als monoton bezeichnet, da sie folgendes Monotonie-Axiom $op(\varphi \wedge \psi) \supset (op\varphi \wedge op\psi)$ erfüllt. An Stelle des Axioms lässt sich auch die Regel der Monotonie (*RM*) einführen: Falls $\varphi \supset \psi$ beweisbar, so ist $op\varphi \supset op\psi$ beweisbar. Eine Axiomatisierung der Operatoren \diamond , αdes :, und αbel : in *bdi-stit* lautet dann wie folgt. Für die Motivation der Axiome für den Überzeugungsoperator, siehe Abschnitt 4.1, und für den Wunschoperator, siehe Abschnitt 4.2.

¹Die Operatoren *des*, *bel* und *int* benötigen als Argument zudem eine Akteursvariable. Der Einfachheit in der Notation halber sei diese bei einem Ausdruck $op\varphi$ implizit angenommen.

6.1. Vollständige Axiomatisierung der *bdi-stit*

(F, F_b, F_d)	$\neg \diamond \perp, \neg \alpha \text{ bel} : \perp, \neg \alpha \text{ des} : \perp$
(RM)	$(\varphi \supset \psi) / (\diamond \varphi \supset \diamond \psi)$
(RM_b)	$(\varphi \supset \psi) / (\alpha \text{ bel} : \varphi \supset \alpha \text{ bel} : \psi)$
(RM_d)	$(\varphi \supset \psi) / (\alpha \text{ des} : \varphi \supset \alpha \text{ des} : \psi)$

An dieser Stelle werden noch kein spezifizierter Intentionsoperator und noch keine Interaktionen mit den Stit-Operatoren betrachtet. Für den Moment wird der *int*-Operator durch die folgenden Axiome und die Monotonieregel axiomatisiert. Dies ist äquivalent zu einer *KD*-Axiomatisierung in *BDICTL**. Die in Abschnitt 4.2 und 4.3 motivierten Axiome erfüllt dann der in Abschnitt 6.4 zu spezifizierende *dint*-Operator.

(A_i)	$(\alpha \text{ int} : \varphi \wedge \alpha \text{ int} : \psi) \supset \alpha \text{ int} : (\varphi \wedge \psi)$
(F_i)	$\neg \alpha \text{ int} : \perp$
(RM_i)	$(\varphi \supset \psi) / (\alpha \text{ int} : \varphi \supset \alpha \text{ int} : \psi)$

Da ein *bdi-stit* Modell auf einem Rahmen der Stit Theorie konstruiert wird, werden die Axiome aus [18, 19, 167, 168] für die deliberative Stit-Logik übernommen. Außer den Axiomen der klassischen Aussagenlogik gelten die folgenden Axiome.

$(A1)$	$\Box \varphi \supset \varphi, \neg \Box \varphi \supset \Box \neg \Box \varphi, \Box(\varphi \supset \psi) \supset (\Box \varphi \supset \Box \psi)$
$(A2)$	$\alpha \text{ cstit} : \varphi \supset \varphi, \neg \alpha \text{ cstit} : \varphi \supset \alpha \text{ cstit} : \neg \alpha \text{ cstit} : \varphi$ $\alpha \text{ cstit} : (\varphi \supset \psi) \supset (\alpha \text{ cstit} : \varphi \supset \alpha \text{ cstit} : \psi)$
$(A3)$	$\Box \varphi \supset \alpha \text{ cstit} : \varphi$
$(A4)$	$\alpha = \alpha, (\alpha = \beta) \supset (\beta = \alpha), ((\alpha = \beta) \wedge (\beta = \gamma)) \supset (\alpha = \gamma)$
$(A5)$	$(\alpha = \beta) \supset (\varphi \supset \varphi[\alpha/\beta])^2$
(AIA_k)	$(\Delta(\beta_0, \dots, \beta_k) \wedge \diamond \beta_0 \text{ cstit} : \psi_0 \wedge \dots \wedge \diamond \beta_k \text{ cstit} : \psi_k)$ $\supset \diamond (\beta_0 \text{ cstit} : \psi_0 \wedge \dots \wedge \beta_k \text{ cstit} : \psi_k)$

²Die Substitution muss nicht vollständig vorgenommen werden. Es ist möglich, sowohl alle Vorkommen

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Die Axiome (AIA_k) garantieren die Unabhängigkeit für jeweils $k \in \mathbb{N}$ beliebige Akteure. Die Formel $\Delta(\beta_0, \dots, \beta_k)$ steht als Abkürzung dafür, dass die Akteure, die durch die Akteursvariablen β_0, \dots, β_k benannt werden, paarweise verschieden sind,

$$\Delta(\beta_0, \dots, \beta_k) = \bigwedge_{\substack{0 \leq i, j \leq k \\ i \neq j}} \neg(\beta_i = \beta_j).$$

Die Regel der Necessitation und Modus Ponens sowie das Axiom der Wahlmöglichkeiten APC_n (Axiom of possible Choices) kommen hinzu.

(RN)	$\varphi / \Box\varphi$
(MP)	$\varphi, \varphi \supset \psi / \psi$
(APC_n)	$(\Diamond \alpha \text{ cstit} : \varphi_1 \wedge \Diamond(\alpha \text{ cstit} : \varphi_2 \wedge \neg\varphi_1) \wedge \dots \wedge$ $\Diamond(\alpha \text{ cstit} : \varphi_n \wedge \neg\varphi_1 \wedge \dots \wedge \neg\varphi_{n-1})) \supset (\varphi_1 \vee \dots \vee \varphi_n)$

Das Axiom (APC_n) besagt, dass jeder Akteur in jeder Situation lediglich n verschiedene Handlungsalternativen hat. Falls das Axiom (APC_n) für ein beliebiges, aber festes n in die Axiomatisierung aufgenommen wurde, wird die erhaltene Sprache mit \mathcal{L}_n bezeichnet. Offensichtlich gilt für $n \geq 1$, dass $\mathcal{L}_{n+1} \subseteq \mathcal{L}_n$.

Anhand dieser Axiome und Regeln lassen sich die folgenden Theoreme ableiten, die im Weiteren noch nützlich sein sollen.

(Ni)	$\alpha \text{ int} : \top$
(Tc)	$\alpha \text{ cstit} : (\varphi \wedge \psi) \equiv (\alpha \text{ cstit} : \varphi \wedge \alpha \text{ cstit} : \psi)$
(Ti)	$\alpha \text{ int} : (\varphi \wedge \psi) \equiv (\alpha \text{ int} : \varphi \wedge \alpha \text{ int} : \psi)$ $(\alpha \text{ int} : \varphi \vee \alpha \text{ int} : \psi) \supset \alpha \text{ int} : (\varphi \vee \psi)$
(T)	$\Diamond(\varphi \wedge \psi) \supset (\Diamond\varphi \wedge \Diamond\psi)$
(Tb)	$\alpha \text{ bel} : (\varphi \wedge \psi) \supset (\alpha \text{ bel} : \varphi \wedge \alpha \text{ bel} : \psi)$
(Td)	$\alpha \text{ des} : (\varphi \wedge \psi) \supset (\alpha \text{ des} : \varphi \wedge \alpha \text{ des} : \psi)$

oder auch nur einige Vorkommen von α mit β zu ersetzen.

Konstruktion des kanonischen Rahmens und Modells der *bdi-stit* Logik

Zuerst wird ein Vollständigkeitsbeweis präsentiert, indem die in [19] definierten kanonischen Rahmen für die deliberative Stit-Theorie erweitert werden, um einen Rahmen für ein kanonisches *bdi-stit* Modell zu erhalten. Es wird sich zeigen, dass jede widerspruchsfreie *bdi-stit* Formelmengung erfüllbar in einem solchen kanonischen Modell ist. Im zweiten Teil wird analog zum Vorgehen für die deliberative Stit-Theorie die endliche Modelleigenschaft nachgewiesen: Jede falsifizierbare Formel kann in einem endlichen Modell falsifiziert werden. Aus der endlichen Axiomatisierbarkeit und der Falsifizierbarkeit in endlichen Modellen resultiert die Entscheidbarkeit.³

Die Korrektheit des Systems von Axiomen und Regeln ist offensichtlich und für die Axiomenschemata (APC_n) und (AIA_k) in [19] gezeigt worden. Daher beschäftigt sich dieser Abschnitt lediglich mit dem Nachweis der Vollständigkeit. Dafür wird ein *bdi-stit* Modell konstruiert, welches ein kanonisches Modell der *dstit* Logik, wie in [19] vorgeschlagen, mit kanonischen Modelleigenschaften kombiniert, wie sie für Modelle klassischer monotoner Modallogik Operatoren üblich sind, vgl. [39].

Da ein kanonisches *dstit* Modell die komplexere Struktur besitzt, bildet es die Basis für ein kanonisches *bdi-stit* Modell, die geeignet erweitert wird, um Überzeugungs-, Wunsch- und Intentionoperatoren zu interpretieren. Die Konstruktion einer BT + AC Struktur wurde von Xu in [19, 167, 168] definiert. Diese wird für das kanonische *bdi-stit*-Modell aufgegriffen. Zuvor werden folgende Begriffe und Relationen definiert. Die Menge $W_{\mathcal{L}_n}$ sei die Menge aller maximalen \mathcal{L}_n konsistenten Formelmengen. Die folgenden Relationen und Teilmengen sind auf dieselbe Art und Weise definierbar wie für die *dstit* Logik Ldm_n , wobei der Index n kennzeichnet, dass das Axiom APC_n zu den Axiomenschemata hinzugenommen wird.

Zu einer gegebenen Teilmenge $W \subseteq W_{\mathcal{L}_n}$ sei die Relation \cong_W auf der Menge der Akteursvariablen durch $\alpha_1 \cong_W \alpha_2$ gdw. $\alpha_1 = \alpha_2 \in w$ für alle $w \in W$ definiert.

Lemma 10

1. Die Relation \cong_W ist eine Äquivalenzrelation für jede Teilmenge $W \subseteq W_{\mathcal{L}_n}$.

BEWEIS: Vgl. [19]. Dass es sich hierbei tatsächlich um eine Äquivalenzrelation handelt, ist auf die Axiome (A4) zurückzuführen, die Reflexivität, Symmetrie und Transitivität widerspiegeln. \square

Ebenso ist es möglich, auf jeder beliebigen Teilmenge $X \subseteq W_{\mathcal{L}_n}$ eine Relation durch eine ebenfalls beliebige Menge an Akteursvariablen A zu definieren:

³In [127] ist der Vollständigkeits- und Endlichkeitsnachweis der *bdi-stit* Logik veröffentlicht.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

$w \cong_A w'$ genau dann, wenn für alle $\alpha_1, \alpha_2 \in A$ gilt, $\alpha_1 = \alpha_2 \in w$ gdw. $\alpha_1 = \alpha_2 \in w'$.

Lemma 10 (Fortsetzung)

2. Bei der Relation \cong_A handelt es sich für jedes $X \subseteq W_{\mathcal{L}_n}$ und eine beliebige Menge von Akteursvariablen A um eine Äquivalenzrelation.

BEWEIS: Dies ist offensichtlich anhand der Definition. \square

Für die nächsten beiden Aussagen wird die Menge $W \subseteq W_{\mathcal{L}_n}$ fixiert, um eine Relation $R \subseteq W \times W$ wie folgt zu definieren: wRw' gdw. $\{\varphi \mid \Box\varphi \in w\} \subseteq w'$.

Lemma 10 (Fortsetzung)

3. Für jede Teilmenge $W \subseteq W_{\mathcal{L}_n}$ ist die so definierte Relation $R \subseteq W \times W$ eine Äquivalenzrelation.

BEWEIS: Vgl. [19]. Analog zu Lemma 10.1. resultiert die Reflexivität und Euklidizität aus den Axiomen (A1). \square

Folglich kann jede Teilmenge W vollständig in Äquivalenzklassen $\{X_i\}_{i \in I}$ hinsichtlich der Relation R eingeteilt werden. Sei $X \in \{X_i\}_{i \in I}$. Falls $(\alpha = \beta) \in w$ für ein $w \in X$, dann ist $(\alpha = \beta)$ auch in allen $w' \in X$.⁴ Dies erleichtert die Schreibweise, so dass anstatt der einzelnen Akteursvariablen die Repräsentanten $\{\beta_j\}_{j \in J} = \{[\alpha]_X\}$ der Äquivalenzrelation \equiv_X verwendet werden, um die folgende Relation R_{β_j} auf X für jedes $j \in J$ zu definieren: $wR_{\beta_j}w'$ gdw. $\{\varphi \mid \beta_j cstit : \varphi \in w\} \subseteq w'$.⁵

Lemma 10 (Fortsetzung)

4. Sei $X \in \{X_i\}_{i \in I}$. Dann ist die Relation $R_{\beta_j} \subseteq X \times X$ eine Äquivalenzrelation für jedes $j \in J$.

BEWEIS: Vgl. [19]. Die Axiome (A2) geben die Reflexivität und Euklidizität an. \square

Die Definition von R_{β_j} hängt offensichtlich von der Wahl der Menge X ab. Daher wird für die folgenden Betrachtungen X fest gewählt. Es sei gesagt, dass dies für jede Äquivalenzklasse von R gilt. Mit E_{β_j} werden die Äquivalenzklassen zu gegebener Relation R_{β_j} auf X bezeichnet.

Lemma 11 Seien $X, \{\beta_j\}_{j \in J}, R, R_{\beta_j}, E_{\beta_j}$ für alle $j \in \mathbb{N}$ entsprechend der vorangegangenen Definitionen gegeben, dann gilt für alle $w \in X$ und $\varphi \in \mathcal{L}_n$:

⁴Sei $\alpha = \beta \in w$. Dann gilt mit (A4), $\alpha = \alpha \in w$, (RN), $\Box(\alpha = \alpha) \in w$, (A5), $(\alpha = \beta \supset (\Box(\alpha = \alpha) \supset \Box(\alpha = \beta))) \in w$, (MP), $\Box(\alpha = \beta) \in w$.

⁵Die Abkürzung $\beta_j cstit : \varphi$ bezeichnet, dass für ein $\alpha \in \beta_j$, $\alpha cstit : \varphi \in w$. Daher gilt dies mit Axiom (A5) für alle $\tilde{\alpha} \in \beta_j$, d.h., $\tilde{\alpha} cstit : \varphi \in w$.

6.1. Vollständige Axiomatisierung der *bdi-stit*

- (i) $\Box\varphi \in w$ *gdw.* $\varphi \in w'$ für alle $w' \in X$
 gdw. $\Box\varphi \in w'$ für alle $w' \in X$
- (ii) $\beta_j \text{ cstit} : \varphi \in w$ *gdw.* $\varphi \in w'$ für alle w' mit $wR_{\beta_j}w'$
 gdw. $\beta_j \text{ cstit} : \varphi \in w'$ für alle w' mit $wR_{\beta_j}w'$
- (iii) $\beta_j \text{ dstit} : \varphi \in w$ *gdw.* $\varphi \in w'$ für alle w' mit $wR_{\beta_j}w'$
 und $\neg\varphi \in w''$ für einige $w'' \in X$
- (iv) Für eine beliebige Funktion f von $\{\beta_j\}_{j \in J}$ in die Menge aller Äquivalenzklassen $\{E_{\beta_j}\}_{j \in J}$ unter der Bedingung $f(\beta_j) \in E_{\beta_j}$ gilt:

$$\bigcap_{j \in J} f(\beta_j) \neq \emptyset.$$

- (v) Sei $n \geq 1$, dann gibt es höchstens n verschiedene Äquivalenzklassen zu R_{β_j} für jedes $j \in J$, d.h.,

$$|E_{\beta_j}| \leq n.$$

BEWEIS: Vgl. [19]. Punkt (i) folgt aus den Axiomen (A1) und Regel (RN). Punkt (ii) gebraucht die Axiome (A2) und (A3) sowie (RN). Behauptung (iii) resultiert aus der Definition des *dstit*-Operators und der Anwendung von (i) und (ii). Offensichtlich ist, dass (iv) durch (AIA_k) und (v) durch (APC_n) gerechtfertigt werden. Als Beispiel sei (i) angeführt, $\Box\varphi \in w$ *gdw.* $\Box\varphi \in w'$ für alle w' , da R transitiv and reflexiv ist. Falls $\Box\varphi \in w$, dann $\varphi \in w'$ für alle $w' \in X$ gemäß Definition der Relation R . Für die Rückrichtung nehme man an, dass es ein $w'' \in X$ mit $\Box\varphi \notin w''$ gibt, so dass $\neg\Box\varphi \in w''$. Da w'' maximal ist, folgt mit (A1), dass $\neg\Box\varphi \in w'$ für alle w' . Zudem ist R eine Äquivalenzrelation, so dass weiterhin $\varphi \supset \Box\neg\Box\neg\varphi$ in w' aufgrund der Symmetrie von R ist. Falls $\varphi \in w'$, ist aufgrund der Reflexivität auch $\neg\Box\neg\varphi \in w'$. Folglich ist $\neg(\Box\neg\varphi \vee \Box\varphi) \in w'$. Aufgrund des Theorems $(\Box\varphi \vee \Box\psi) \supset \Box(\varphi \vee \psi)$ ⁶ gilt, dass $\neg\Box(\varphi \vee \neg\varphi) \in w'$. Dies ist widersprüchlich, da $\Box(\varphi \vee \neg\varphi)$ gemäß (RN) in jeder maximal konsistenten Menge ist. Somit ist die Annahme falsch. Für alle w'' gilt, dass $\Box\varphi \in w''$ für alle $w'' \in X$.

□

Theorem 12 (Vollständigkeit) Jede \mathcal{L}_n -widerspruchsfreie Menge Ψ von *bdi-stit* Formeln ist in einem *bdi-stit* Modell erfüllbar.

⁶Formel $\varphi \supset (\varphi \vee \psi)$ ist aus den Axiomen A1 und (RN) $\Box(\varphi \supset (\varphi \vee \psi))$ ableitbar und somit in jeder maximal konsistenten Menge enthalten.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

BEWEIS: Sei $W_{\mathcal{L}_n}$ die Menge aller maximalen, bezüglich \mathcal{L}_n widerspruchsfreien Formelmengen. Sei $A = \{\alpha \mid \text{Akteursvariable } \alpha \text{ kommt in } \psi \text{ vor, wobei } \psi \in \Psi\}$. Sei \cong_A die nach Lemma 10.2. definierte Äquivalenzrelation auf $W_{\mathcal{L}_n}$. Mit W wird die Äquivalenzklasse bezeichnet, so dass für alle Akteursvariablen gilt, falls $\beta, \tilde{\beta} \notin A$ und $\alpha, \tilde{\alpha} \in A$, dann $\alpha = \tilde{\alpha} \notin \Psi$ gdw. $\alpha = \tilde{\alpha} \notin w$ und $\beta = \tilde{\beta} \in w$ und $\alpha = \beta \notin w$ für alle $w \in W$. Sei weiterhin $\{X_i\}_{i \in I}$ die Menge aller Äquivalenzklassen der Relation R auf W , vgl. Lemma 10.3.

Das Fundament des Modells ist sein Rahmen $\mathcal{F} = (Tree, \leq, \mathcal{A}, N, C, B, D, I)$, der im Fall des *bdi-stit* Modells auf einer BT Struktur $(Tree, \leq)$ definiert ist. Er wird wie folgt gewählt:

- $Tree := \{w \mid w \in W\} \cup \{X_i \mid i \in I\} \cup \{W\}$
- $\leq := rtrcl^7 \left(\{(W, X_i), (X_i, w) \mid w \in X_i, i \in I\} \right)$
- $\mathcal{A} := \{\alpha \mid \alpha \text{ gehört zu einer beliebigen, festgewählten Menge von Repräsentanten der Relation } \cong_W \text{ auf allen Akteursvariablen.}\}^8$
- Für alle $i \in I$ wird die Wahlzelle jedes Akteurs $\alpha \in \mathcal{A}$ im Moment $m \in Tree$ durch das Folgende festgelegt.
 - $C(\alpha, w) := \{h_w\}$, wobei h_w der einzige Geschichtsverlauf durch den Moment w mit $h_w = \{w, X_i, W\}$ ist, wobei X_i die Äquivalenzklasse ist, zu der w gehört.
 - $C(\alpha, W) := \{h_w \mid w \in W\}$
 - Gemäß Lemma 10.4. gibt es eine Äquivalenzklasse β_j mit $\alpha \in \beta_j$ und eine Äquivalenzrelation $R_{\beta_j}^i$ auf X_i . Die Klassen von $R_{\beta_j}^i$ auf X_i werden mit $E_{\beta_j}^i$ bezeichnet, so dass $C(\alpha, X_i) := \{H \mid \exists e : e \in E_{\beta_j}^i \text{ und } H = \{h_w \mid w \in e\}\}$.

Da eine eindeutige Abbildung zwischen den Mengen $w \in W$ und allen Geschichtsverläufen in $(Tree, \leq)$ existiert, sind die folgenden Mengen wohldefiniert. Für alle $m \in Tree$, $w \in W$ und $\alpha \in \mathcal{A}$ wird Folgendes festgelegt.

- $|\varphi| := \bigcup_{i \in I} \{(X_i, h_{w'}) \mid \varphi \in w', X_i \in h_{w'}\} \in N_{(m, h_w)}$ gdw. $\diamond \varphi \in w$

⁷Die Abkürzung *rtrcl* steht für den reflexiven und transitiven Abschluss der Relation.

⁸Sowohl für die Variablen als auch für die Akteure werden der Übersichtlichkeit halber dieselben griechischen Buchstaben α, β, \dots verwendet. Da im Folgenden die Akteursvariablen durch sich selbst interpretiert werden, scheint dies nur sinnvoll. Im Allgemeinen gilt jedoch $|\mathcal{A}| \neq |A|$.

6.1. Vollständige Axiomatisierung der bdi-stit

- $|\varphi| \in B_{(m,h_w)}^\alpha$ gdw. $\alpha \text{ bel} : \varphi \in w^9$
- $|\varphi| \in D_{(m,h_w)}^\alpha$ gdw. $\alpha \text{ des} : \varphi \in w^9$
- Für die Intentionsrelation sei auf W die Relation $S_\alpha \subseteq W \times W$ für alle $\alpha \in \mathcal{A}$ mit $wS_\alpha w'$ gdw. $\{\varphi \mid \alpha \text{ int} : \varphi \in w\} \subseteq w'$ definiert. Die Mengen werden I_s^α für jede Situation $s = (m, h_w)$ mit $I_{(m,h_w)}^\alpha = \{(w', h_{w'}) \mid wS_\alpha w'\}$ festgelegt. Für alle $w \in W$ gibt es ein $w' \in W$ mit $wS_\alpha w'$, so dass $I_{(m,h_w)}^\alpha \neq \emptyset$ für beliebiges $m \in \text{Tree}$. Anhand des Axioms (D_i) , der Ableitungsregel (RM_i) und des Theorems (T_i) ist offensichtlich, dass für jedes $w \in W$ die Menge $S = \{\varphi \mid \alpha \text{ int} : \varphi \in w\}$ widerspruchsfrei ist. Folglich gibt es eine maximale, widerspruchsfreie Menge w' mit $S \subseteq w'$.

In Analogie zu [19] erfüllt der Rahmen \mathcal{F} die Bedingung der Akteursunabhängigkeit. Zu gegebenem Moment $m \in \text{Tree}$ sei Select_m die Menge aller Funktionen von \mathcal{A} in die Teilmengen von H_m , der Menge der Geschichten, die durch den Moment m verlaufen, wobei gelte, dass $\sigma(\alpha) \in C_m^\alpha$ für alle $\sigma \in \text{Select}_m$. So erfüllt der Rahmen \mathcal{F} diese Bedingung genau dann, wenn für jeden Moment m und jedes $\sigma \in \text{Select}_m$ gilt

$$\bigcap_{\alpha \in \mathcal{A}} \sigma(\alpha) \neq \emptyset.$$

Für $m = w$ und jedes beliebige $w \in W$ oder $m = W$ ist diese Bedingung offensichtlich erfüllt. Sei $i \in I$ beliebig, σ_{X_i} eine beliebige Funktion von \mathcal{A} in $\mathcal{P}(H_{X_i})$ mit $\sigma_{X_i}(\alpha) \in C_{X_i}^\alpha$ für alle $\alpha \in \mathcal{A}$, dann gibt es aufgrund der obigen Definition von $C_{X_i}^\alpha$ eine Äquivalenzklasse $e_j \in E_{\beta_j}^i$ mit $\sigma_{X_i}(\alpha) = \{h_w \mid w \in e_j\}$. Die Funktion f_i wird durch $f_i(\alpha) = e_j \in E_{\beta_j}^i$ festgelegt, wobei $\alpha \in \beta_j$ für $j \in J$. Dann gilt für alle $\alpha, \tilde{\alpha} \in \beta_j$, dass $f_i(\alpha) = f_i(\tilde{\alpha})$, so dass es eine wohldefinierte Funktion $\tilde{f}_i : \{\beta_j\}_{j \in J} \rightarrow \bigcup_{j \in J} E_{\beta_j}^i$ gibt. Außerdem gilt $w \in f_i(\alpha)$ gdw. $h_w \in \sigma_{X_i}(\alpha)$ und nach Lemma 11. (iv), dass

$$\bigcap_{\alpha \in \mathcal{A}} f_i(\alpha) = \bigcap_{j \in J} \tilde{f}_i(\beta_j) \neq \emptyset, \quad \text{so dass} \quad \bigcap_{\alpha \in \mathcal{A}} \sigma_{X_i}(\alpha) \neq \emptyset.$$

Da $|C_w^\alpha| = |C_W^\alpha| = 1$ für alle $w \in W$, $\alpha \in \mathcal{A}$ und $|C_{X_i}^\alpha| = |E_{\beta_j}^i|$ für alle $i \in I$, folgt mit Lemma 11. (v), dass $|C_m^\alpha| \leq n$ für jedes $\alpha \in \mathcal{A}$ und $m \in \text{Tree}$, vgl. [19]. Folglich hat jeder Akteur α höchstens n mögliche Wahlzellen in einem solchen Rahmen \mathcal{F} .

Um vom Rahmen \mathcal{F} zum Modell $\mathcal{M} = (\mathcal{F}, v)$ überzugehen, bedarf es einer Interpretationsfunktion v . Diese bildet jede Akteursvariable β aus A in \mathcal{A} ab, $v(\beta) = \alpha \in \mathcal{A}$,

⁹Aufgrund des Axioms (A5) folgt für alle $\beta \in [\alpha]_W$, $\beta \text{ op} : \varphi \in w$ gdw. $\alpha \text{ op} : \varphi \in w$ für $\text{op} \in \{\text{cstit}, \text{dstit}, \text{bel}, \text{des}, \text{int}\}$ und für alle $w \in W$.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

falls $\beta \in [\alpha]_w$, und ordnet jeder atomaren Formel eine Teilmenge von Situationen zu, so dass $s = (m, h_w) \in v(p)$ gdw. $p \in w$ für alle $m \in Tree$. Offensichtlich ist dann $v(\top) = S$ und $v(\perp) = \emptyset$. Für eine beliebige Akteursvariable, beliebige Geschichten $h \in H$ und $w \in W$ gilt, dass $h \in C_{X_i}^\alpha(h_w)$ gdw. ein $e \in E_{[\alpha]_{X_i}}^i$ existiert und $w, w' \in e$, wobei $h = h_{w'}$. Außerdem gilt, falls $w, w' \in e \in E_{[\alpha]_{X_i}}^i$, dann $wR_{[\alpha]_{X_i}}^i w'$ nach Definition, so dass $h \in C_{X_i}^\alpha(h_w)$ gdw. $wR_{[\alpha]_{X_i}}^i w'$.

Durch Induktion über den Formelaufbau wird gezeigt, dass

$$(X_i, h_w) \models \varphi \quad \text{gdw.} \quad \varphi \in w$$

für jede *bdi-stit* Formel $\varphi \in \Psi$ und für alle $w \in X_i, i \in I$.¹⁰

$$(X_i, h_w) \models p \quad \text{gdw.} \quad (X_i, h_w) \in v(p) \quad \text{gdw.} \quad p \in w.$$

$$(X_i, h_w) \models (\alpha = \beta) \quad \text{gdw.} \quad v(\alpha) = v(\beta) \quad \text{gdw.} \quad \alpha \cong_w \beta \quad \text{gdw.} \quad \alpha = \beta \in w.$$

$$(X_i, h_w) \models \neg\varphi \quad \text{gdw.} \quad (X_i, h_w) \not\models \varphi \quad \text{gdw.} \quad \varphi \notin w \quad \text{gdw.} \quad \neg\varphi \in w.$$

$$(X_i, h_w) \models \varphi \wedge \psi \quad \text{gdw.} \quad (X_i, h_w) \models \varphi \quad \text{und} \quad (X_i, h_w) \models \psi \\ \text{gdw.} \quad \varphi \in w \quad \text{und} \quad \psi \in w \quad \text{gdw.} \quad (\varphi \wedge \psi) \in w.$$

$$(X_i, h_w) \models \Box\varphi \quad \text{gdw.} \quad \text{für alle } h \in H_{X_i}: (X_i, h) \models \varphi \\ \text{gdw.} \quad \text{für alle } h \in H_{X_i} \text{ existiert ein } w' \in X_i: \\ h = h_{w'} \quad \text{und} \quad \varphi \in w' \\ \text{gdw.} \quad \text{für alle } w' \in X_i: \varphi \in w'$$

$$\text{Lemma 11. (i)} \quad \text{gdw.} \quad \Box\varphi \in w.$$

$$(X_i, h_w) \models \beta \text{ cstit} : \varphi \quad \text{gdw.} \quad \beta \in [\alpha]_w, \text{ für alle } h \in C_{X_i}^\alpha(h_w) : (X_i, h) \models \varphi \\ \text{gdw.} \quad \text{für alle } w' \in X_i : \text{falls } wR_{[\alpha]_{X_i}}^i w', \text{ dann } \varphi \in w'$$

$$\text{Lemma 11. (ii)} \quad \text{gdw.} \quad \alpha \text{ cstit} : \varphi \in w$$

$$\text{Axiom (A5)} \quad \text{gdw.} \quad \beta \text{ cstit} : \varphi \in w.$$

$$(X_i, h_w) \models \Diamond\varphi \quad \text{gdw.} \quad \text{es gibt } U \in N_{(X_i, h_w)} \text{ mit } \emptyset \neq U \subseteq \|\varphi\| \\ \text{gdw.} \quad \text{es gibt } \psi \text{ mit } \emptyset \neq \|\psi\| \subseteq \|\varphi\| \quad \text{und} \quad \Diamond\psi \in w$$

$$(*) \quad \text{gdw.} \quad \Diamond\varphi \in w.$$

Es wird (*) gezeigt. Falls $\Diamond\varphi \in w$, dann $|\varphi| \in N_{(X_i, h_w)}$ mit $w \in X_i$. Es gibt eine Formel nämlich $\psi = \varphi$ und $|\varphi| \subseteq \|\varphi\|$ mit Induktion. Angenommen es gibt ψ mit $\emptyset \neq \|\psi\| \subseteq \|\varphi\|$ und $\Diamond\psi \in w$. Da $|\psi| = \bigcup\{(X_i, h_{w'}) \mid \psi \in w', X_i \in h_{w'}, i \in I\}, W =$

¹⁰Im Unterschied zu dem Beweis von Xu in [19, 167, 168] wird aufgrund der mentalen Zustandsoperatoren die Aussage für alle Äquivalenzklassen benötigt.

6.1. Vollständige Axiomatisierung der bdi-stit

$\bigcup\{w \mid (X_i, h_w) \in S, i \in I\}$ und $|\psi| \subseteq \|\varphi\|$, gilt für alle $w' \in W, i \in I$: Falls $(X_i, h_{w'}) \models \psi$, dann $(X_i, h_{w'}) \models \varphi$. Nach Induktion gilt für alle $w' \in W$, falls $\psi \in w'$, dann $\varphi \in w'$. Folglich $(\psi \supset \varphi) \in w'$. Nach Anwendung der Regel (RM) gilt, dass $(\diamond\psi \supset \diamond\varphi) \in w'$ für alle $w' \in W$. Da $\diamond\psi \in w$, also auch $\diamond\varphi \in w$.

$$\begin{aligned} (X_i, h_w) \models \beta bel : \varphi & \text{ gdw. es gibt } U \in B_{(X_i, h_w)}^\alpha, \beta \in [\alpha]_W \text{ und } U \subseteq \|\varphi\| \\ & \text{gdw. es gibt } \psi \text{ mit } |\psi| \subseteq \|\varphi\| \text{ und } \alpha bel : \psi \in w \\ (**) & \text{ gdw. } \beta bel : \varphi \in w. \end{aligned}$$

$$\begin{aligned} (X_i, h_w) \models \beta des : \varphi & \text{ gdw. es gibt } U \in D_{(X_i, h_w)}^\alpha, \beta \in [\alpha]_W \text{ und } U \subseteq \|\varphi\| \\ & \text{gdw. es gibt } \psi \text{ mit } |\psi| \subseteq \|\varphi\| \text{ und } \alpha des : \psi \in w \\ (**) & \text{ gdw. } \beta des : \varphi \in w. \end{aligned}$$

Auf dieselbe Weise wie für den *bel*-Operator wird (**) für den *des*-Operator gezeigt. Falls $\beta bel : \varphi \in w$, dann mit Axiom (A5) $\alpha bel : \varphi \in w$, so dass $|\varphi| \in B_{(X_i, h_w)}^\alpha$ mit $w \in X_i$. Dann existiert $\psi = \varphi$, für das gilt, $|\varphi| \subseteq \|\varphi\|$ nach Induktion. Es gibt ψ mit $\emptyset \neq |\psi| \subseteq \|\varphi\|$ und $\alpha bel : \psi \in w$. Da $|\psi| = \bigcup\{(X_i, h_{w'}), i \in I \mid \psi \in w', X_i \in h_{w'}\}$, $W = \bigcup\{w \mid (X_i, h_w) \in S, i \in I\}$ und $|\psi| \subseteq \|\varphi\|$, gilt für alle $w' \in W, i \in I$: Falls $(X_i, h_{w'}) \models \psi$, dann $(X_i, h_{w'}) \models \varphi$. D.h., falls $\psi \in w'$, dann $\varphi \in w'$, also auch $(\psi \supset \varphi) \in w'$. Mit Regel (RM_b) gilt, $(\alpha bel : \psi \supset \alpha bel : \varphi) \in w'$ für alle $w' \in W$. Da $\alpha bel : \psi \in w$, also auch $\alpha bel : \varphi \in w$. Wiederum mit Axiom (A5) gilt $\beta bel : \varphi \in w$.

$$\begin{aligned} (X_i, h_w) \models \beta int : \varphi & \text{ gdw. } I_{(X_i, h_w)}^\alpha \subseteq \|\varphi\| \text{ mit } \beta \in [\alpha]_W \\ & \text{gdw. für alle } s = (w', h'_w): \text{ Falls } wS_\alpha w', \text{ dann } s \models \varphi. \\ & \text{gdw. für alle } s = (w', h'_w): \text{ Falls } wS_\alpha w', \text{ dann } \varphi \in w'. \\ (***) & \text{ gdw. } \beta int : \varphi \in w. \end{aligned}$$

Zuletzt wird (***) bewiesen. Falls $\beta int : \varphi \in w$, dann $\alpha int : \varphi \in w$ und für alle $w' \in W$ mit $wS_\alpha w'$ gilt, dass $\varphi \in w'$. Für alle $s = (w', h'_w)$ gelte, falls $wS_\alpha w'$, dann $\varphi \in w'$. Aufgrund der Maximalität von $w \in W$ und Axiom (D_i) gilt, dass entweder $\alpha int : \varphi \in w$ oder $\alpha int : \neg\varphi \in w$ oder beide Formeln nicht in w sein können. Die Annahme, dass $\alpha int : \neg\varphi \in w$, ist kontradiktorisch, da dann $\neg\varphi \in \{\psi \mid \alpha int : \psi \in w\} \subseteq w'$ für ein w' mit $wS_\alpha w'$ gelte. Die Annahme $\neg\alpha int : \varphi \in w$ und $\neg\alpha int : \neg\varphi \in w$ führt dazu, dass $\varphi, \neg\varphi \notin \{\psi \mid \alpha int : \psi \in w\}$. Da w maximal ist, ist die Menge $\{\psi \mid \alpha int : \psi \in w\}$ abgeschlossen unter Implikation aufgrund der Regel (RM_i) und Theorems (T_i), so dass $\{\varphi\} \cup \{\psi \mid \alpha int : \psi \in w\}$ und $\{\neg\varphi\} \cup \{\psi \mid \alpha int : \psi \in w\}$ beide widerspruchsfrei sind. Dann gäbe es aber eine maximale Welt w'' mit $\{\neg\varphi\} \cup \{\psi \mid \alpha int : \psi \in w\} \subseteq w''$ und $wS_\alpha w''$. Dies widerspricht jedoch $\varphi \in w'$ für alle $wS_\alpha w'$. Folglich kann nur $\alpha int : \varphi \in w$ bzw. $\beta int : \varphi \in w$ gelten.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Da es mindestens eine widerspruchsfreie Menge $w_0 \in W$ mit $\Psi \subseteq w_0$ nach Lindenbaums Lemma gibt und dieses w_0 einer Äquivalenzklasse X_{i_0} zugeordnet werden kann, erfüllt obiges kanonisches Modell \mathcal{M} jede Formel aus Ψ , denn $(X_{i_0}, h_{w_0}) \models \varphi$ gdw. $\varphi \in w_0$. \square

Die endliche Modelleigenschaft

Nachdem soeben gezeigt wurde, dass obige Axiomatisierung schwach vollständig ist, stellt sich die Frage: Ist die *bdi-stit* Logik auch entscheidbar? Falls gezeigt werden kann, dass zu jedem Satz, der kein Theorem der Sprache ist, ein endliches Modell existiert, das diesen Satz falsifiziert, d.h. dass es die Negation des Satzes erfüllt, dann ist zusammen mit der endlichen Axiomatisierung der Sprache erwiesen, dass es in endlich vielen Schritten entscheidbar ist, ob es sich bei einer vorgegebenen Formel um ein Theorem handelt oder nicht, vgl. [39].

Zu gegebenem Satz φ nach Definition 7 wird ein endlicher Rahmen \mathcal{F}_{FIN} konstruiert und eine spezielle Interpretationsfunktion ν hinzugefügt, so dass für jede Teilformel von φ (einschließlich φ) entscheidbar ist, ob der Satz in dem endlichen Modell (\mathcal{F}_{FIN}, ν) erfüllt wird oder nicht. Hierfür wird die Filtrationsmethode wie in [19, 168] für die *dstit* Logik verwendet und mit der Filtrationsmethode für klassische Modallogiken kombiniert, vgl. [39]. Als Ausgangspunkt wird der kanonische Rahmen $\mathcal{F} = (Tree, \leq, \mathcal{A}, N, C, B, D, I)$ des vorangegangenen Abschnittes genommen.

Die erste Filtration wird über alle Welten durch die Menge aller Teilformeln und deren anhand der Axiomatisierung gewonnenen abgeleiteten Formeln vollzogen. Dabei eingeschlossen sind die Formeln, die sich anhand der Axiome (AIA_k) und (Ai) für alle $1 \leq i \leq 5$ ergeben, vgl. [19]. Danach erfolgt eine weitere Filtration über die erhaltenen endlich vielen Äquivalenzklassen zur Relation R durch die Menge aller *dstit* bzw. *cstit*-Operatoren. Anhand der zweiten Filtration ist es möglich, die wahläquivalenten Geschichten zu definieren. Die notwendigen Mengen von Teilformeln für die jeweiligen Filtrationen sind wie folgt zu konstruieren.

$$\begin{aligned} \Sigma_\varphi &= \{ \psi \mid \psi \text{ Teilformel von } \varphi \} \\ \Sigma_i &= \{ \neg\beta \text{ int} : \neg\psi \mid \beta \text{ int} : \psi \in \Sigma_\varphi \} \\ \Sigma_d &= \{ \beta \text{ cstit} : \psi, \neg\Box\psi \mid \beta \text{ dstit} : \psi \in \Sigma_\varphi \} \cup \{ \neg\Box\neg\psi \mid \Diamond\psi \in \Sigma_\varphi \} \\ \Sigma_c &= \{ \beta \text{ cstit} : \neg\beta \text{ cstit} : \psi \mid \beta \text{ cstit} : \psi \in \Sigma_d \} \\ \Sigma_e &= \{ \psi \mid \psi \text{ Teilformel einer Formel aus } \Sigma_\varphi \cup \Sigma_d \cup \Sigma_c \} \end{aligned}$$

Sei $A = \{\beta_j\}_{j \in J}$ die Menge aller Akteursvariablen, die in φ vorkommen. Dann sei für

jeden Akteur β_i die folgende endliche Menge Σ_a definiert.

$$\begin{aligned}\Sigma_{\beta_i} &= \{ \psi_0 \wedge \dots \wedge \psi_m \mid 0 \leq m \leq |\Sigma_e|, 0 \leq j \leq m, \beta_i \text{ cstit} : \psi_j \in \Sigma_e \} \\ \Sigma_p &= \{ \diamond(\beta_0 \text{ cstit} : \psi_0 \wedge \dots \wedge \beta_n \text{ cstit} : \psi_n) \mid \beta_0, \dots, \beta_n \text{ paarweise verschieden,} \\ &\quad n \geq 0, \psi_i \in \Sigma_{\beta_i}, \text{ für alle } 0 \leq i \leq n \} \\ \Sigma_a &= \{ \psi \mid \psi \text{ Teilformel einer Formel aus } \Sigma_e \cup \Sigma_i \cup \Sigma_p \}\end{aligned}$$

Auf W ¹¹ ist eine Äquivalenzrelation \equiv_{Σ_a} definiert. Für alle $w, w' \in W$ gelten, dass $w \equiv_{\Sigma_a} w'$ gdw. für alle $\psi \in \Sigma_a$ gilt, dass $(m, h_w) \models \psi$ gdw. $(m, h_{w'}) \models \psi$. Mit \tilde{W} wird die Menge aller Repräsentanten der zugehörigen Äquivalenzklassen bezeichnet. Mit \tilde{X}_i wird diejenige Teilmenge von \tilde{W} bezeichnet, die aus allen Repräsentanten besteht, die zur Äquivalenzklasse X_i der Relation R gehören. Dies ist konsistent. Falls die Relation R auf W angewandt wird und die Äquivalenzrelation \equiv_{Σ_a} wiederum auf jeder Klasse X_i angewandt wird, entsteht eine Branching Time Struktur. Da Σ_a endlich ist, ist es folglich auch \tilde{W} und damit jedes \tilde{X}_i . Der folgende Rahmen $\mathcal{F}_{FIN} = (Tree, \leq, \tilde{\mathcal{A}}, \tilde{\mathcal{C}}, \tilde{\mathcal{N}}, \tilde{\mathcal{B}}, \tilde{\mathcal{D}}, \tilde{\mathcal{I}})$ ist somit endlich.

- $Tree := \{w \mid w \in \tilde{W}\} \cup \{\tilde{X}_i \mid i \in I\} \cup \{\tilde{W}\}$ ist endlich.
- $\leq := rtrcl(\{(\tilde{W}, \tilde{X}_i), (\tilde{X}_i, w) \mid w \in \tilde{X}_i, i \in I\})$,
so dass eine eindeutige Abbildung der Menge der Geschichten \tilde{H} und der Menge \tilde{W} besteht, wobei $\tilde{H} = \{h_w \mid w \in \tilde{W}\}$.
- Die Menge der Akteure sei $\tilde{\mathcal{A}} = \{\alpha \mid \alpha \text{ kommt in } \varphi \text{ vor}\}$ und ist damit endlich.¹²
- Für alle $\alpha \in \tilde{\mathcal{A}}$ legen die Relationen $\equiv_{\Sigma_e}^\alpha$ auf jeder Menge \tilde{X}_i mit $w \equiv_{\Sigma_e}^\alpha w'$ gdw. für alle $\alpha \text{ cstit} : \psi \in \Sigma_e$ fest. Dann gilt, dass $\alpha \text{ cstit} : \psi \in w$ gdw. $\alpha \text{ cstit} : \psi \in w'$. Mit $\tilde{U}_{[\alpha]\tilde{X}_i}^i$ wird die Menge aller Äquivalenzklassen auf \tilde{X}_i bezeichnet. Die Funktion $\tilde{\mathcal{C}}$, die die wahläquivalenten Geschichten festlegt, wird wie folgt definiert.
 - $\tilde{\mathcal{C}}(\alpha, w) := \{\{h_w\}\}$, wobei $h_w = \{w, \tilde{X}_i, \tilde{W}\}$ die einzige Geschichte ist, die durch den Moment w verläuft,
 - $\tilde{\mathcal{C}}(\alpha, \tilde{W}) := \{\{h_w \mid w \in \tilde{W}\}\}$,
 - $\tilde{\mathcal{C}}(\alpha, \tilde{X}_i) := \{H \mid \exists e : e \in U_{[\alpha]\tilde{X}_i}^i \text{ und } H = \{h_w \mid w \in e\}\}$.

¹¹Für die Festlegung von W siehe die Konstruktion des kanonischen Rahmens S.247f.

¹²Eigentlich müsste wie im obigen Rahmen mit den Äquivalenzklassen bezüglich der Relation \equiv_W auf A argumentiert werden. Da die Vorgehensweise analog zu obiger Handhabung verläuft, lasse ich dies zugunsten der Lesbarkeit weg. Akteure werden nun vollständig mit Akteursvariablen gleichgesetzt.

6. Vollständigkeit und Erweiterungen von bdi-stit

Aufgrund der eineindeutigen Beziehung zwischen der Menge aller $w \in \tilde{W}$ und der Menge der Geschichten in $(Tree, \leq)$ sind die folgenden Definitionen ebenfalls wohldefiniert. Für alle $\diamond\psi, \alpha bel:\psi, \alpha des:\psi \in \Sigma_\varphi, m \in Tree, w \in \tilde{W}$ und $\alpha \in \tilde{\mathcal{A}}$ seien

- $|\psi| := \bigcup_{i \in I} \{(w', h_{w'}) \mid \psi \in w'\} \in N_{(m, h_w)}$ gdw. $\diamond\psi \in w$,
- $|\psi| \in \tilde{B}_{(m, h_w)}^\alpha$ gdw. $\alpha bel:\psi \in w$,
- $|\psi| \in \tilde{D}_{(m, h_w)}^\alpha$ gdw. $\alpha des:\psi \in w$.
- Für alle $w \in \tilde{W}$ und $\alpha \in \tilde{\mathcal{A}}$ wird $t_w^\alpha = \{\psi \mid \alpha int:\psi \in w\}$ gesetzt, so dass es aufgrund der Widerspruchsfreiheit nach Axiomen $(T_i), (D_i)$ mindestens ein $\tilde{w}^\alpha \in \tilde{W}$ mit $t_w^\alpha \cap \Sigma_\alpha \subseteq \tilde{w}^\alpha$ gibt. Somit sind für alle $m \in Tree$ die folgenden Mengen nicht leer

$$\tilde{I}_{(m, h_w)}^\alpha = \{(\tilde{w}^\alpha, h_{\tilde{w}^\alpha}) \mid \tilde{w}^\alpha \in \tilde{W}, t_w^\alpha \cap \Sigma_\alpha \subseteq \tilde{w}^\alpha\}.$$

Lemma 13 Sei $\mathcal{F} = (Tree, \leq, \mathcal{A}, N, C, B, D, I)$ der kanonische Rahmen. Man wähle eine feste, beliebige Äquivalenzklasse X_i aus, wobei \tilde{X}_i die zugehörige Klasse im endlichen Rahmen $\mathcal{F}_{FIN} = (Tree, \leq, \tilde{\mathcal{A}}, \tilde{N}, \tilde{C}, \tilde{B}, \tilde{D}, \tilde{I})$ nach Filtration sei. Bezeichne $U_{[\alpha]\tilde{X}_i}^i$ die Menge aller Äquivalenzklassen von \equiv_{Σ_e} auf \tilde{X}_i für alle $\alpha \in \tilde{\mathcal{A}}$, dann gelten:

(i) Falls $\square\psi \in \Sigma_\alpha, w \in \tilde{X}_i$,

$$\square\psi \in w \text{ gdw. } \psi \in w' \text{ für alle } w' \in \tilde{X}_i.$$

(ii) Falls $\alpha cstit:\psi \in \Sigma_e, w \in \tilde{X}_i$,

$$\alpha cstit:\psi \in w \text{ gdw. } \psi \in w' \text{ für alle } w' \in \tilde{X}_i \text{ mit } w \equiv_{\Sigma_e}^\alpha w'.$$

(iii) Für alle Äquivalenzklassen $e_\alpha \in \tilde{U}_{[\alpha]\tilde{X}_i}^i$ gilt

$$\bigcap_{\alpha \in \tilde{\mathcal{A}}} e_\alpha \neq \emptyset.$$

(iv) Sei $\varphi \in \mathcal{L}_n$ mit $n \geq 1$, dann gilt für alle $j \in [0, |\tilde{\mathcal{A}}|]$

$$\left| U_{[\alpha]\tilde{X}_i}^i \right| \leq n.$$

BEWEIS: Der Beweis läuft analog in [19]. Die Klasse \tilde{X}_i ist dabei festgewählt. \square

Dieser endliche Rahmen erfüllt die Bedingung der Unabhängigkeit der Akteure. Für alle Momente $m \in \{\tilde{W}, w \mid w \in \tilde{W}\}$ ist es offensichtlich, dass für eine beliebige Funktion $\sigma_m : \tilde{A} \rightarrow \tilde{C}_m$ die Schnittmenge $\bigcap \{\sigma_m(\alpha) \mid \alpha \in \tilde{\mathcal{A}}\}$ nicht leer ist. Falls $m = \tilde{X}_i$, dann gibt es für alle $\alpha \in \tilde{\mathcal{A}}$ eine Klasse $e_\alpha \in U_{[\alpha]_{\tilde{X}_i}}^i$ mit $\sigma_m(\alpha) = e_\alpha$, so dass die Eigenschaft 13. (iii) erfüllt ist, d.h. $\bigcap \{\sigma_{\tilde{X}_i}(\alpha) \mid \alpha \in \tilde{\mathcal{A}}\} \neq \emptyset$. Für festes $n \geq 1$ und \mathcal{L}_n ist das zugehörige Axiom (APC_n) ebenfalls erfüllt, da $|\tilde{C}_{\tilde{W}}^\alpha| = |\tilde{C}_w^\alpha| = 1$ und $|\tilde{C}_{\tilde{X}_i}^\alpha| = |U_{[\alpha]_{\tilde{X}_i}}^i| \leq n$ für alle $\alpha \in \tilde{\mathcal{A}}$ nach Lemma 13.(iv). Für ein beliebiges $\psi \in \Sigma_\varphi$ kann gezeigt werden, dass für jede Klasse \tilde{X}_i

$$\mathcal{M}_{FIN}(\tilde{X}_i, h_w) \models \psi \quad \text{gdw.} \quad \psi \in w,$$

wobei $\mathcal{M}_{FIN} = (\mathcal{F}_{FIN}, v)$ und v die Interpretationsfunktion des kanonischen Modells auf \tilde{W} beschränkt ist. Der Beweis läuft über Induktion.

$$\begin{aligned} (\tilde{X}_i, h_w) \models p & \quad \text{gdw.} \quad (\tilde{X}_i, h_w) \in v(p) \quad \text{gdw.} \quad p \in w. \\ (\tilde{X}_i, h_w) \models \neg\psi & \quad \text{gdw.} \quad (\tilde{X}_i, h_w) \not\models \psi \quad \text{gdw.} \quad \psi \notin w \quad \text{gdw.} \quad \neg\psi \in w. \\ (\tilde{X}_i, h_w) \models \chi \wedge \psi & \quad \text{gdw.} \quad \text{gdw.} \quad \chi \in w \text{ und } \psi \in w \quad \text{gdw.} \quad (\chi \wedge \psi) \in w. \\ (\tilde{X}_i, h_w) \models \Box\psi & \quad \text{gdw.} \quad \text{für alle } h \in H_{\tilde{X}_i} \text{ gilt: } (\tilde{X}_i, h) \models \psi \\ & \quad \text{gdw.} \quad \text{für alle } w' \in \tilde{X}_i : \psi \in w' \stackrel{13.(i)}{\Leftrightarrow} \Box\psi \in w. \\ (\tilde{X}_i, h_w) \models \alpha \text{ cstit} : \psi & \quad \text{gdw.} \quad \text{für alle } h \in \tilde{C}_{\tilde{X}_i}^\alpha(h_w) \text{ gilt: } (\tilde{X}_i, h) \models \psi \\ & \quad \text{gdw.} \quad \psi \in w' \text{ für alle } w' \in e_\alpha \in U_{[\alpha]_{\tilde{X}_i}}^i \text{ mit } w \in e_\alpha \\ & \quad \text{gdw.} \quad \text{für alle } w' \in \tilde{X}_i: \text{ falls } w, w' \in e_\alpha, \text{ dann } \psi \in w' \\ & \quad \text{gdw.} \quad \alpha \text{ cstit} : \psi \in w. \\ (\tilde{X}_i, h_w) \models \Diamond\psi & \quad \text{gdw.} \quad \text{es existiert } U \in \tilde{N}_{(\tilde{X}_i, h_w)} : U \subseteq \|\psi\| \\ & \quad \text{gdw.} \quad \text{es gibt } \chi \in \Sigma_\varphi \text{ mit } |\chi| \subseteq \|\psi\| \text{ und } \Diamond\chi \in w \\ & \quad \text{gdw.} \quad \Diamond\psi \in w. \end{aligned}$$

Falls $|\chi| \subseteq \|\psi\|$, dann $|\chi| \subseteq |\psi|$, d.h. für alle $w' \in \tilde{W} : (\psi \supset \chi) \in w'$. Angenommen es gibt ein $w \in \tilde{W}$ mit $(\psi \supset \chi) \notin w$. Da \tilde{W} eine vollständige Menge aller Repräsentanten der \equiv_{Σ_a} -Äquivalenzklassen ist, existiert ein $\tilde{w} \in \tilde{W}$ mit $w \equiv_{\Sigma_a} \tilde{w}$. Im kanonischen Modell \mathcal{M} gilt, dass $(\psi \supset \chi) \notin w$. Dann $\mathcal{M}, (X_i, h_w) \models \neg(\psi \supset \chi)$ und $\mathcal{M}, (X_i, h_{\tilde{w}}) \models (\psi \supset \chi)$. Demnach gilt, dass entweder $\mathcal{M}, (X_i, h_w) \not\models \psi$ oder $\mathcal{M}, (X_i, h_w) \models \chi$ und $\mathcal{M}, (X_i, h_{\tilde{w}}) \models \psi$ und $\mathcal{M}, (X_i, h_{\tilde{w}}) \not\models \chi$. Dies widerspricht aber $w \equiv_{\Sigma_a} \tilde{w}$ für $\chi, \psi \in \Sigma_a$. Daher gilt für alle $w \in \tilde{W}$, dass $\chi \supset \psi \in w$, und nach Regel

6. Vollständigkeit und Erweiterungen von *bdi-stit*

(*RM*), dass $\diamond\chi \supset \diamond\psi \in w$ somit $\diamond\psi \in w$. Die Gegenrichtung ist offensichtlich. Man setze $\psi = \chi$. Anhand ähnlicher Betrachtungen erhält man die folgenden Äquivalenzen.

$$\begin{aligned}
 (X_i, h_w) \models \alpha \text{ bel} : \psi & \text{ gdw. es gibt ein } U \in \tilde{B}_{(X_i, h_w)}^\alpha \text{ mit } \emptyset \neq U \subseteq \|\psi\| \\
 & \text{ gdw. es gibt } \chi \in \Sigma_\varphi \text{ mit } \emptyset \neq |\chi| \subseteq \|\psi\| \text{ und } \alpha \text{ bel} : \chi \in w \\
 & \text{ gdw. } \alpha \text{ bel} : \psi \in w. \\
 (X_i, h_w) \models \alpha \text{ des} : \psi & \text{ gdw. es gibt } U \in \tilde{D}_{(X_i, h_w)}^\alpha \text{ mit } \emptyset \neq U \subseteq \|\psi\| \\
 & \text{ gdw. es gibt } \chi \in \Sigma_\varphi \text{ mit } \emptyset \neq |\chi| \subseteq \|\psi\| \text{ und } \alpha \text{ des} : \chi \in w \\
 & \text{ gdw. } \alpha \text{ des} : \psi \in w. \\
 (X_i, h_w) \models \alpha \text{ int} : \psi & \text{ gdw. } \tilde{I}_{(X_i, h_w)}^\alpha \subseteq \|\psi\| \\
 & \text{ gdw. für alle } w' \in \tilde{W}: \text{ falls } t_w^\alpha \cap \Sigma_a \subseteq w', \text{ dann } \psi \in w' \\
 & \text{ gdw. } \alpha \text{ int} : \psi \in w.
 \end{aligned}$$

Wenn angenommen wird, dass $\alpha \text{ int} : \psi \notin w$, dann gilt $\neg\alpha \text{ int} : \psi \in w$. Es sind nur die Fälle (i) $\alpha \text{ int} : \neg\psi \in w$ oder (ii) $\neg\alpha \text{ int} : \neg\psi \in w$ möglich. Falls (i), dann $\neg\psi \in t_w^\alpha$ und $\neg\psi \in t_w^\alpha \cap \Sigma_a \subseteq w'$, da $\Sigma_i \subseteq \Sigma_a$. Sei (ii) der Fall, dass $\neg\alpha \text{ int} : \psi \in w$ und $\neg\alpha \text{ int} : \neg\psi \in w$, dann sind $\psi, \neg\psi \notin t_w^\alpha$. Dann gibt es ein w'' mit $t_w^\alpha \cap \Sigma_a \subseteq w''$ und $\neg\psi \in w''$. Dieser Widerspruch widerlegt die Annahme und führt zu $\alpha \text{ int} : \psi \in w$.

Nach diesen Betrachtungen steht fest, dass analog zu der *dstit* Logik auch die *bdi-stit* Logik mit endlich viele Axiomen und Regeln vollständig axiomatisiert und entscheidbar ist.

6.2 Ein Tableauekalkül für *bdi-stit*

In diesem Abschnitt soll ein Tableauekalkül für die *bdi-stit* Logik vorgestellt werden. Dafür wird der bereits in [159] entwickelte Tableauekalkül für die *dstit* Logik um Tableauregeln für die monotonen, klassischen Operatoren erweitert.¹³ Dieser Tableauekalkül stellt ein weiteres mögliches Beweissystem zur Verfügung, um Formeln der *bdi-stit* Logik daraufhin zu untersuchen, ob sie kontingenterweise wahr oder falsch, tautologisch oder kontradiktorisch sind. Ebenso ist es anhand dieses Kalküls möglich, den Beweis der Korrektheit und Vollständigkeit der *bdi-stit* Logik auf einfachem Weg zu führen.

Ein Tableau basiert auf einer Baumstruktur. Es besitzt einen Wurzelknoten, Zwei-

¹³Die Regeln für diese klassischen Operatoren wurden bereits in [128] veröffentlicht. Leider bedurfte es der Korrektur der Negationsbeseitigungsregeln für die klassischen Operatoren. In diesem Abschnitt finden sich die korrigierten Regeln.

ge,¹⁴ an den Verzweigungen Knoten und am Ende jedes Zweiges ein sogenanntes Blatt. In jedem Knoten, jedem Blatt und der Wurzel stehen Ausdrücke, die ich im Weiteren auch *syntaktische Informationen* nenne. An einem neu entstandenen Blatt kann durch Anwendung einer Tableauregel der Baum erweitert werden. Wie in [159] wird bei den Tableauregeln zwischen *Strukturregeln* und *Operatorregeln* unterschieden. Strukturregeln beziehen sich auf die darunter liegende BT + AC Struktur und liefern syntaktische Informationen als Beschriftung an Knoten, die die Einführung neuer Situationen, Geschichten, Momente und Akteure bzw. die Beziehungen zwischen diesen im korrespondierenden Modell darstellen. Unter einer Information bzw. einem Ausdruck verstehe ich jede Symbolkette, die durch Regelanwendung dem Zweig hinzugefügt wird. Der Tableukalkül und damit die Informationen an den Knoten und Blättern eines Tableaus bilden die Grundlage für eine syntaktische Folgerungsbeziehung. Daraus resultiert die Bezeichnung der Ausdrücke als syntaktische Informationen. Aufgrund der noch nachzuweisenden Vollständigkeit des Tableukalküls kann eine Interpretation der Symbolketten eines bestimmten Tableaus ein Modell induzieren, das die Formeln des Wurzelknotens erfüllt.

Im Gegensatz zu den Strukturregeln, die Informationen über die zwischen den Momenten, Geschichten und Akteuren bestehenden Relationen auf der induzierten BT + AC Struktur liefern, geben die Operatorregeln vor, welche Teilformeln bzw. Informationen über Situationen und Akteure durch Auswertung von Formeln auf dem Zweig entstehen und betrachtet werden müssen, um die Formel bzw. die Formelmenge des Wurzelknotens in dem zum Zweig korrespondierenden Modell bzw. dem vom Zweig induzierten Modell zu erfüllen.

Ziel der Regelanwendungen ist es, möglichst viele Zweige zu schließen bzw. zu vervollständigen. Ein Zweig ist *geschlossen*, wenn auf dem Zweig Informationen vorliegen, dass eine Formel und ihre Negation in ein und derselben Situation wahr sind. Dies ist offensichtlich kontradiktorisch und wäre in einem Modell nicht erfüllbar. Dem Zweig ist kein Modell *getreu*, da nicht alle Formeln auf dem Zweig erfüllt werden können. Durch den Zweig wird somit kein Modell *induziert*. Liegen keine widersprüchlichen Informationen auf einem Zweig vor, wird der Zweig als *offen* betrachtet und steht für weitere Regelanwendungen zur Verfügung. Ein Zweig heißt *vollständig*, wenn keine Regelanwendung mehr möglich ist oder es nur noch möglich ist, durch eine wiederholte Folge von immer gleichen Regelanwendungen Informationen auf dem Zweig zu erhalten, die in noch zu definierender Weise äquivalent zu

¹⁴Aufgrund der Verwendung des Begriffes ‘Pfad’ in *BDICTL** übersetze ich ‘tableau branch’ mit ‘Tableauzweig’.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

<i>sub</i>	<i>ref</i>	<i>sym</i>	<i>tran</i>
$\varphi, (m, h_i)$	·	$\alpha = \beta$	$\alpha = \beta$
$\alpha = \beta$			$\beta = \gamma$
↓	↓	↓	↓
$\varphi(\alpha/\beta), (m, h_i)$	$\alpha = \alpha$	$\beta = \alpha$	$\alpha = \gamma$

Tabelle 6.1: Identitätsregeln für Akteursvariablen

Informationen sind, die durch dieselbe Folge von Regelanwendungen sich bereits auf dem Zweig befinden. Insbesondere sind geschlossene Zweige vollständig. Wenn alle Zweige vollständig sind, heißt das Tableau *vollständig*. Ebenso gilt, wenn alle Zweige des Tableaus geschlossen sind, ist das Tableau *geschlossen*. Anderenfalls ist es *offen*. Dieser Abschnitt gliedert sich somit wie folgt. Es werden die Strukturregeln und Operatorregeln kurz vorgestellt. Es wird erläutert, wie und wann die Regeln korrekt angewandt werden. Schlussendlich wird die Korrektheit und Vollständigkeit dieser Tableauregeln nachgewiesen.

Strukturregeln

Da die durch einen offenen, vollständigen Zweig induzierten Modelle auf BT Strukturen basieren, muss ein Zweig eines Tableaus Informationen über die Ordnung der Momente und damit die Menge der Geschichten enthalten, die durch den jeweiligen Moment verlaufen. Zudem sollen die Wahlmöglichkeiten des Akteurs zu einem Zeitpunkt dargestellt und die Wahläquivalenz von Geschichtsverläufen für einen beliebigen Akteur deutlich gemacht werden, d.h. die BT + AC Struktur muss mittels der Strukturregeln sicher gestellt werden. So sollte sich die Unabhängigkeit der Akteure in der abzulesenden BT + AC Struktur wiederfinden lassen. Dafür sind Strukturregeln nötig, siehe hierfür die Tabelle 6.1 für die Identitätsregeln der Akteursvariablen und die Tabellen 6.2 und 6.3 für die Strukturregeln der Moment und Geschichtsverläufe.

Die Strukturregeln in Tabelle 6.1 und in Tabelle 6.3 sind selbsterklärend. Zum Teil sind die Regeln *REF*, *SYM*, *TRAN*, *IND* aus [159] übernommen. In dem durch das Tableau induzierten Modell wird jede Akteursvariable von der Bewertungsfunktion v mit sich selbst interpretiert. Die Untersagung der Anwendung der Regel *ref* für Akteure, die sich nicht auf dem Zweig befinden, verhindert, dass es notwendig ist, Identitäten zwischen Akteuren festzuhalten, die sich nicht auf dem Zweig befin-

6.2. Ein Tablea kalkül für bdi-stit

ID	ID_{ref}	ID_{sym}	ID_{tran}	M_{tran}	$M_{antisym}$	ONE
$m_i = m_k$	$m \in h$	$m_k = m_i$	$m_k = m_i$	$m_k \triangleleft m_i$	$m_i \triangleleft m_k$	$m_i = m_k$
$\varphi, (m_i, h)$	\downarrow	\downarrow	$m_i = m_l$	$m_i \triangleleft m_l$	$m_k \triangleleft m_i$	$m_k \in h$
\downarrow	$m = m$	$m_i = m_k$	\downarrow	\downarrow	\downarrow	\downarrow
$\varphi, (m_k, h)$			$m_k = m_l$	$m_k \triangleleft m_l$	$m_i = m_k$	$m_i \in h$

H_{uni}	H_{lin}			H_{con}		
$m_n \triangleleft m_i$	$m_i \triangleleft m_l, m_k \triangleleft m_l$			$m_i \in h_k, m_j \in h_l$		
$m_i \in h$	\swarrow	\downarrow	\searrow	\swarrow	\downarrow	\searrow
\downarrow	$m_i \triangleleft m_k$	$m_i = m_k$	$m_k \triangleleft m_i$	$m_i \triangleleft m_j$	$m_j \triangleleft m_i$	$m_n \triangleleft m_i$
$m_n \in h$						$m_n \triangleleft m_j$
						n neu

Tabelle 6.2: Strukturregeln für die BT+AC Struktur I

den, um die Erfüllbarkeit bzw. Nichterfüllbarkeit einer Formelmenge zu überprüfen.¹⁵ Während die Identitätsregeln Informationen über die Akteursvariablen und somit über die Akteure im Modell liefern, stellen die weiteren Strukturregeln in Tabelle 6.2 und in Tabelle 6.3 Beziehungen zwischen den Momenten, Geschichten, Situationen und Akteuren da. So wird das Symbol \triangleleft im anhand des Tableaus erzeugten Modell mit der Ordnungsrelation auf der Menge der Momente interpretiert. Die für einen Akteur α in einem Moment m äquivalenten Geschichten in der BT + AC Struktur des Modells werden anhand der Informationen erzeugt, die durch Symbolketten \triangleleft_m^α dargestellt werden, die die Informationen umfassen, für welchen Akteur α in welchem m Moment die zwei Geschichten wahläquivalent sind. Dass es sich um eine Äqui-

¹⁵Ich unterscheide bei den Strukturregeln zwischen Identitätsregeln und weiteren Strukturregeln, obwohl ich ebenso wie bei dem Endlichkeitsbeweis in Abschnitt 6.1 der Einfachheit halber annehme, dass verschiedene Akteursvariablen durch die Interpretationsfunktion des Modells verschiedenen Akteuren zugeordnet werden, so dass auf die Identitätsregeln verzichtet werden kann. Die Äquivalenzklassen der Akteure sind dann einelementig. Sollte von einem vollständigen Tableau das induzierte Modell konstruiert werden, würden die durch die Identitätsregeln eingeführten Identitäten zwischen Akteuren wiedergegeben werden, indem die Menge der Akteure beibehalten wird und die in Identitätsbeziehungen stehenden Akteursvariablen durch die Bewertungsfunktion des Modells auf denselben Akteur abgebildet werden. Würde eine Regelanwendung dazu führen, dass eine Variable mehr als die bis dato vorliegenden Akteursvariablen im Modell auf dem Zweig benötigt, kann der Definitionsbereich von der Bewertungsfunktion erweitert werden, ohne die Menge der Akteure des Modells zu ändern.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

<i>ENV</i>	<i>REF</i>	<i>SYM</i>	<i>IND</i>	<i>SER</i>
$(m_i, h_i) \in \{s_\varphi\}$	$m \in h_i$	$h_k \triangleleft_m^\alpha h_i$	$h_{l_1} \triangleleft_m^{\alpha_1} h_{l_1}$	$m \in h$
↓	↓	↓	$\dots h_{l_k} \triangleleft_m^{\alpha_k} h_{l_k}$	↓
$\varphi, (m_i, h_i)$	$h_i \triangleleft_m^\alpha h_i$	$h_i \triangleleft_m^\alpha h_k$	↓	$(m_l, h_l) \in I_{(m,h)}^\alpha,$
	<i>TRAN</i>	<i>UNI</i>	$m \triangleleft m_n, n \text{ neu}$	$m_l \in h_l, m_k \in h_k,$
			$m \in h_n, m_n \in h_n,$	$I_{(m,h)}^\alpha \in N_{(m_k, h_k)}$
			$h_{l_1} \triangleleft_m^{\alpha_1} h_n \dots$	$l, k \text{ neu}$
	$h_i \triangleleft_m^\alpha h_k$	$h_i \triangleleft_{m_l}^\alpha h_k$	$h_{l_k} \triangleleft_m^{\alpha_k} h_n$	
	$h_k \triangleleft_m^\alpha h_l$	$m_l = m_j$		
	↓	↓		
	$h_i \triangleleft_m^\alpha h_l$	$h_i \triangleleft_{m_j}^\alpha h_k$		

Tabelle 6.3: Strukturregeln für die BT+AC Struktur II

valenzrelation handelt, sichern die Strukturregeln in Tabelle 6.3. Sie beinhalten die Informationen über die Reflexivität (Regel *REF*), die Symmetrie (Regel *SYM*) und die Transitivität (Regel *TRAN*) der für jeden Akteur wahläquivalenten Geschichten in einem vollständigen Zweig eines Tableaus. Die Regel *UNI* garantiert, dass, falls die Symbole m_l und m_j denselben Moment im Modell repräsentieren, für alle Akteure und alle Geschichten dieselben Geschichten in m_l und m_j wahläquivalent sind.

Dass es sich bei der Relation, die durch \triangleleft erzeugt wird, um eine Teilrelation einer Ordnungsrelation \leq einer BT Struktur handelt, wird durch die Strukturregeln in Tabelle 6.2 sicher gestellt. Die Relation ist transitiv. Sie erfordert zudem eine gewisse Konnektivität hinsichtlich früherer Momente und schließt eine gewisse Konnektivität hinsichtlich späterer Momente aus, vgl. Definition 1 einer BT Struktur. Die Regeln in Tabelle 6.2 sichern daher, dass je zwei Momente einen gemeinsamen Vorgänger haben (Regel H_{con}), dass die Geschichtsverläufe eine eindeutige Vergangenheit haben (Regeln H_{lin}, H_{uni}) und dass die Relationssymbole \triangleleft eine transitive Relation induzieren (Regeln M_{tran}), die jedoch echte Kreise ausschließt (Regel $M_{antisym}$). Auch wenn zwei Geschichten einen gemeinsamen Punkt in der Vergangenheit besitzen und somit verbunden sind (Regel H_{con}), gibt es für zwei Momente, die nicht vergleichbar sind, keinen gemeinsamen Moment in der Zukunft. Gäbe es einen gemeinsamen Moment, dann müssten nach Regel H_{lin} die Momente vergleichbar sein und möglicherweise zusammenfallen (Regel $M_{antisym}$). Die Regel H_{uni} stellt sicher, dass alle Geschichten, die durch spätere Momente verlaufen, auch durch alle vorherigen Momente verlau-

fen. Die anderen Regeln in Tabelle 6.2 sind selbsterklärend. Sie stellen die Identität zwischen Momenten als eine Äquivalenzrelation dar (Regeln ID_{ref} , ID_{sym} , ID_{tran}). Sie sichern, dass in identischen Momenten mit denselben Geschichten dieselben Formeln wahr sind (Regel ID) bzw. dass durch identische Momente dieselben Geschichten verlaufen (Regel ONE).

Die Strukturregel REF in Tabelle 6.3 garantiert, dass für jede Geschichte die Information auf dem Zweig verwendet werden kann, dass für jeden Akteur α jede Geschichte zu sich selbst wahläquivalent ist. Die Regel SYM liefert die Information, dass für jeden Akteur in jedem Moment gilt, dass, wenn die Information – die Geschichte mit Index i ist wahläquivalent zur Geschichte mit Index k – auf dem Zweig sich befindet, dann die Information – die Geschichte mit Index k ist wahläquivalent zur Geschichte mit Index i – ebenfalls auf dem Zweig ist. Die Regel $TRANS$ liefert die Informationen, um den transitiven Abschluss der wahläquivalenten Geschichten auf einem Zweig zu sichern. Die Regel SER sorgt dafür, dass die Information auf dem Zweig ist, dass es mindestens eine Situation geben muss, die kompatibel mit dem ist, was ein Akteur beabsichtigt. Die Regel IND sichert, dass sich die Informationen für die Unabhängigkeit der Akteure auf dem Zweig befinden. Für auf dem Zweig befindliche Informationen der Art, dass durch einen Moment m k beliebige Geschichten verlaufen, und dass sich für jede beliebige Kombination von k Akteuren auf dem Zweig Informationen befinden, dass jede Geschichte für einen anderen der k Akteure zu sich selbst wahläquivalent ist, liefert die Regel IND Informationen, wonach eine neu hinzugenommene Geschichte auf dem Zweig eingeführt wird, so dass die neu eingeführte Geschichte zu jeder der k Geschichten hinsichtlich des ihr zugeordneten Akteurs wahläquivalent ist. Diese Regel IND , entsprechend angewendet, garantiert, dass das induzierte Modell bzw. die darunterliegende $BT + AC$ Struktur die Unabhängigkeitsbedingung der Akteure erfüllt. Die Regel ENV stellt das Folgende sicher. Wenn eine Umgebung eines Umgebungssystems bei der Beseitigung eines Nachbarschaftsoperators gebildet wurde, so dass die Umgebung mit φ kompatibel ist, dann ist jede neu in die Umgebung eingeführte Situation ebenfalls mit φ kompatibel.

Einschränkungen der Regelanwendungen

Weiterhin ist es vonnöten, einige Bedingungen für die Regelanwendungen aufzustellen. Die Indizes, die für die Nummerierung der Geschichten und Momente verwendet werden, sind natürliche Zahlen und für jede hinzugenommene Information über einen neuen Moment m_k , eine neue Geschichte h_k bzw. neue Situation $m_k \in h_k$, wird als Index k die kleinste natürliche Zahl genommen, die in keiner Information des Tableaus

6. Vollständigkeit und Erweiterungen von *bdi-stit*

verwandt wird, die auf einen Moment bzw. auf eine Geschichte verweist.

Sollte eine Regelanwendung einer Strukturregel allein zu Informationen führen, die sich bereits auf dem Zweig befinden, so ist diese Regel nicht mehr anwendbar. Ebenfalls ist es nicht nötig, eine Information ein zweites Mal in denselben Zweig zu schreiben, wenn sie sich durch eine Regelanwendung zwar erneut ergibt, aber sich bereits auf dem Zweig befindet. Insbesondere die Regeln H_{con} und H_{lin} seien hier hervorgehoben. Ihre Anwendbarkeit wird weiter eingeschränkt. Seien im Falle von Regel H_{con} bereits alle der in einem Teilzweig erzeugten Ausdrücke auf dem Zweig, dann ist die Regel H_{con} nicht mehr anwendbar, d.h. alle drei Teilzweige werden nicht erzeugt. Seien beispielsweise $m_i \in h_k$ und $m_j \in h_l$ auf dem Zweig und außerdem sei $m_i \triangleleft m_j$ auf dem Zweig, dann ist H_{con} nicht mehr anwendbar, d.h. auch die anderen Teilzweige werden nicht erzeugt. Dasselbe gilt, falls $m_i \in h_k$ und $m_j \in h_l$ auf dem Zweig sind und es ein m_n auf dem Zweig gibt, für das die Ausdrücke $m_n \triangleleft m_j$ und $m_n \triangleleft m_i$ bereits auf dem Zweig sind. Dann ist H_{con} ebenfalls nicht anwendbar. Für Regel H_{lin} gilt dies analog. Sei für $m_i \triangleleft m_l$ und $m_k \triangleleft m_l$ bereits einer der in den Teilzweigen genannten Ausdrücke $m_i = m_k$, $m_i \triangleleft m_k$, $m_k \triangleleft m_i$ auf dem Zweig, dann kann die Regel H_{lin} nicht mehr angewandt werden. Ziel dieser Regeleinschränkung ist es, nicht unnötig viele Momente zu erzeugen, die äquivalente Strukturen bilden.

Diese Einschränkung betrifft jedoch nur die Strukturregeln. Für Regelanwendungen zur Beseitigung eines Operators, Tabelle 6.4 und Tabelle 6.5, bei denen sich der Baum verzweigt und sich die Information nur in einem Teilzweig wiederholt, in dem anderen jedoch nicht, werden beide Teilzweige mit allen Informationen erzeugt. Beispielsweise seien die Ausdrücke φ , s und $\varphi \vee \psi$, s auf dem Zweig. Dann ist die Beseitigungsregel für \vee immer noch anwendbar bzw. muss sogar noch angewandt werden, solange der Zweig offen ist, auch wenn die Informationen des einen durch die Regelanwendung erzeugten Teilzweiges bereits auf dem Zweig vorliegen.

Ziel der Regelanwendungen sollte es sein, möglichst minimale, vollständige Tableaux zu generieren. Minimal wird hierbei im Sinne einer möglichst geringen Blatt- und Knotenanzahl verstanden. Dafür ist es sinnvoll, Regelanwendungen zu vermeiden, die auf eine unnötige Einführung sich gleich verhaltener Akteure hinauslaufen. Daher werden die Regeln *ref*, *REF* und *SER* insoweit beschränkt, dass sie nur Anwendung finden, wenn die in ihnen genannten Akteursvariablen sich bereits auf dem Zweig befinden.

$\neg\neg\varphi, s$	$(\varphi \wedge \psi), s$	$\neg(\varphi \wedge \psi), s$	
\downarrow	\downarrow	\swarrow	\searrow
φ, s	φ, s	$\neg\varphi, s$	$\neg\psi, s$
	ψ, s		
$\Box\varphi, (m, h_i), m \in h_k$	$\neg\Box\varphi, (m, h_i)$		
\downarrow	\downarrow		
$\varphi, (m, h_k)$	$\neg\varphi, (m, h_k),$		
	$m \in h_k, m_k \in h_k,$		
	$m \triangleleft m_k, k \text{ neu}$		
$\alpha \text{ dstit}: \varphi, (m, h_i),$	$\neg\alpha \text{ dstit}: \varphi, (m, h_i), m \in h_l$		
$h_i \triangleleft_m^\alpha h_k$	\swarrow	\searrow	
\downarrow	$\varphi, (m, h_l)$	$\neg\varphi, (m, h_k)$	
$\varphi, (m, h_k), m \triangleleft m_l,$	$m \in h_k, m_k \in h_k,$		
$m \in h_l, m_l \in h_l,$	$h_i \triangleleft_m^\alpha h_k,$		
$\neg\varphi, (m, h_l), l \text{ neu}$	$m \triangleleft m_k, k \text{ neu}$		

Tabelle 6.4: Operatorregeln für eine Stit-Logik mit *dstit*-Operator

Operatorregeln

Die Tabelle 6.4 stellt die Beseitigungsregeln der *dstit*-Operatoren und ihrer Negationen dar, die aus [159] entnommen sind. Die Symbole s bzw s_k in den Regeltabellen verwende ich abkürzend für ein Moment-Geschichte Paar. Auf dem Tableau steht statt des s ein Ausdruck (m_i, h_k) . Die Regeln der aussagenlogischen Junktoren sind selbsterklärend. Für manche Operatoren ist zu beachten, dass eine Regelanwendung sich über den ganzen nachfolgenden Zweig erstrecken kann. Sobald ein Notwendigkeitsoperator durch eine Regelanwendung beseitigt wird, ist die Anwendung nicht notwendigerweise abgeschlossen. Für die Beseitigungsregeln des *dstit*-, *int*- und \Box -Operators und für die Beseitigungsregeln der negierten Modaloperatoren gilt es, die Regelanwendung auf jede neu hinzukommende Information auszudehnen, die für die Auswertung des jeweiligen Operators in Frage kommt.

So ist es beispielsweise nötig, wenn eine Regel auf die Knoteninformationen $\Box\varphi, (m, h_l), m \in h_i$ und zwar für alle i , die sich mit $m \in h_i$ auf dem Zweig befinden, angewandt wird und in einem späteren Knoten auf dem Zweig ein Ausdruck der Form $m \in h_k$ eingeführt wird, dass die Regel auch auf den Ausdruck $m \in h_k$ Anwendung fin-

6. Vollständigkeit und Erweiterungen von *bdi-stit*

$\begin{array}{c} \diamond\varphi, s \\ \downarrow \\ (m_l, h_l) \in \{s_\varphi\}, m_l \in h_l, \\ \{s_\varphi\} \in N_s, l \text{ neu} \end{array}$	$\begin{array}{c} \neg\diamond\varphi, s, \{s_\psi\} \in N_s \\ \downarrow \\ \neg\varphi, (m_l, h_l), m_l \in h_l, \\ (m_l, h_l) \in \{s_\psi\}, l \text{ neu} \end{array}$
$\begin{array}{c} \alpha \text{ bel}: \varphi, s \\ \downarrow \\ (m_l, h_l) \in \{s_\varphi\}, m_l \in h_l, \{s_\varphi\} \in B_s^\alpha, \\ \{s_\varphi\} \in N_{(m_k, h_k)}, m_k \in h_k, k, l \text{ neu} \end{array}$	$\begin{array}{c} \neg\alpha \text{ bel}: \varphi, s, \{s_\psi\} \in B_s^\alpha, \\ \downarrow \\ \neg\varphi, (m_l, h_l), m_l \in h_l, \\ (m_l, h_l) \in \{s_\psi\}, l \text{ neu} \end{array}$
$\begin{array}{c} \alpha \text{ des}: \varphi, s \\ \downarrow \\ \varphi, (m_l, h_l), m_l \in h_l, \{s_\varphi\} \in D_s^\alpha, \\ \{s_\varphi\} \in N_{(m_k, h_k)}, m_k \in h_k, k, l \text{ neu} \end{array}$	$\begin{array}{c} \neg\alpha \text{ des}: \varphi, s, \{s_\psi\} \in D_s^\alpha, \\ \downarrow \\ \neg\varphi, (m_l, h_l), m_l \in h_l, \\ (m_l, h_l) \in \{s_\psi\}, l \text{ neu} \end{array}$
$\begin{array}{c} \neg\alpha \text{ int}: \varphi, s \\ \downarrow \\ \neg\varphi, (m_k, h_k), m_k \in h_k, (m_k, h_k) \in I_s^\alpha, k \text{ neu} \end{array}$	$\begin{array}{c} \alpha \text{ int}: \varphi, s, s' \in I_s^\alpha \\ \downarrow \\ \varphi, s' \end{array}$

Tabelle 6.5: Operatorregeln für die mentalen Zustandsoperatoren in *bdi-stit* Logik

det, da diese k -te Geschichte unter die obigen mit i aufgezählten Geschichten fällt. Die Regelanwendung ist somit erst abgeschlossen, wenn der Zweig vollständig ist. Insbesondere befindet sich keine Information über eine Geschichte, die durch m verläuft, auf dem Zweig, auf die die Regel nicht angewendet wurde. Jede neu eingeführte Situation in demselben Moment muss für die Beseitigung des Notwendigkeitsoperators auf dem Zweig berücksichtigt werden.

In Tabelle 6.5 sind die Regeln für den Überzeugungs-, Wunsch- und Möglichkeitsoperator abgebildet. Da es sich um monotone, klassische Operatoren handelt, werden Umgebungen mit mindestens einer Situation eingeführt, um sicherzustellen, dass es eine Umgebung gibt, die mit dem kompatibel ist, was der Akteur wünscht, glaubt bzw. was möglicherweise Inhalt eines mentalen Zustandes ist. Im Falle der Auswertung der Information über die Negation eines solchen Operators werden die bestehenden

Umgebungen um eine Situation erweitert, die die negierte Teilformel wahr macht.

Für die Beseitigungsregel des \diamond -Operators gilt, dass $\diamond\varphi$, s erfüllt wird, wenn Informationen über eine neue Umgebung $\{s_\varphi\}$ mit mindestens einer Situation (m_l, h_l) vorliegen, die kompatibel zu der Teilformel φ ist. Dass die Situation (m_l, h_l) kompatibel zu der durch die Umgebung repräsentierten Formel φ ist, liefert die Regel *ENV*, die besagt, dass jede in einer Umgebung eingeführte Situation die Formel erfüllt.

Für die Beseitigung der negierten Operatoren, z.B. $\neg\diamond\psi$, ist es nötig, folgende Informationen auf dem Tableau hinzuzufügen. In allen Umgebungen $\{s_\varphi\}$ wird eine neue Situation (m_l, h_l) eingeführt, in der die Formel ψ nicht wahr ist. Weiterhin muss die Information auf dem Zweig vorliegen, dass die ursprüngliche Formel φ , für die die Umgebung $\{s_\varphi\}$ eingeführt wurde, auch in der neu eingeführten Situation (m_l, h_l) wahr ist, vgl. Regel *ENV*. Die Regeln für den Wunsch- und Überzeugungsoperator sind analog zu den Regeln des \diamond -Operators.

Mithilfe der Tableauregeln wird der syntaktische Folgerungsbegriff zwischen einer Formelmenge Δ und einer Formel φ definiert. Es wird ein Wurzelknoten mit Informationen über zwei Situationen und einer Geschichte kreiert. Die Informationen beinhalten, dass alle Formeln aus Δ bzw. eine endliche Teilmenge von Δ und die Negation der abzuleitenden Formel φ in einer der beiden Situationen bestehen. Daraufhin werden die Tableauregeln solange angewandt, bis entweder der Baum geschlossen oder keine Regel mehr anwendbar bzw. sinnvoll anwendbar ist.

Definition 14 Sei $\Delta \cup \{\varphi\}$ eine Menge von bdi-stit Formeln. Dann gilt, dass $\Delta \vdash \varphi$ (φ ist syntaktisch aus Δ ableitbar) genau dann, wenn es ein geschlossenes Tableau mit einem Wurzelknoten der Form $\Delta^0 \cup \{m \in h_0, m \triangleleft m_0, m_0 \in h_0\} \cup \{\neg\varphi, (m, h_0)\}$ gibt, wobei Δ^0 eine endliche Menge von Informationen aus $\{\psi, (m, h_0) \mid \psi \in \Delta\}$ ist.

Wenn es ein vollständiges Tableau mit eben diesem Wurzelknoten gibt, das geschlossen ist, so ist φ aus der jeweiligen Teilmenge von Δ ableitbar. Wenn ein vollständiges Tableau mit diesem Wurzelknoten mindestens einen offenen Zweig hat, so ist φ nicht aus der im Wurzelknoten genannten Formelmenge ableitbar und anhand des offenen Zweiges ist ein Gegenmodell konstruierbar, dass in einer Situation die Teilmenge aus Δ und $\neg\varphi$ erfüllt.

Dabei werden die Informationen des Zweiges in ein sogenanntes *induziertes Modell* $(M, \leq, \mathcal{A}, N, Ch, Bel, Des, Int, \nu)$ übersetzt, siehe Definition 19, S.276. Die Symbole $m_i \triangleleft m_k$ legen die Ordnung der auf dem Zweig vorkommenden Momenten fest. Im Modell wird der transitive und reflexive Abschluss der Menge gebildet, so dass die Relation \leq gewonnen wird. Die Ausdrücke der Form $m \in h$ bedeuten, dass es sich

6. Vollständigkeit und Erweiterungen von *bdi-stit*

bei (m, h) um eine Situation handelt. Die Informationen $h_i \triangleleft_m^\alpha h_k$ werden interpretiert, um die Wahläquivalenz von Geschichten mit Hinsicht auf einen Akteur α und einen Moment m darzustellen. Der Wertebereich der Funktion C wird anhand der Informationen $h_i \triangleleft_m^\alpha h_k$ abgelesen. Als Akteure werden alle Akteursvariablen angesehen, die mit v im Modell durch sich selbst interpretiert werden $v(\alpha) = \alpha$.¹⁶

In den Regeln für die klassischen Operatoren zuzüglich der Regel *ENV* signalisieren Ausdrücke der Form $\{s_\varphi\} \in B_s^\alpha$ und $s_k \in \{s_\varphi\}$ für einige Indizes k , dass es eine Umgebung in dem Überzeugungssystem des Akteurs α in der Situation s gibt, die die Situationen s_k enthält, wobei φ in diesen Situationen s_k wahr ist. Die Funktion B bildet jedes Akteur-Situationspaar auf eine Menge bestehend aus Situationsmengen ab, die sich anhand des Zweiges ablesen lassen. Dasselbe gilt für die Funktion D und Ausdrücke der Form $\{s_\varphi\} \in D_s^\alpha$ und $s_k \in \{s_\varphi\}$ für einige Indizes k . Um die Funktion I festzulegen, die jedem Akteur-Situationspaar eine Menge von kompatiblen Situationen und damit eine Umgebung zuordnet, werden die Informationen $s_l \in I_s^\alpha$ des Zweiges interpretiert. Die Bewertungsfunktion v ergibt sich aus den Literalen und den Situationsinformationen, die auf dem Tableau für jede propositionale Variable und jede Situation stehen. Die Information p, s auf dem Zweig liefert $s \in v(p)$ im Modell. Wenn die Information $\neg p, s$ vorliegt, muss im Modell sichergestellt werden, dass $s \notin v(p)$. Wenn dies nicht möglich ist, ist der Zweig geschlossen. Falls keine Information auf einem Zweig bezüglich der Bewertung einer Aussagevariable in einer Situation vorliegt, kann diese Aussagevariable beliebig in dieser Situation interpretiert werden.

Vollständigkeit des Tableaukalküls für *bdi-stit* Logik

Die Vollständigkeit des Tableaukalküls wird auf übliche Art und Weise bewiesen, indem gezeigt wird, falls eine Formel φ aus der Formelmenge Δ nicht syntaktisch ableitbar ist, dass es zu jedem vollständigen Tableau mit einem Wurzelknoten, der aus $\Delta \cup \{\neg\varphi\}$ gebildet wird, einen offenen Zweig gibt und eine Situation in dem durch den Zweig induzierten Modell existiert, in dem die Formelmenge $\Delta \cup \{\neg\varphi\}$ erfüllt wird.

Bevor jedoch die Vollständigkeit des Kalküls bewiesen werden soll, wird Korrektheit nachgewiesen, indem für eine beliebige erfüllbare Menge $\Delta \cup \{\varphi\}$ von *bdi-stit* Formeln gezeigt wird, dass es kein geschlossenes Tableau für $\Delta \cup \{\neg\varphi\}$ gibt, da auch nach einer beliebigen Regelanwendung das Modell mindestens einem Zweig *getreu*

¹⁶Solange keine Informationen über Wahläquivalenzen in einem Moment für einen Akteur vorliegen, können die Geschichten, die durch diesen Moment laufen, für einen Akteur alle als wahläquivalent betrachtet werden. Dies sichert auf triviale Weise die Unabhängigkeit der Akteure.

bleibt. Für eine beliebige Menge $\Delta \cup \{\varphi\}$ von Formeln gilt somit, dass $\Delta \not\models \varphi$ impliziert, dass $\Delta \not\vdash \varphi$. Dafür wird definiert, wann ein Modell einem Tableauezweig getreu ist. Nach beliebiger Regelanwendung auf einem Zweig, dem ein Modell getreu ist, ist das Modell mindestens einer Erweiterung des Zweiges ebenfalls getreu. Wansing [159] hat Korrektheits- und Vollständigkeitsnachweise für einen Tableauealkül der deliberativen Stit-Logik erbracht, der über einen Tableauealkül für eine Fusion normaler monomodaler Logiken aufgrund der Unabhängigkeit der Akteure hinausgeht, so dass insbesondere die Strukturregeln zur Konstruktion der darunter liegenden BT + AC Strukturen zu beachten sind.

Für die Konstruktion des induzierten Modells sei an die Definition der „truth sets“ $\|\varphi\|$ erinnert.¹⁷ Das ist die Menge an Situationen in einem gegebenem Modell, in denen die Formel φ wahr ist. Da anhand der Wahrheitsbedingungen, vgl. Definition 8, S.236, feststeht, dass im induzierten Modell nur eine Teilmenge der „truth sets“ in einem Nachbarschaftssystem existieren muss, können u.U. in einer Umgebung endlich abzählbar viele Situationen betrachtet werden, um eine Formel $\diamond\varphi$ zu erfüllen, wobei φ möglicherweise in unendlichen vielen Situationen wahr ist. Insbesondere, wenn die BT Struktur unendlich viele Situationen enthält, lässt sich der Tableauealkül vereinfachen, da nicht mehr alle Situationen betrachtet werden müssen.

Da die Strukturregeln getrennt von den Operatorregeln untersucht werden, wird ein weiterer Begriff eingeführt, bevor definiert wird, wann ein Modell einem Zweig eines Tableaus getreu ist. Anhand der Informationen über mögliche Situationen auf dem Zweig lässt sich eine Menge $\bar{S} = \{(m_i, h_k) \mid m_i \in h_k \text{ auf } b\}$ für einen Tableauezweig definieren.

Sei eine BT Struktur mit (M, \leq) gegeben, so werden zwei Projektionen definiert, die die Menge der Situationen S der BT Struktur auf die Menge der zugehörigen Momente bzw. der zugehörigen Geschichten abbilden, $\pi_1 : S \rightarrow M$ und $\pi_2 : S \rightarrow H$ mit $\pi_1((m, h)) = m$ und $\pi_2((m, h)) = h$.¹⁸ Somit kann festgehalten werden, wann ein Zweig in eine BT Struktur eingebettet werden kann.

Definition 15 Sei (M, \leq) eine BT Struktur und b der Zweig eines Modells. Dann ist ein Zweig in (M, \leq) einbettbar, wenn es eine Funktion $f : \bar{S} \rightarrow M \times H$ gibt, so dass die folgenden Bedingungen erfüllt sind.

1. $f(m, h) \in S$, falls $m \in h$ auf dem Zweig b ist.

¹⁷Für die Definition vgl. S.236 bzw. in [39, S.38].

¹⁸Die doppelten Klammern werden im Folgenden weggelassen. Der Übersichtlichkeit halber werden einfache Klammern verwendet, sofern es eindeutig und ersichtlich ist.

6. Vollständigkeit und Erweiterungen von bdi-stit

2. $\pi_1(f(m_i, h)) \leq \pi_1(f(m_k, h'))$, falls $m_i \triangleleft m_k$, $m_i \in h$ und $m_k \in h'$ auf dem Zweig b vorkommen.
3. $f(m_i, h) = f(m_k, h)$ für alle $m_i \in h$ und $m_k \in h$ auf dem Zweig b , falls $m_i = m_k$ auf b vorkommt.

Falls es eine solche Einbettungsfunktion gibt, gilt für alle Ausdrücke der Form $m \in h$ auf dem Zweig, dass $\pi_2(f(m, h)) \in H_{\pi_1(f(m, h))}$. Abkürzend schreibe ich daher auch für (m, h) auf dem Zweig ein s und $f(s)$ anstelle von $f(m, h)$. Ebenso verwende ich s_l und $f(s_l)$ statt (m_l, h_l) und $f(m_l, h_l)$. Mithilfe dieser Definition und Festlegungen wird definiert, wann ein Modell einem Zweig getreu ist.

Definition 16 Sei $\mathcal{M} = (M, \leq, \mathcal{A}, N, Ch, Bel, Des, Int, \nu)$ ein bdi-stit Modell und b ein Zweig eines Tableaus. Das Modell \mathcal{M} ist dem Zweig b genau dann getreu, wenn es eine Funktion $f : \bar{S} \rightarrow M \times H$ gibt, die den Zweig b in die BT Struktur des Modells \mathcal{M} einbettet und die folgenden Bedingungen erfüllt. Hierbei sei $f(\bar{U}) = \{f(s) \mid s \in U\}$, und $U \subseteq S$.

1. Für jeden Ausdruck $\varphi, (m, h)$ auf dem Zweig b gilt, dass $\mathcal{M}, f(m, h) \models \varphi$.
2. Falls $h_i \triangleleft_m^\alpha h_k$ auf b ist, dann $\pi_2(f(m, h_k)) \in Ch_{\pi_1(f(m, h_k))}^{\nu(\alpha)}(\pi_2(f(m, h_i)))$.
3. Falls $\{s_\varphi\} \in N_s$ auf dem Zweig b ist, existiert $U \in N_{f(s)}$ mit $U \subseteq \|\varphi\|$ und für alle $s_k \in \{s_\varphi\}$ auf dem Zweig b gilt, dass $f(s_k) \in U$.
4. Falls $\{s_\varphi\} \in B_s^\alpha$ auf dem Zweig b ist, existiert $U \in Bel_{f(s)}^{\nu(\alpha)}$ mit $U \subseteq \|\varphi\|$ und für alle $s_k \in \{s_\varphi\}$ auf dem Zweig b gilt, dass $f(s_k) \in U$.
5. Falls $\{s_\varphi\} \in D_s^\alpha$ auf dem Zweig b ist, existiert $U \in Des_{f(s)}^{\nu(\alpha)}$ mit $U \subseteq \|\varphi\|$ und für alle $s_k \in \{s_\varphi\}$ auf dem Zweig b gilt, dass $f(s_k) \in U$.
6. Falls $(m_k, h_k) \in I_{(m, h)}^\alpha$ auf dem Zweig b ist, dann gilt $f(m_k, h_k) \in Int_{f(m, h)}^{\nu(\alpha)}$.
7. Falls s und α auf dem Zweig b sind, existiert eine Situation $(m, h) \in Int_{f(s)}^{\nu(\alpha)}$.

Die Funktion f , die die eben genannten Bedingungen erfüllt, zeigt, dass das Modell \mathcal{M} dem Zweig b getreu ist. Im Folgenden wird für jede Tableauregel einzeln gezeigt, dass ein Modell, welches einem Zweig getreu ist, nach einer Regelanwendung mindestens einer Extension des Zweiges getreu ist.

Lemma 17 Sei $\mathcal{M} = (M, \leq, \mathcal{A}, N, Ch, Bel, Des, Int, \nu)$ ein Modell und b ein Tableauezweig. Falls das Modell \mathcal{M} dem Zweig b getreu ist und eine Regel auf b angewendet wird, dann gibt es einen erweiterten Zweig b' von b nach der Regelanwendung, dem \mathcal{M} ebenfalls getreu ist.

BEWEIS: Die Induktionsannahme lautet wie folgt. Wenn f die Funktion sei, die zeigt, dass $\mathcal{M} b$ getreu ist, ist das Modell einem der nach Regelanwendung erweiterten Zweige von b getreu. Für die Identitätsregeln und die Regeln für die Junktorenregeln der klassischen Aussagenlogik ist das offensichtlich. In diesen Fällen zeigt die Funktion f selbst, dass das Modell einer Erweiterung b' getreu ist. Alle anderen Regeln werden gesondert betrachtet.

Angenommen die \Box -Regel wird auf $\Box\varphi, (m, h_i)$ auf dem Zweig b angewandt. Der Zweig wird durch $\varphi, (m, h_k)$ für alle Indizes k mit $m \in h_k$ auf dem Zweig b erweitert. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(m, h_i) \models \Box\varphi$. Nach Definition 8 gilt, $\mathcal{M}, (\pi_1(f(m, h_i)), h') \models \varphi$ für alle $h' \in H_{\pi_1(f(m, h_i))}$. Da $\pi_2(f(m, h_k)) \in H_{\pi_1(f(m, h_i))}$, gilt $\mathcal{M}, f(m, h_k) \models \varphi$ für alle h_k und folglich zeigt f , dass $\mathcal{M} b'$ getreu ist.

Angenommen die $\neg\Box$ -Regel wird auf $\neg\Box\varphi, (m, h_i)$ angewandt, so dass b durch $m \triangleleft m_k, m \in h_k, m_k \in h_k$ und $\neg\varphi, (m, h_k)$ für einen neuen Index k zum Zweig b' erweitert wird. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(m, h_i) \models \neg\Box\varphi$. Somit gibt es $h' \in H_{\pi_1(f(m, h_i))}$ mit $\mathcal{M}, (\pi_1(f(m, h_i)), h') \models \neg\varphi$. Sei f' eine Erweiterung von f , die für den neuen Index k festlegt, dass $f'(m, h_k) = (\pi_1(f(m, h_i)), h')$ und $f'(m_k, h_k) = (m', h')$ für einen Moment $m' \in h'$ mit $\pi_1(f(m, h_i)) \leq m'$. Dann gelten $\mathcal{M}, f'(m, h_k) \models \neg\varphi$, $\pi_2(f(m, h_k)) \in H_{\pi_1(f(m, h_k))}$ und $\pi_2(f(m_k, h_k)) \in H_{\pi_1(f(m_k, h_k))}$. Somit zeigt die Funktion f' , dass \mathcal{M} dem erweiterten Zweig b' getreu ist.

Wird die $\neg dstit$ -Regel auf $\neg\alpha dstit : \varphi, (m, h_i)$ angewandt, so wird b geteilt. Auf der einen Seite wird b um $\varphi, (m, h_l)$ für jeden Index l mit $m \in h_l$ auf dem Zweig zu b' erweitert. Auf der anderen Seite wird b durch $m \triangleleft m_k, m \in h_k, m_k \in h_k, h_i \triangleleft_m^\alpha h_k$ und $\neg\varphi, (m, h_k)$ für einen neuen Index k zu b'' erweitert. Nach Induktionsannahme gilt $\mathcal{M}, f(m, h_i) \models \neg\alpha dstit : \varphi$. Entweder ist die Positivbedingung (+) oder die Negativbedingung (-) des $dstit$ -Operators in \mathcal{M} verletzt. Das bedeutet, dass entweder ein $h' \in Ch_{\pi_1(f(m, h_i))}^{\nu(\alpha)}(\pi_2(f(m, h_i)))$ mit $(\pi_1(f(m, h_i)), h') \models \neg\varphi$ existiert. Oder es gilt für alle $h'' \in H_{\pi_1(f(m, h_i))}$, dass $(\pi_1(f(m, h_i)), h'') \models \varphi$. Im letzteren Fall ist die Bedingung (-) in \mathcal{M} verletzt. Es gilt $\mathcal{M}, f(m, h_l) \models \varphi$ für alle Indizes l mit $m \in h_l$ auf dem Zweig b , da $\pi_2(f(m, h_l)) \in H_{\pi_1(f(m, h_i))}$ für jeden Ausdruck $m \in h_l$ gilt. Für jedes $m \in h_l$ existiert ein $h'' \in H_{\pi_1(f(m, h_i))}$. In diesem Fall zeigt f selbst, dass $\mathcal{M} b'$ getreu ist. Folglich ist das Modell einer Erweiterung von b nämlich b' getreu.

Im anderen Zweig b'' ist die Bedingung (+) in \mathcal{M} verletzt. Betrachte man die Erwei-

6. Vollständigkeit und Erweiterungen von *bdi-stit*

terung f' von f , so dass $f'(m, h_k) = (\pi_1(f(m, h_i)), h')$ und $f'(m_k, h_k) = (m', h')$ für den neuen Index k und für eine beliebige Situation (m', h') im Modell mit $\pi_1(f'(m, h_k)) < m'$ gelten. Somit ist $\mathcal{M}, f'(m, h_k) \models \neg\varphi$ und $\pi_2(f'(m, h_k)) \in H_{\pi_1(f'(m, h_k))}$, $\pi_2(f'(m, h_k)) = h'$ mit $h' \in \text{Choice}_{\pi_1(f'(m, h_k))}^{\nu(\alpha)}(\pi_2(f'(m, h_i))) \subseteq H_{\pi_1(f'(m, h_i))}$. Daher zeigt f' , dass $\mathcal{M} b''$ getreu ist. Daher ist \mathcal{M} mindestens einer der Erweiterungen von b nach der Regelanwendung getreu.

Die *dstit*-Regel auf einen Ausdruck $\alpha \text{ dstit} : \varphi, (m, h_i)$ angewandt, erweitert b zu einem Zweig b' , auf dem für jeden Ausdruck $h_i \triangleleft_m^\alpha h_k$ auf b der Ausdruck $\varphi, (m, h_k)$ und zudem $m \triangleleft m_l$, $m \in h_l$, $m_l \in h_l$ und $\neg\varphi, (m, h_l)$ für einen neuen Index l hinzukommen. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(m, h_i) \models \alpha \text{ dstit} : \varphi$. Das bedeutet, dass sowohl die (+) als auch die (-) Bedingung erfüllt sind. Für alle $h' \in \text{Ch}_{\pi_1(f(m, h_i))}^{\nu(\alpha)}(\pi_2(f(m, h_i)))$ gilt, dass $\mathcal{M}, (\pi_1(f(m, h_i)), h') \models \varphi$ und dass es $h'' \in H_{\pi_1(f(m, h_i))}$ mit $\mathcal{M}, (\pi_1(f(m, h_i)), h'') \models \neg\varphi$ gibt. Sei die Erweiterung f' für alle Ausdrücke auf b wie f definiert. Nach Definition 16.2 gilt $\pi_2(f'(m, h_k)) \in \text{Choice}_{\pi_1(f'(m, h_i))}^{\nu(\alpha)}(\pi_2(f'(m, h_i)))$ für jeden Index k mit $h_i \triangleleft_m^\alpha h_k$ und $\mathcal{M}, f'(m, h_k) \models \varphi$. Für den Index l setze $f'(m, h_l) = (\pi_1(f(m, h_i)), h'')$ und $f'(m_l, h_l) = (m', h')$ für ein beliebiges $(m', h') \in S$ und $\pi_1(f(m, h_i)) \leq m'$. Somit gilt $\mathcal{M}, f'(m, h_l) \models \neg\varphi$ und $\pi_2(f'(m, h_l)) \in H_{\pi_1(f'(m, h_i))}$. Folglich zeigt f' , dass $\mathcal{M} b'$ getreu ist.

Angenommen die \diamond -Regel wird auf den Ausdruck $\diamond\varphi, (m, h)$ auf b angewandt, dann erweitert sich b zu b' um $(m_l, h_l) \in \{s_\varphi\}$, $m_l \in h_l$, $\{s_\varphi\} \in N_s$ für einen neuen Index l . Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(m, h) \models \diamond\varphi$. Nach Definition 8 existiert ein $U \in N_{f(m, h)}$, wobei $U \neq \emptyset$ und für alle $s' \in U$ gilt, dass $\mathcal{M}, s' \models \varphi$. Wähle f' wie f für alle Ausdrücke auf b und setze $f'(m_l, h_l) = s'$ für ein $s' \in U \subseteq S$. Demzufolge gilt $\mathcal{M}, f'(m_l, h_l) \models \varphi$ und Definition 16.3 ist erfüllt. D.h. f' zeigt, dass $\mathcal{M} b'$ getreu ist.

Falls ein Ausdruck $\neg\diamond\varphi, s$ auf b ist, so bewirkt die $\neg\diamond$ -Regel, dass sich der Zweig b für alle $\{s_\psi\} \in N_s$ um die Informationen $\neg\varphi, (m_l, h_l)$, $m_l \in h_l$ und $(m_l, h_l) \in \{s_\psi\}$ für jedes ψ um einen neuen Index l zu b' erweitert. Da $\mathcal{M}, f(s) \models \neg\diamond\varphi$, gilt für alle $U \in N_{f(s)}$ im Modell, dass es eine Situation $s_U \in U$ mit $\mathcal{M}, s_U \models \neg\varphi$ gibt. Außerdem gilt, da f zeigt, dass $\mathcal{M} b$ getreu ist, dass es zu jedem $\{s_\psi\} \in N_s$ auf dem Zweig ein $U_\psi \in N_{f(s)}$ mit $U_\psi \subseteq \|\psi\|$ gibt. Wenn f' wie f gewählt wird und für die neuen Indizes l festgelegt wird, dass $f'(m_l, h_l) = s_{U_\psi}$ für ein $s_{U_\psi} \in U_\psi$ mit $\mathcal{M}, s_{U_\psi} \models \neg\varphi$ und $\mathcal{M}, s_{U_\psi} \models \psi$, dann zeigt f' , dass $\mathcal{M} b'$ getreu ist.

Wird die *bel*-Regel auf einen Ausdruck $\alpha \text{ bel} : \varphi, s$ angewandt, dann erweitert sich b zu b' um $(m_l, h_l) \in \{s_\varphi\}$, $m_l \in h_l$, $\{s_\varphi\} \in B_s^\alpha$ und $\{s_\varphi\} \in N_{(m_k, h_k)}$, $m_k \in h_k$ für neue Indizes k, l . Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(s) \models \alpha \text{ bel} : \varphi$. So existiert nach Definition 8 ein $U \in \text{Bel}_{f(s)}^{\nu(\alpha)}$, wobei $U \neq \emptyset$ und für alle $s' \in U$ gilt, dass $\mathcal{M}, s' \models \varphi$. Da

\mathcal{M} ein *bdi-stit* Modell ist, existiert eine Situation s'' mit $U \in N_{s''}$. Wähle f' wie f für alle Ausdrücke auf b und setze $f'(m_k, h_k) = s''$ und $f'(m_l, h_l) = s'$ für ein $s' \in U \subseteq S$. Demzufolge gilt $\mathcal{M}, f'(m_l, h_l) \models \varphi$ und Definition 16.4 ist erfüllt. D.h. f' zeigt, dass $\mathcal{M} b'$ getreu ist.

Wird die \neg -*bel*-Regel auf einen Ausdruck $\neg\alpha \text{ bel}:\varphi$, s auf b angewandt, so folgt, dass sich der Zweig b für alle $\{s_\psi\} \in B_s^\alpha$ um die Informationen $\neg\varphi$, (m_l, h_l) , $m_l \in h_l$ und $(m_l, h_l) \in \{s_\psi\}$ für jedes ψ um einen neuen Index l zu b' erweitert. Da $\mathcal{M}, f(s) \models \neg\alpha \text{ bel}:\varphi$, gilt für alle $U \in \text{Bel}_{f(s)}^{\nu(\alpha)}$ im Modell, dass es eine Situation $s_U \in U$ mit $\mathcal{M}, s_U \models \neg\varphi$ gibt. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt außerdem, dass es zu jedem $\{s_\psi\} \in B_s^\alpha$ auf dem Zweig ein $U_\psi \in \text{Bel}_{f(s)}^{\nu(\alpha)}$ mit $U_\psi \subseteq \|\psi\|$ gibt. Wenn f' wie f gewählt wird und für die neuen Indizes l festgelegt wird, dass $f'(m_l, h_l) = s_{U_\psi}$ für ein $s_{U_\psi} \in U_\psi \in \text{Bel}_{f(s)}^{\nu(\alpha)}$ mit $\mathcal{M}, s_{U_\psi} \models \neg\varphi$, dann zeigt f' , dass $\mathcal{M} b'$ getreu ist.

Die Fälle der *des*- und \neg -*des*-Operatorbeseitigungsregeln verlaufen analog zu den *bel*- und \neg -*bel*-Regeln.

Da der Intensionsoperator relational interpretiert wird, ist der Beweis für die Regeln $\alpha \text{ int}:$ und $\neg\alpha \text{ int}:$ ähnlich zum Beweis der Regeln \Box und $\neg\Box$. Angenommen die Regel $\alpha \text{ int}:$ wird auf einen Ausdruck $\alpha \text{ int}:\varphi$, s angewandt. Dann wird der Zweig b mit φ , s_k für alle $s_k \in I_s^\alpha$ auf b zu b' erweitert. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(s) \models \alpha \text{ int}:\varphi$. Das bedeutet, dass für alle $s' \in \text{Int}_{f(s)}^{\nu(\alpha)}$ gilt, dass $\mathcal{M}, s' \models \varphi$. Da $f(s_k) \in \text{Int}_{f(s)}^{\nu(\alpha)}$ für alle $s_k \in I_s^\alpha$ auf dem Zweig b , zeigt die Funktion f , dass $\mathcal{M} b'$ getreu ist.

Angenommen es wird die Regel $\neg\alpha \text{ int}:$ auf einen Ausdruck $\neg\alpha \text{ int}:\varphi$, s auf b angewandt, so dass der Zweig mit $m_k \in h_k$, $\neg\varphi$, (m_k, h_k) und $(m_k, h_k) \in I_s^\alpha$ für einen neuen Index k zu b' erweitert wird. Da f zeigt, dass $\mathcal{M} b$ getreu ist, gilt $\mathcal{M}, f(s) \models \neg\alpha \text{ int}:\varphi$, d.h. es gibt ein $(m', h') \in \text{Int}_{f(s)}^{\nu(\alpha)}$ und $\mathcal{M}, (m', h') \models \neg\varphi$. Wähle f' wie f und setze für den neuen Index k $f'(m_k, h_k) = (m', h')$, dann ist $\mathcal{M}, f'(m_k, h_k) \models \neg\varphi$ und f' zeigt, dass $\mathcal{M} b'$ getreu ist.

Es bleibt zu zeigen, dass f zeigt, dass \mathcal{M} unter den Strukturregeln b getreu bleibt. Für die Regeln aus Tabelle 6.1 gilt dies offensichtlich. Wenn $\alpha = \beta$ auf b , gilt $\nu(\alpha) = \nu(\beta)$ unabhängig von f . Für die Strukturregeln ID , ID_{ref} , ID_{sym} , ID_{tran} , $M_{antisym}$, M_{tran} , und ONE aus Tabelle 6.2 ist dies ebenfalls offensichtlich.

Für die Regeln H_{lin} , H_{uni} , H_{con} resultiert die Behauptung daraus, dass f den Zweig in die BT Struktur einbettet. Sei beispielsweise für die Regel H_{lin} $m_i \triangleleft m_k$, $m_l \triangleleft m_k$ auf dem Zweig, dann gilt für alle $m_i \in h$, $m_k \in h'$, $m_l \in h''$ auf dem Zweig, dass $\pi_1(f(m_i, h)) \leq \pi_1(f(m_k, h'))$ und $\pi_1(f(m_l, h'')) \leq \pi_1(f(m_k, h'))$. Aufgrunddessen, dass das Modell auf einer BT Struktur basiert, gilt entweder $\pi_1(f(m_i, h)) = \pi_1(f(m_l, h''))$ oder $\pi_1(f(m_i, h)) < \pi_1(f(m_l, h''))$ oder $\pi_1(f(m_l, h)) < \pi_1(f(m_i, h''))$, so dass nach einer

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Regelanwendung H_{lin} f zeigt, dass \mathcal{M} einer Extension b' von b getreu ist. Für die Regel H_{uni} gilt dies offensichtlich. Für eine Regelanwendung H_{con} sei $m_i \in h_k, m_j \in h_l$ auf dem Zweig, dann gilt für $\pi_1(f(m_i, h_k))$ und $\pi_1(f(m_j, h_l))$, dass die beiden Momente entweder in der BT Struktur vergleichbar sind. Für diese beiden Erweiterungen zeigt f , dass \mathcal{M} b getreu ist. Oder sie sind nicht vergleichbar, dann gibt es einen früheren Moment m' auf (M, \leq) , für den gilt, dass $m' < \pi_1(f(m_j, h_l))$ und $m' < \pi_1(f(m_i, h_k))$, so dass f' für alle Ausdrücke auf b wie f definiert ist und $f'(m_n, h_k) = (m', \pi_2(f(m_i, h_k)))$ bzw. $f'(m_n, h_l) = (m', \pi_2(f(m_j, h_l)))$ gewählt wird. Dann zeigt f' , dass \mathcal{M} dieser Erweiterung von b getreu ist.

Für die Regeln in Tabelle 6.3 ist Korrektheit zum Teil bereits in [159] gezeigt worden. D.h. die Regeln *REF*, *SYM*, *TRAN* sind korrekt, da für jeden Akteur $v(\alpha)$ und jedem Moment m im Modell gilt, dass $Ch_m^{v(\alpha)}$ eine Äquivalenzrelation ist, somit zeigt eine Funktion f , die Definition 16.2 erfüllt, dass das Modell auch einer Erweiterung nach Regelanwendung *REF*, *SYM*, *TRAN*, dem Zweig getreu ist. Für die Regel *UNI* gilt dies ebenso. Für $m_l = m_j$ auf b und für jede Funktion f , die zeigt, dass \mathcal{M} b getreu ist, gelten auch $f(m_l, h) = f(m_j, h)$ und somit $Ch_{\pi_1(f(m_j, h))}^{v(\alpha)} = Ch_{\pi_1(f(m_l, h))}^{v(\alpha)}$.

Desgleichen wurde in [159] gezeigt, dass für die Regel *IND* Korrektheit gilt. Es gibt ein $h' \in H_{\pi_1(f(m, h_i))}$, für das gilt, dass $h' \in Ch_{\pi_1(f(m, h_i))}^{v(\alpha_1)}(\pi_2(f(m, h_i)))$ für alle $i = 1, \dots, k$, da \mathcal{M} eine BT + AC Struktur zugrunde liegt, die die Unabhängigkeitsbedingung erfüllt, vgl. Abschnitt 1.2, S.16, bzw. [17, 18, 19]. Setze f' wie f und $f'(m, h_n) = (\pi_1(f(m, h_i)), h')$, dann zeigt f' , dass \mathcal{M} b getreu ist.

Angenommen für einen Ausdruck $m \in h$ auf b wird die Regel *SER* angewandt, so dass Ausdrücke der Form $(m_l, h_l) \in I_{(m, h)}^\alpha$, $m_l \in h_l$, $m_k \in h_k$ und $I_{(m, h)}^\alpha \in N_{(m_k, h_k)}$ auf dem Zweig für neue Indizes l, k eingeführt werden. Da \mathcal{M} b getreu ist, existiert mindestens eine Situation $s' \in Int_{f(m, h)}^{v(\alpha)}$ und eine Situation s'' mit $Int_{f(m, h)}^{v(\alpha)} \in N_{s''}$. Setze f' für alle Ausdrücke auf b wie f und $f'(m_l, h_l) = s'$ und $f'(m_k, h_k) = s''$, dann zeigt f' , dass \mathcal{M} der Erweiterung von b getreu ist.

Für die Regel *ENV* ist die Behauptung ebenfalls offensichtlich. Sie wird auf Ausdrücke der Form $s_i \in \{s_\varphi\}$ angewandt, so dass b um den Ausdruck φ , s_i erweitert wird. Nach Definition 16.3-5 zeigt f , dass \mathcal{M} dieser Erweiterung getreu ist. □

Mittels dieses Lemmas lässt sich die Korrektheit des Tableauekalküls zeigen.

Theorem 18 Falls $\Delta \not\models \varphi$, dann gilt $\Delta \not\vdash \varphi$.

BEWEIS: Falls $\Delta \not\models \varphi$, dann existiert ein Modell \mathcal{M} und eine Situation s in \mathcal{M} , so dass für alle $\psi \in \Delta$ gilt, dass $\mathcal{M}, s \models \psi$ aber $\mathcal{M}, s \not\models \varphi$. Es wird ein Tableau erzeugt, das

aus der Formelmenge $\Delta \cup \{\varphi\}$ seinen Wurzelknoten bildet, $\Delta^0 \cup \{\neg\psi, (m, h_0)\} \cup \{m \in h_0, m \triangleleft m_0, m_0 \in h_0\}$, so dass das Modell dem Zweig getreu ist, der nur aus dem Wurzelknoten besteht. Zum Beispiel zeigt die Funktion f die Getreue des Modells, die $f(m, h_0) = s$ und $s = f(m_0, h_0)$ abbildet. Nach vorangegangenem Lemma ist das Modell mindestens einer Erweiterung des Zweiges nach jeder Regelanwendung getreu. D.h. selbst wenn das Tableau vollständig ist, gibt es einen Zweig b' , dem das Modell \mathcal{M} getreu ist. Dann kann nicht jeder Zweig geschlossen sein. Denn wäre b' geschlossen, wären Ausdrücke der Form $\chi, s', \neg\chi, s'$ auf b' . Da \mathcal{M} b' getreu ist, gibt es eine Funktion f' , so dass $\mathcal{M}, f'(s') \models \chi$ und $\mathcal{M}, f'(s') \models \neg\chi$. Dies ist ein Widerspruch. Infolgedessen kann es kein vollständiges Tableau für den Wurzelknoten $\Delta^0 \cup \{\neg\psi, (m, h_0)\} \cup \{m \in h_0, m \triangleleft m_0, m_0 \in h_0\}$ geben, dessen Zweige alle geschlossen sind. Nach Definition 14 bedeutet das, dass φ aus Δ nicht syntaktisch ableitbar, d.h., $\Delta \not\vdash \varphi$. \square

Um die Vollständigkeit zu zeigen, d.h. dass $\Delta \models \varphi \Rightarrow \Delta \vdash \varphi$ impliziert, wird für ein beliebiges, vollständiges und offenes Tableau ein Modell anhand eines offenen Zweiges b des Tableaus definiert und gezeigt, dass dieses durch den Zweig b induzierte Modell die Formeln erfüllt, die auf dem Zweig b vorkommen.

Dieses Modell zu definieren, erfordert zuvor einige Erklärungen bzw. Einführungen weiterer Symbole für bestimmte Mengen bzw. Funktionen. Sei $M = \{m \mid m \in h_i \text{ auf } b\}$ die Menge der auf dem Zweig b vorkommenden Informationen bezüglich der Momente. Die Regeln $ID_{ref}, ID_{sym}, ID_{tran}$ in Tabelle 6.2 garantieren im Falle eines vollständigen Tableauezweiges b , dass auf M durch $m_i = m_k$ eine Äquivalenzrelation definiert wird. D.h. $m_i \in [m]$ gdw. $m_i = m$ auf dem Zweig b vorkommt. Sei $Tree$ die Menge der Äquivalenzklassen $Tree = \{[m] \mid m \in h \text{ auf } b\}$. Somit lassen sich die Äquivalenzklassen durch die Momente im Modell interpretieren. Sei $HT = \{h_l \mid m \in h_l \text{ kommt auf } b \text{ vor}\}$. Dann ist mit eben definierter Äquivalenzrelation auf M eine Äquivalenzrelation auf HT gegeben. Für zwei $h_l, h_k \in HT$ gilt, dass $h_l \equiv h_k$ gdw. für alle $m \in M$ gilt, dass $m \in h_l$ auf b gdw. $m \in h_k$ auf b vorkommt. Die Äquivalenzklassen werden mit $[h_l]$ für alle $h_l \in HT$ notiert.

Die Umgebungen der Nachbarschaften werden aus den folgenden, wohldefinierten Mengen gebildet: $U_{([m],[h])}^\varphi = \{([m_l],[h_l]) \mid (m_l, h_l) \in (m_i, h)_\varphi, m_i = m \text{ auf } b\}$. Nach Regel *ONE* ist mit $m_i = m$ und $m \in h$ auch $m_i \in h$ auf einem vollständigen Zweig. Da $[m]$ eine Äquivalenzklasse repräsentiert, gilt, falls $m = m_k$ auf b , dass für alle h und φ , die auf b vorkommen, $U_{([m],[h])}^\varphi = U_{([m_k],[h])}^\varphi$. Die Regel *ENV* sichert, dass für alle $([m'], [h']) \in U_{([m],[h])}^\varphi$ gilt, dass $\varphi, (m', h')$ auf b ist.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Aufgrund von Regel *UNI* gilt, dass für alle $m_l \in [m_k]$ gilt, dass $h_i \triangleleft_{m_l}^\alpha h_k$ auf b genau dann, wenn $h_i \triangleleft_{m_k}^\alpha h_k$ auf b . Daher können die Relationen $\triangleleft_{[m]}^\alpha$ wie folgt definiert werden: $[h_i] \triangleleft_{[m]}^\alpha [h_k]$ gdw. ein Ausdruck $h_i \triangleleft_{m_j}^\alpha h_k$ für ein beliebiges $m_j \in [m]$ und für beliebige $m_j \in h_i, m_j \in h_k$ auf b vorkommt. Damit ist es möglich, das durch einen vollständigen Zweig b induzierte Modell wie folgt zu definieren.

Definition 19 Sei b ein offener Zweig in einem vollständigen Tableau. Dann wird das Modell $\mathcal{M}_b = (Tree, \leq, \mathcal{A}, \bar{N}, Choice, Bel, Des, Int, v)$ als durch b induziert bezeichnet, wobei die folgenden Definitionen gelten.

1. $Tree := \{ [m] \mid m \in h \text{ auf } b \}$
2. $\leq := trcl \left\{ ([m_i], [m_j]) \mid m_i \triangleleft m_j \text{ auf } b, [m_i], [m_j] \in Tree \right\}$ ¹⁹
3. $\mathcal{A} := \{ \alpha \mid \alpha \text{ ist eine Akteursvariable auf } b \}$ ²⁰
4. $Choice_{[m]}^\alpha := \left\{ ([h_k], [h_l]) \mid h_k \triangleleft_{[m]}^\alpha h_l \right\}$ für alle $\alpha \in \mathcal{A}$ und $[m] \in Tree$.
5. $N_{([m],[h])} := \left\{ U_{([m'],[h'])}^\varphi \mid \{(m', h')_\varphi\} \in N_{(m,h)} \text{ auf } b \right\}$ für alle $m \in h$ auf b .
6. $Bel_{([m],[h])}^\alpha := \left\{ U_{([m],[h])}^\varphi \mid \{(m, h)_\varphi\} \in B_{(m,h)}^\alpha \text{ auf } b \right\}$ für alle α und $m \in h$ auf b .
7. $Des_{([m],[h])}^\alpha := \left\{ U_{([m],[h])}^\varphi \mid \{(m, h)_\varphi\} \in D_{(m,h)}^\alpha \text{ auf } b \right\}$ für alle α und $m \in h$ auf b .
8. $Int_{([m],[h])}^\alpha := \left\{ ([m'], [h']) \mid (m', h') \in I_{(m,h)}^\alpha \text{ auf } b \right\}$ für alle $m \in h$ und α auf b .
9. a) $v(\alpha) := \alpha$.
b) $v(p) := \{ ([m], [h]) \mid p, (m, h) \text{ kommt auf } b \text{ vor} \}$.
c) $([m], [h]) \notin v(p)$, falls $\neg p, (m_k, h_i)$ für ein $h_i \in [h]$ und für $m_k \in [m]$ auf b vorkommt.
d) Für alle anderen Literale und Situationen über $(Tree, \leq)$ ist die Wahl von v beliebig.²¹

¹⁹Hierbei steht *trcl* für die Bildung des reflexiven und transitiven Abschlusses der Relation.

²⁰Ich nehme an, dass zwei unterschiedliche Akteursvariablen zwei verschiedene Akteure repräsentieren, so dass keine Gleichheit von Akteuren auf dem Zweig möglich ist. Anderenfalls würde \mathcal{A} die Repräsentanten der Äquivalenzklassen beinhalten.

²¹Unter der Einschränkung, dass v geeignet gewählt wird, so dass kein Widerspruch entsteht.

Aufgrund der Strukturregeln M_{tran} , $M_{antisym}$, H_{lin} , H_{uni} und H_{con} in Tabelle 6.2 sowie des reflexiven, transitiven Abschlusses ist die geordnete Menge $(Tree, \leq)$ eine BT Struktur. Aufgrund der Vollständigkeit des Zweiges ist die Relation \leq transitiv und antisymmetrisch. Für zwei beliebige Momente $[m_i]$, $[m_k]$ gilt nach Regel H_{con} , dass entweder $[m_i] \leq [m_k]$ oder $[m_k] \leq [m_i]$ oder ein weiterer Moment $[m_n]$ auf dem Zweig ist, so dass $[m_n] \leq [m_i]$ und $[m_n] \leq [m_k]$ gelten. Damit ist die Definition 1 einer BT Struktur erfüllt, siehe Abschnitt 1.1 S.13.

Aufgrund der Vollständigkeit des Zweiges und der Regeln ONE , H_{con} , H_{uni} ist die Menge der Geschichten H über der BT Struktur $(Tree, \leq)$ eineindeutig und strukturerhaltend auf die Menge $\{[h_i] \mid h_i \text{ kommt auf } b \text{ vor.}\}$ abbildbar, so dass diese mit diesen Symbolen identifiziert werden. Die Menge der Geschichten, die durch einen Moment verlaufen, wird mit der Menge $H_{[m]} = \{[h_k] \mid m \in h_k \text{ kommt auf } b \text{ vor}\}$ identifiziert. Die Menge der Situationen wird mit $S = \{([m], [h]) \mid m \in h \text{ auf } b\}$ gekennzeichnet. Somit sind eben getroffene Definitionen für das induzierte Modell wohldefiniert.

Durch die Strukturregeln in Tabelle 6.3 vor allem durch die Regeln REF , SYM und $TRAN$ und da das Tableau vollständig ist, ist die Relation $\simeq_{[m]}^\alpha$ eine Äquivalenzrelation auf $H_{[m]}$, wobei $H_{[m]}$ die Menge der Geschichten durch den Moment $[m]$ ist. Daher ist $Choice_{[m]}^\alpha([h])$ die Äquivalenzklasse von h auf b . Nach Regel IND erfüllt die so erhaltene BT + AC Struktur die Unabhängigkeitsbedingung, vgl. [159].

Aufgrund von Regel SER sind die Mengen $Int_{([m],[h])}^\alpha$ für jeden Akteur α und jede Situation $([m], [h])$ nicht leer. Im Folgenden bedeutet s eine Situation aus S , d.h. $s = ([m], [h])$. Das folgende Lemma ist für jeden offenen Zweig in einem vollständigen Tableau gültig.

Lemma 20 *Sei b ein offener Zweig in einem vollständigen Tableau und $\mathcal{M}_b = (Tree, \leq, \mathcal{A}, \bar{N}, Choice, Bel, Des, Int, \nu)$ das durch den Zweig b induzierte Modell. Es gilt,*

$$\text{falls } \varphi, s \text{ auf } b, \text{ dann } \mathcal{M}_b, s \models \varphi.$$

BEWEIS: Der Beweis ist ein Induktionsbeweis, wobei die Induktion über die Anzahl der in φ vorkommenden Operatoren läuft. Er verläuft für die Operatoren der Stit-Logik analog zu dem in [159] angegebenen Beweis für die *dstit* Logik. Angenommen φ wäre eine atomare Formel. Für ν gelten $([m], [h]) \in \nu(\varphi)$ nach Definition 19 und somit $\mathcal{M}_b, ([m], [h]) \models \varphi$.

Die Induktionshypothese (IH) lautet, dass für jede Formel, die von der Anzahl her weniger Operatoren beinhaltet als φ , die Aussage gilt. Der Beweis macht öfter implizit Gebrauch von der Vollständigkeit des Zweiges und somit von der Behauptung, dass

6. Vollständigkeit und Erweiterungen von *bdi-stit*

für Situationen $([m], [h])$ und $([m_i], [h])$ mit $m \in [m_i]$ gilt, dass aufgrund der Regelanwendungen *ONE* und *ID* dieselben Formeln in den Situationen wahr sind.

Sei $\varphi = \neg p$ eine Negation einer atomaren Formel und $\neg p, (m, h)$ kommt auf b vor. Nach Definition 19 gilt $\mathcal{M}_b, ([m], [h]) \not\models p$, und somit $\mathcal{M}_b, ([m], [h]) \models \varphi$.

Falls φ eine Formel der folgenden Form $\neg\neg\psi$, $\psi \wedge \chi$ und $\neg(\psi \wedge \chi)$ ist, folgt die Behauptung aus (IH) und der Vollständigkeit des Zweiges.

Sei $\varphi = \Box\psi$. Falls $\varphi, (m, h)$ auf b vorkommt, dann gilt für jedes h_k mit $m \in h_k$ ebenfalls aufgrund der Vollständigkeit von b , dass $\psi, (m, h_k)$ auf b ist, und aufgrund von (IH) folgt, dass $\mathcal{M}_b, ([m], [h_k]) \models \psi$. Nach Definition der Menge der Geschichten gilt, dass jede Geschichte durch $H_{[m]}$ auf eine Äquivalenzklasse $[h_k]$ zurückzuführen ist, und somit $\mathcal{M}_b, ([m], [h]) \models \Box\psi$.

Sei $\varphi = \neg\Box\psi$. Falls $\varphi, (m, h)$ auf b ist, dann gibt es einen Ausdruck $m \in h_k$ auf b , wobei $\neg\psi, (m, h_k)$ ebenfalls auf b ist. Nach (IH) und Definition von $H_{[m]}$ gilt $[h_k] \in H_{[m]}$ und $\mathcal{M}_b, ([m], [h_k]) \models \neg\psi$, woraus folgt, dass $\mathcal{M}_b, ([m], [h]) \models \neg\Box\psi$.

Sei $\varphi = \alpha dstit : \psi$. Falls $\varphi, (m, h)$ auf b , dann sind für alle h_k mit $h \triangleleft_m^\alpha h_k$ auf b die Ausdrücke $\psi, (m, h_k)$ ebenfalls auf b . Außerdem gibt es eine Geschichte h_l , für die $m \in h_l$ und $\neg\psi, (m, h_l)$ auf b vorkommen. Nach (IH) und Definition von \mathcal{M}_b folgt, dass $\mathcal{M}_b, ([m], [h]) \models \alpha dstit : \psi$.

Sei $\varphi = \neg\alpha dstit : \psi$. Falls $\varphi, (m, h)$ auf b , dann gibt es für den vollständigen Zweig zwei Möglichkeiten. Entweder gibt es eine Geschichte h_k mit $h \triangleleft_m^\alpha h_k$ und $\neg\psi, (m, h_k)$ auf b . Oder es gilt für alle h_l mit $m \in h_l$ auf b , dass $\psi, (m, h_l)$ auf b ist. Im ersten Fall ist die Bedingung (+) des *dstit*-Operators nicht erfüllt. Es folgt aufgrund der Definition von \mathcal{M}_b , dass $[h_k] \in Choice_{[m]}^\alpha([h])$ und aufgrund von (IH), dass $\mathcal{M}_b, ([m], [h]) \models \neg\alpha dstit : \psi$. Im anderen Fall ist die Bedingung (-) nicht erfüllt, da für alle $[h_l] \in H_{[m]}$ gilt, dass $\mathcal{M}_b, ([m], [h_l]) \models \psi$. Es folgt wiederum $\mathcal{M}_b, ([m], [h_l]) \models \neg\alpha dstit : \psi$. In beiden möglichen Fällen gilt somit die Behauptung.

Sei $\varphi = \alpha int : \psi$. Falls $\varphi, (m, h)$ auf b , dann gilt für alle $(m_i, h_i) \in I_{(m,h)}^\alpha$ auf b , dass auch $\psi, (m_i, h_i)$ auf b ist. Aufgrund von (IH) gilt für alle $([m_i], [h_i]) \in Int_{([m],[h])}^\alpha$, dass $\mathcal{M}_b, ([m_i], [h_i]) \models \psi$. Somit gilt $\mathcal{M}_b, ([m], [h]) \models \alpha int : \psi$.

Sei $\varphi = \neg\alpha int : \psi$. Falls $\varphi, (m, h)$ auf b vorkommt, gibt es einen Ausdruck $m_k \in h_k$ auf b für einen bis zu der Tiefe des Zweiges neuen Index k , so dass $\neg\psi, (m_k, h_k)$ und $(m_k, h_k) \in I_{(m,h)}^\alpha$ ebenfalls auf b vorkommen. Nach (IH) und Definition von \mathcal{M}_b sind somit $([m_k], [h_k]) \in Int_{([m],[h])}^\alpha$ und $\mathcal{M}_b, ([m_k], [h_k]) \models \neg\psi$ offensichtlich, so dass die Behauptung $\mathcal{M}_b, ([m], [h]) \models \neg\alpha int : \psi$ offensichtlich ist.

Sei $\varphi = \diamond\psi$. Falls $\varphi, (m, h)$ auf b vorkommt, dann kommen folgende Informationen $\{(m, h)_\psi\} \in N_{(m,h)}$ und $(m_l, h_l) \in \{(m, h)_\psi\}$ für einen bis dato neuen Index l ebenfalls

auf b vor. Aufgrund der Vollständigkeit des Zweiges folgt aus Regel *ENV*, dass auch $\psi, (m_l, h_l)$ auf b ist. Nach (IH) gilt $\mathcal{M}_b, ([m_l], [h_l]) \models \psi$. Ebenso sorgt Regel *ENV* dafür, dass für jeden weiteren Ausdruck $(m_k, h_k) \in \{(m, h)_\psi\}$ auf b ebenfalls $\psi, (m_k, h_k)$ auf b vorkommt. Auf diese Ausdrücke ist ebenfalls (IH) anwendbar, so dass nach Definition von \mathcal{M}_b $U_{([m], [h])}^\psi \in N_{([m], [h])}$ und für alle $([m_k], [h_k]) \in U_{([m], [h])}^\psi$ gilt, dass $\mathcal{M}_b, ([m_k], [h_k]) \models \psi$. D.h. $\mathcal{M}_b, ([m], [h]) \models \diamond\psi$.

Sei $\varphi = \neg\diamond\psi$. Falls $\varphi, (m, h)$ auf b vorkommt, dann sind für alle Ausdrücke $\{(m, h)_\chi\} \in N_{(m, h)}$ auf b die Ausdrücke $\neg\psi, (m_l, h_l), (m_l, h_l) \in \{(m, h)_\chi\}$ für χ ebenfalls auf b für jeden neu eingeführten Index l . Aufgrund der Definition der Umgebungssysteme $N_{([m], [h])}$ in \mathcal{M}_b gilt, dass für alle $U_{([m], [h])}^\chi \in N_{([m], [h])}$ ein $([m_l], [h_l]) \in U_{([m], [h])}^\chi$ existiert, für das nach (IH) $\mathcal{M}_b, ([m_l], [h_l]) \models \neg\psi$ gilt. D.h. $\mathcal{M}_b, ([m], [h]) \models \neg\diamond\psi$.

Sei $\varphi = \alpha\text{bel}:\psi$. Falls $\varphi, (m, h)$ auf b vorkommt, kommen folgende Informationen $\{(m, h)_\psi\} \in B_{(m, h)}^\alpha$ und $(m_l, h_l) \in \{(m, h)_\psi\}$ für einen neuen Index l auf b vor. Aufgrund der Vollständigkeit des Zweiges und Regel *ENV* gilt für jeden Ausdruck $(m_k, h_k) \in \{(m, h)_\psi\}$ auf b , dass auch $\psi, (m_k, h_k)$ auf b vorkommt. Auf diese Ausdrücke ist (IH) anwendbar, so dass nach Definition 19 eine nicht-leere Umgebung $U_{([m], [h])}^\psi \in \text{Bel}_{([m], [h])}^\alpha$ existiert und für alle $([m_k], [h_k]) \in U_{([m], [h])}^\psi$ gilt, dass $\mathcal{M}_b, ([m_k], [h_k]) \models \psi$. D.h. $\mathcal{M}_b, ([m], [h]) \models \alpha\text{bel}:\psi$.

Sei $\varphi = \neg\alpha\text{bel}:\psi$. Falls $\varphi, (m, h)$, dann sind für alle Ausdrücke $\{(m, h)_\chi\} \in B_{(m, h)}^\alpha$ auf b die Ausdrücke $\neg\psi, (m_l, h_l)$ und $(m_l, h_l) \in \{(m, h)_\chi\}$ ebenfalls auf b mit neu eingeführtem Index l für jedes auf b vorkommendes χ . Aufgrund der Definition der Umgebungssysteme $\text{Bel}_{([m], [h])}^\alpha$ in \mathcal{M}_b gilt, dass für alle $U_{([m], [h])}^\chi \in \text{Bel}_{([m], [h])}^\alpha$ somit ein Index l eine Situation kennzeichnet, für die gilt, dass $([m_l], [h_l]) \in U_{([m], [h])}^\chi$ und nach (IH) $\mathcal{M}_b, ([m_l], [h_l]) \models \neg\psi$. D.h. $\mathcal{M}_b, ([m], [h]) \models \neg\alpha\text{bel}:\psi$.

Der Fall $\varphi = \alpha\text{des}:\psi$ ist analog zu den Fällen $\varphi = \diamond\psi$ bzw. $\varphi = \alpha\text{bel}:\psi$. Dasselbe gilt für den Fall $\varphi = \neg\alpha\text{des}:\varphi$ und die Fälle $\varphi = \neg\diamond\psi$ bzw. $\varphi = \neg\alpha\text{bel}:\psi$. □

Mittels dieses Lemmas lässt sich folgendes Theorem beweisen, das die schwache Vollständigkeit der *bdi-stit* Logik festhält. Dies ist nur die schwache Vollständigkeit aufgrunddessen, dass $\Delta \vdash \psi$ nur für endliche Mengen Δ definiert ist.

Theorem 21 Falls $\Delta \not\vdash \psi$, dann $\Delta \not\models \psi$.

BEWEIS: Angenommen $\Delta \not\vdash \psi$. Nach Definition 14 bedeutet das, dass jedes vollständige Tableau mit Wurzelknoten $\Delta^0 \cup \{\neg\psi, (m, h_0)\} \cup \{m \in h_0, m \triangleleft m_0, m_0 \in h_0\}$ mindestens einen offenen Zweig b hat. Sei \mathcal{M}_b das durch den Zweig b induzierte Modell, dann gilt gemäß vorangegangenen Lemma, dass dieses Modell jede Formel auf dem

6. Vollständigkeit und Erweiterungen von *bdi-stit*

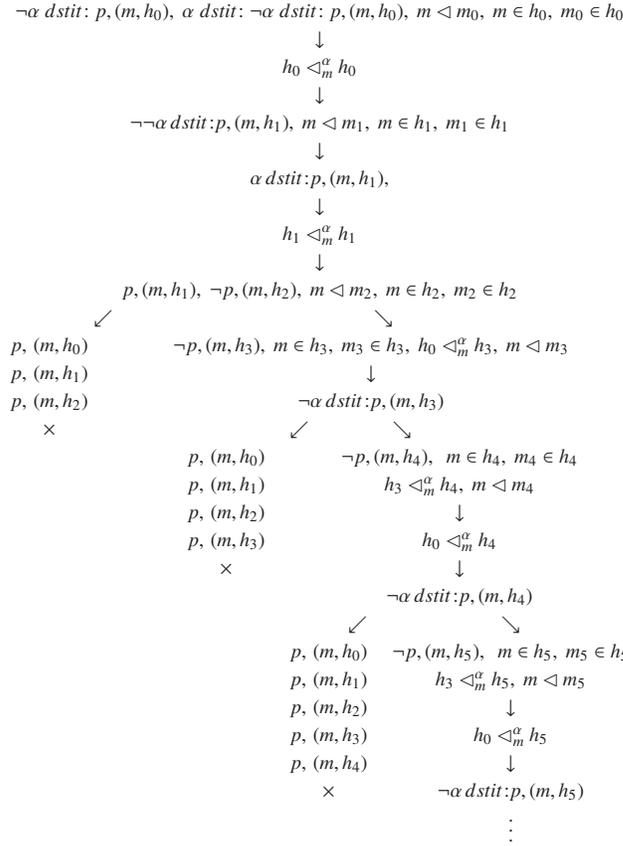


Tabelle 6.6: Ein Beispiel für ein unendliches, vollständiges, offenes Tableau

Zweig erfüllt. Insbesondere werden alle Formeln im Wurzelknoten in der Situation $([m], [h_0])$ des induzierten Modells erfüllt: $\mathcal{M}_b, ([m], [h_0]) \models \varphi$ für jede Formel $\varphi \in \Delta$ und $\mathcal{M}_b, ([m], [h_0]) \models \neg\psi$, so dass $\mathcal{M}_b, ([m], [h_0]) \not\models \psi$. Infolgedessen folgt die Behauptung. \square

Folgerung Falls es ein vollständiges Tableau zum Wurzelknoten $\Delta^0 \cup \{\neg\psi, (m, h_0)\} \cup \{m \in h_0, m \triangleleft m_0, m_0 \in h_0\}$ gibt, welches offen (geschlossen) ist, dann ist jedes vollständige Tableau zu diesem Wurzelknoten offen (geschlossen).

Somit ist die Korrektheit und die schwache Vollständigkeit des Tableauealküls nachgewiesen. Dass die *bdi-stit* Logik entscheidbar ist, wurde bereits in Abschnitt 6.1 gezeigt.

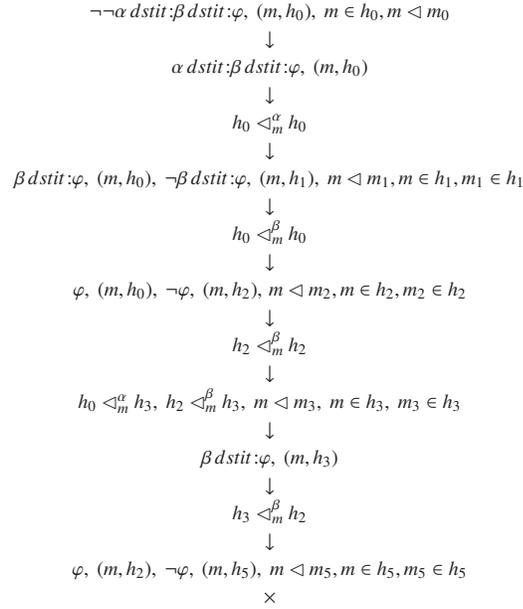


Tabelle 6.7: Handlungsunabhängigkeit der Akteure

Beispiele und Besonderheiten der *bdi-stit* Tableaux

Wie leicht zu sehen ist, kann es zu unendlichen Tableaux kommen, z.B. durch wiederholte Regelanwendung von *SER* bzw. *TRAN*. Dies ist keine Besonderheit des *bdi-stit* Tableukalküls. Die Regel, die die Transitivität einer darunterliegenden Relation sichert, führt bei normalen Operatoren zu unendlichen Tableaux [110, S.43ff]. Das Tableau in Tabelle 6.6 kann durch die Transitivitätsregel ad infinitum verlängert werden. Das Beispiel ist aus [159] entnommen.

Wie Priest in [110] festhält, ist ein offenes, unendliches Tableau kein hinreichender Grund anzunehmen, dass jedes Modell, welches ein Gegenbeispiel zu der Ableitung liefern soll, unendlich ist. Für das Beispiel in Tabelle 6.6 gibt es ein endliches Modell, welches die Ableitung $\{\alpha dstit:\neg\alpha dstit:p\} \vdash \alpha dstit:p$ als ungültig markiert. Der Grund für das unendliche Tableau ist, dass durch die $\neg dstit$ -Beseitigungsregel immer wieder eine neue Geschichte eingeführt wird, die wahläquivalent ist, in der der Akteur p jedoch nicht verursacht hat. Die Tableauregeln sehen nicht vor, dass es bereits eine Geschichte auf dem Zweig gibt, die dies sicher stellt. Daher ist es möglich, die Geschichten, die in dem Tableau mit h_4 und h_5 neu gekennzeichnet sind, mit der Geschichte h_3 zu identifizieren.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Ein endliches Modell, welches ein Gegenbeispiel zu der Ableitung ist und durch das Tableau erzeugt wird, ist $M_b = \{m, m_0, m_1, m_2, m_3\}$, $\leq = \text{trcl}\{(m, m_i) \mid i = 0, \dots, 3\}$, $\mathcal{A} = \{\alpha\}$, $C(\alpha) = \{\{h_0, h_3\}, \{h_1\}, \{h_2\}\}$ und $v(p) = \{(m, h_1)\}$, die restlichen Informationen über *Choice*, *Int*, *Bel*, *Des* etc. können beliebig, aber einem *bdi-stit* Modell entsprechend ergänzt werden. Für alle Indizes $i, j \geq 3$ befinden sich die folgenden Ausdrücke $\neg p$, (m, h_i) , $\neg \alpha \text{ dstit} : p$, (m, h_i) , $m_i \in h_i$, $m \in h_i$, $h_0 <_m^\alpha h_i$, $h_i <_m^\alpha h_j$ auf dem rechten Teilzweig b . Die eingeführten Ausdrücke $m_i \in h_i$ liefern dieselben und somit keine neuen Informationen auf dem Teilzweig. Sie sind in gewisser Weise als äquivalent zu betrachten. Ein Ausdruck $m_k \in h_i$ ist äquivalent zu einem Ausdruck $m_j \in h_l$ gdw. für die Funktion f , die die Menge $M \cup HT$ auf b mit $f(m_k) = m_j$, $f(h_i) = h_l$ und für alle anderen $w \in M \cup HT$ mit $f(w) = w$ auf sich abbildet, gilt, dass, falls die Elemente von $M \cup HT$ auf b durch ihre Bilder ersetzt werden, keine Änderung der Informationen auf b vorkommt. In dem Beispiel in Tabelle 6.6 sind nach dieser Definition die Ausdrücke $m_3 \in h_3$ äquivalent zu jedem $m_i \in h_i$ mit $i \geq 3$. Im induzierten Modell kann somit die Äquivalenzklasse von $[m_3]$ auch m_i enthalten, obwohl sie nicht auf dem Baum sind. Damit ist das Modell endlich, auch wenn der Zweig, der das Modell induziert, unendlich ist.

Die Unabhängigkeitsbedingung der Akteure sichert in der Stit-Theorie, dass Akteure ihre Möglichkeiten, etwas hervorzubringen, nicht gegenseitig einschränken. Sie sorgt zudem dafür, dass es nicht möglich ist, dass Akteure kontradiktorische Sachverhalte im selben Moment hervorbringen. Eine weitere Folge der Unabhängigkeitsbedingung ist, dass Akteure sich nicht in ein und demselben Moment gegenseitig dazu bringen können, etwas zu bewirken, sondern die Verursachung der Handlung in einem Moment immer noch beim Akteur selbst liegt. Die Unabhängigkeit der Akteuren im instantan ausgewerteten *dstit*-Operator sorgt dafür, dass die sogenannte ‘no other-agent-stit’-Formel $\alpha \text{ dstit} : \beta \text{ dstit} : \varphi$ nicht erfüllbar ist, solange die Variablen α und β durch verschiedene Akteure interpretiert werden, vgl. z.B. [19, 159]. Infolgedessen kann insbesondere kein Akteur α dafür sorgen, dass im selben Moment ein anderer Akteur β Kontrolle über seine Überzeugungen, Intentionen oder Wünsche im selben Moment ausübt, siehe hierfür das geschlossene Tableau in Tabelle 6.7. Für die Formeln φ kann jede beliebige Formel substituiert werden, z.B. auch $\varphi = \beta \text{ bel} : \psi$, $\beta \text{ des} : \psi$, $\beta \text{ int} : \psi$. Dass es möglich ist, dass *bdi-stit* Modelle konfligierende Überzeugungen ein und desselben Akteurs zulassen, zeigt das offene Tableau in Tabelle 6.8, anhanddessen ein Modell induziert wird, welches die Formel $(\alpha \text{ bel} : p \wedge \alpha \text{ bel} : \neg p)$ für ein beliebiges p erfüllt.

Nachdem festgehalten wurde, dass es einen Tableaurekalkül gibt, der vollständig ist,

6.3. Temporale Operatoren in der *bdi-stit* Logik

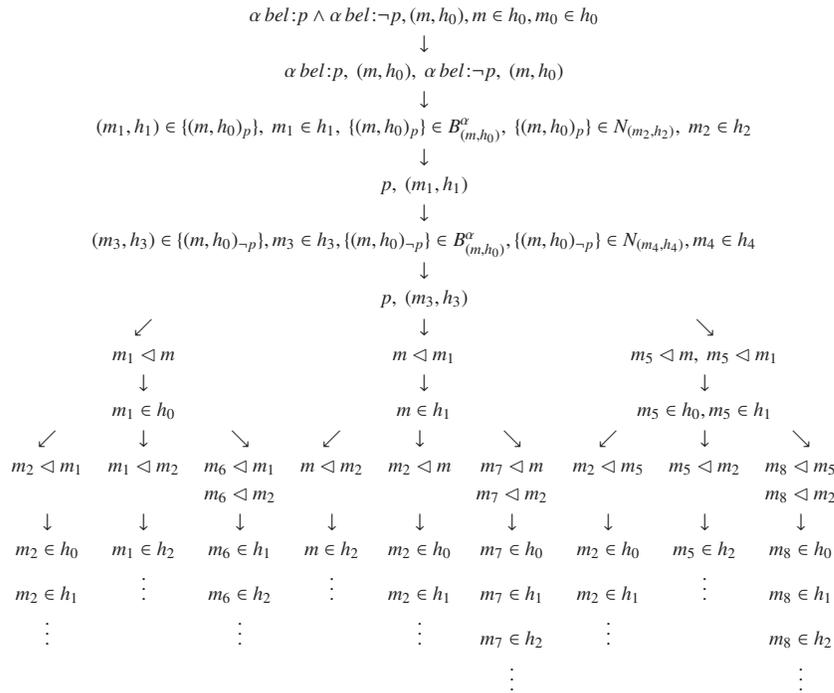


Tabelle 6.8: Offenes Tableau der konfligierenden Überzeugungen

möchte ich das Feld der Tableauregeln wieder verlassen. Im Folgenden diskutiere ich Erweiterungen von *bdi-stit* Logik. Dabei werden im folgenden Abschnitt temporale Operatoren in *bdi-stit* eingeführt. Für diese temporalen Operatoren wird gezeigt, dass es entsprechende Operatoren in *BDICTL** Logiken gibt, so dass die Generalisierungsaussage für *bdi*-Zustandsformeln aus Abschnitt 5.3 auf Pfadformeln ausgedehnt werden kann. In Abschnitt 6.4 spielen die Relationen zwischen Handlungen und Intentionen eines Akteurs eine Rolle, die in Abschnitt 4.3 motiviert wurden. Es wird ein spezifizierter Intensionsoperator eingeführt, der diese Beziehungen zwischen Handlungen und Intentionen geeignet wiedergibt.

6.3 Temporale Operatoren in der *bdi-stit* Logik

Da *bdi-stit* Modellen eine BT Struktur zugrunde liegt, ist es möglich, die üblichen, temporalen Operatoren, die in *CTL** definiert sind, ebenfalls zu definieren. Die Operatoren $F\varphi$ (irgendwann in der Zukunft φ) und $G\varphi$ (für immer in der Zukunft φ) wurden in [17, 19] in die Stit Theorie aufgenommen. Operatoren der Form $X\varphi$ (im nächsten

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Zeitpunkt φ) und $\varphi U \psi$ (solange φ bis ψ) werden in [17, 19] nicht näher betrachtet. Ein Grund, warum der X -Operator nicht aufgenommen wurde, ist, dass u.a. auch Belnap und Horty Zeit als kontinuierlich erachten. Zu jedem Moment gibt es spätere Momente. Aber es lässt sich kein nächster Moment festlegen. Zu jedem späteren Moment m gibt es einen, der früher als m aber später als der aktuelle ist. Temporale Operatoren auf BT Strukturen wurden u.a. in [116, 170, 171] untersucht. Bei der Betrachtung des X -Operators ist die Annahme von abzählbaren Zeitpunkten unumgänglich. Falls man die Einschränkung trifft, dass jede Momentmenge einer beliebigen Geschichte in einer BT Struktur abzählbar und nicht kontinuierlich sei, ist eine Nachfolgerfunktion festlegbar. Dem würden Horty und Belnap u.a. nicht zustimmen. Die Interpretationen von CTL^* -Operatoren in dieser abzählbaren BT Struktur ergeben sich von selbst.²²

Die Nachfolgerfunktion *succ* lässt sich aus einer solchen abzählbaren BT Struktur ablesen. Sei (M, \leq) eine abzählbare BT Struktur und S die Menge aller Situationen über ihr. Da die Momentmenge einer Geschichte nun abzählbar und linear geordnet ist, existiert zu jeder Situation (m, h) ein Moment mit $succ(m, h) := \min\{m' \mid m' \in h, m < m'\}$. Existiert $\min S = m''$, so gilt für alle $m' \in S$ entweder $m'' = m'$ oder $m'' < m'$. Es ist folglich möglich, eine Funktion mit $succ : S \rightarrow M$ zu definieren. Der k -te Nachfolger wird dann mit k -facher Anwendung von *succ* auf den Moment und die wiederholt gleichbleibende Geschichte angegeben, $succ^k(m, h) = \underbrace{succ(\dots(succ(m, h), h) \dots, h), h)}_{k \dots \text{ mal}}$. Sollte zugelassen werden, dass es endliche Geschichten gibt, so kann es Momente geben, die keinen Nachfolger haben. Die Nachfolgerfunktion wäre partiell. Sie wird eine funktionale Relation genannt. Im Weiteren werde ich jedoch generell von der Nachfolgerfunktion sprechen, auch wenn sie möglicherweise nicht überall definiert ist. Mithilfe dieser Funktion lassen sich die Erfüllbarkeitsbedingungen in einem *bdi-stit* Modell für die temporalen Operatoren wie folgt angeben. Sie lehnen sich entsprechend an die Definitionen aus [50, 51, 53] an.²³

Definition 22 *Gegeben sei eine abzählbare BT Struktur und ein Modell \mathcal{M} über dieser Struktur.*

$$\mathcal{M}, (m, h) \models X\varphi \quad \text{gdw. es gibt einen Moment } m' \text{ mit } succ(m, h) = m' \\ \text{und } \mathcal{M}, (m', h) \models \varphi.$$

²²Vergangenheitsoperatoren zu betrachten, ist ebenso möglich, aber nicht Gegenstand dieser Arbeit. Eine Erweiterung von CTL^* um Vergangenheitsoperatoren $PCTL^*$ wurde in z.B. [115] untersucht.

²³Die Wahrheitsbedingungen für den U -Operator variieren in den genannten Veröffentlichungen dahingehend, ob $\psi U \varphi$ eine Situation erfordert, in der ψ und φ wahr sind. Ich folge der Konvention in [50].

6.3. Temporale Operatoren in der bdi-stit Logik

$$\mathcal{M}, (m, h) \models (\psi U \varphi) \text{ gdw. es gibt ein } k \geq 0: \mathcal{M}, (\text{succ}^k(m, h), h) \models \varphi.$$

$$\text{und für alle } 0 \leq j < k: \mathcal{M}, (\text{succ}^j(m, h), h) \models \psi$$

Die Operatoren $F\varphi$ und $G\varphi$ ergeben sich als Abkürzungen für $F\varphi = \top U \varphi$ und demzufolge $G\varphi = \neg(\top U \neg\varphi)$, vgl. [17, 19]. Erachtet man die Zeit als kontinuierlich, verzichtet man infolgedessen auf den X -Operator. Die Wahrheitsbedingung des U -Operators lautet dann wie folgt.

$$\mathcal{M}, (m, h) \models (\psi U \varphi) \text{ gdw. es gibt ein } m < m': \mathcal{M}, (m', h) \models \varphi$$

$$\text{und für alle } m \leq m'' < m': \mathcal{M}, (m'', h) \models \psi.$$

Heutzutage gehen die Meinungen auseinander, ob die Branching Time Struktur notwendigerweise auf kontinuierlicher Zeit oder diskreten Zeitpunkten basieren sollte. In vielen Logiken wird die Zeit als diskret angenommen. Es werden ausgezeichnete Zeitpunkte in der Zeit betrachtet, in denen die Formeln ausgewertet werden, um einen nächsten Moment definieren zu können. Eine abzählbare oder diskrete BT Struktur soll sicherlich nicht ausdrücken, dass der damit modellierte Geschichtsverlauf aus abzählbar vielen Zeitpunkten oder Situationen besteht. Es soll vielmehr ausdrücken, dass die Auswertung von Formeln immer nur zu bestimmten Zeitpunkten erfolgt, so dass eine Beschreibung der Zeit sich auf einzelne Zeitpunkte stützen sollte. In Abschnitt 3.2 wurde argumentiert, dass Handlungen als Ereignisse zwar Zeiträume einnehmen, jedoch die Beschreibung, ob eine Handlung stattgefunden hat, erst in dem mit bzw. nach der Handlung erreichten Zustand bewertet werden kann. Demnach kann zu einem bestimmten Zeitpunkt beschrieben werden, dass eine Handlung vorliegt, auch wenn die Handlung selbst sich über einen bestimmten Zeitraum erstreckt. Die punktuelle Auswertung von Formeln, die Akteuren Handlungen zuschreiben, ist somit unproblematisch. Desgleichen gilt für Zustände. Da ein mentaler Zustand generell in einem Moment vorliegt oder nicht, ist die Bewertung der Beschreibung, ob der Zustand vorliegt oder nicht, ebenfalls punktuell möglich, vgl. Abschnitt 2.1.

In *BDICTL** wird nicht von diskreten Zeitpunkten, sondern von aufeinanderfolgenden Zustände gesprochen, die u.U. wiederkehren. Dies widerspricht dem von mir eingeführten Begriff des Zustandes. Es gäbe die Möglichkeit zu sagen, dass bestimmte Typen von Zuständen zu verschiedenen Zeitpunkten bzw. in verschiedenen Situationen wieder instanziiert werden. Jedoch ist auch dies unplausibel. Ist der Zustandstyp zu umfassend, d.h. würde er durch viele Sachverhalte charakterisiert, ist es unwahrscheinlich, dass er immer und immer wieder eintritt. Ist er nicht umfassend genug, ist unklar, inwieweit er einen Geschichtsverlauf in einer Welt beschreiben kann. Ob eine

6. Vollständigkeit und Erweiterungen von *bdi-stit*

endliche Anzahl von Zustandstypen einen Geschichtsverlauf umfassend beschreibt, wage ich zu bezweifeln. Spräche man von abzählbar unendlichen vielen verschiedenen Zustandstypen, ist die Annahme von wiederkehrenden Zustandstypen überflüssig. Dies könnte ebenso durch nicht wiederkehrende Situationen dargestellt werden.

Dennoch liefert das Bild von aufeinanderfolgenden Zuständen, dass der *BDICTL** zugrunde liegt, eine abzählbare Struktur. Welt-Zustands-Paare stellen die Menge an Situationen zur Verfügung, aus der die Momente und Pfade in einer Welt gebildet werden. Pfade sind als Aneinanderreihung von Zuständen einer Welt definiert und können ebenso wie ein Geschichtsverlauf in einer BT Struktur als Verlauf über die Zeit verstanden werden, vgl. Abschnitt 5.1.

Es gibt die Ansicht, dass es nicht Zustände bzw. Zustandstypen sind, die aneinandergereiht einen Geschichtsverlauf beschreiben. Bei Cohen und Levesque ist ein Geschichtsverlauf eine Aneinanderreihung von Ereignissen (bzw. Ereignistypen). Der Zeitverlauf ist linear geordnet und eine Kette von Ereignissen [47, S.222]. Die Beschreibung der Geschichte über wiederkehrende Ereignisse (Typen) ist jedoch ebenso unplausibel wie die Beschreibung der Geschichte als Kette von Zuständen (Typen). Sicherlich gibt es Modelle von bestimmten Bereichen, in denen immer wieder dieselben Zustandstypen auftreten und eine solche Beschreibung sinnvoll ist, z.B. eine Maschine, die eine begrenzte Anzahl von Zuständen hat, in denen sie sich befinden kann. Für menschliche Akteure ist dies unplausibel. Dass ein Akteur immer wieder in denselben Typ von Zustand kommt und dort wieder dieselben möglichen Verläufe hat, scheint mir als Beschreibung von menschlichen Akteuren unangebracht. Wir altern, wir lernen hinzu. Wir unterliegen somit Veränderungen, die nicht reversibel sind. Anzunehmen, dass es möglich ist, denselben Zustand noch einmal heraufzubeschwören, mag manchmal wünschenswert sein. Es ist aber kein realitätsnahes Modell für menschliche Akteure, auch nicht für idealisierte. Selbst wenn man als Argument anführt, eine Beschreibung bzw. Modellierung sollte möglichst einfach gehalten werden. So ist die Frage, wieviel Vereinfachung nötig ist.

Dass nur eine endliche Anzahl von verschiedenen Zuständen nötig ist, um die Handlungen von Akteuren zu beschreiben, könnte sicherlich dadurch motiviert werden, dass Akteure nur endliche viele Handlungsalternativen haben. Wiedergegeben wird dies in der Logik durch die Annahme, dass die Menge an atomaren Formeln endlich sein kann und die instantanen Handlungsoperatoren durch Relationen ausgewertet werden, die nur endlich viele Handlungsmöglichkeiten offen lassen, vgl. das Axiom endlich vieler Handlungsmöglichkeiten (APC_n). Somit wird für jeden Moment möglicherweise einem Akteur nur eine endliche Anzahl von Zuständen zugeschrie-

ben. Jedoch ändert sich dies mit Einführung der temporalen Operatoren. Über einen Geschichtsverlauf hinweg ist eine (abzählbar) unendlich große Menge an Situationen gegeben, die verschiedene Formelmengen erfüllen, so dass eine Beschreibung solcher Situationen durch Wiederholung von Instanzen desselben Zustandstypes unplausibel ist. Vielmehr scheinen Situationen in verschiedenen Geschichtsverläufen, sobald sie keine gemeinsame Vergangenheit mehr aufweisen, nicht dieselben Zustände zu beinhalten. Denn ein bestehender Zustand zum aktuellen Moment wird auch immer durch den Zustand individuiert, der zuvor vorlag.

In der aktuellen Debatte der Stit-Theorie wird von einigen Vertretern nicht notwendigerweise von verschiedenen Zeitpunkten bzw. Situationen entlang eines Geschichtsverlaufes gesprochen. Broersen [28, 29] benennt „states“ als die Basis seiner Auswertung. Er unterscheidet dabei zwischen „static states“ und „dynamic states“. In eine BT Struktur übertragen, kann man einen *static state* als einen Moment und einen *dynamic state* als Moment-Geschichte-Paar verstehen. Der *dynamic state* ermöglicht somit eine Betrachtung über die Geschichte hinweg, die hierbei wiederum eine Folge von aufeinander folgenden *static states* ist. Diese Ansicht führt dazu, dass es für ihn ebenso wie in der *BDICTL** möglich ist, nur eine endliche Anzahl an Zuständen (*static states*) und dennoch eine unendliche Zeitstruktur in Formen von diesen Geschichten als Zustandsfolgen (*dynamic states*) anzunehmen. Die Bezeichnungen als „dynamic state“ und „static state“ finde ich unglücklich. In Abschnitt 2.1 ist gezeigt worden, dass Zustände weder dynamisch noch statisch sind. Im Falle des „dynamic state“ kann es sich somit in keinem Fall um einen Zustand handeln. Begriffe wie Zeitpunkt oder Moment sind als Bezeichnung für den „static state“ und Situation (Zeitpunkt, der den Geschichtsverlauf beinhaltet) für den „dynamic state“ unproblematischer. Wenn es noch möglich ist, von wiederkehrenden Situationen zu sprechen, so ist es quasi paradox von wiederkehrenden Zeitpunkten zu sprechen. Dies macht den Rückgriff von Broersen auf Ausdrücke wie „static“ und „dynamic state“ verständlich. Es ist deswegen nicht weniger problematisch. Was wiederkehrend ist, da es nicht notwendigerweise zeitlich fixiert ist, ist das Bestehen von bestimmten Sachverhalten.

Dies ist das Bild, dass in der Logik im Allgemeinen gezeichnet wird. Einer Situation, Welt, Zustand etc. wird eine Menge von Sachverhalten zugeordnet, die durch Aussagebuchstaben symbolisiert werden, die in der Situation bestehen. Die Menge aller Aussagenbuchstaben symbolisiert transiente Sachverhalte, die in der einen oder anderen Situation bestehen. Bestehenden Sachverhalten korrespondiert eine Proposition. Um einem Sachverhalt, der durch eine aussagenlogische Formel symbolisiert wird, einen Wahrheitswert zuzuordnen, ist es notwendig, das Modell, den Geschichtsverlauf

6. Vollständigkeit und Erweiterungen von *bdi-stit*

bzw. die Situation anzugeben, in der der Sachverhalt besteht. Dann ist es möglich, dass diese Sachverhalte in der Situation bzw. Welt bestehen, in der nächsten möglicherweise nicht. Diese beliebige Zuordnung von Situationen und Sachverhalten ermöglicht eine breite Auswahl an Modellen für mögliche Geschichtsverläufe und Situationen, in denen ein Großteil transienter Sachverhalte besteht, die in früheren Situationen desselben Geschichtsverlaufes möglicherweise ebenfalls bestanden haben. Dennoch können sich die Situationen unterscheiden in der Menge an nicht-transienten Sachverhalten, die in ihnen bestehen. Sie müssen nicht wie im Falle der Zustände von Broersen oder *BDICTL** wiederkehren. So zeichnet die Annahme, die einer BT Struktur zugrunde liegt, dass kein zeitlicher Moment und somit keine Situation wiederkehren kann, und es dennoch möglich ist, dass in jeder späteren Situation dieselben transienten Sachverhalte erneut bestehen, ein wesentlich plausibleres Bild.

Dass in jeder späteren Situation andere als die aktuellen Möglichkeiten offenstehen, ist in *BDICTL** nicht möglich. Wenn die Pfade in einer Welt immer wieder aus derselben endlichen Zustandsmenge gebildet werden, so gibt es in jedem wiederkehrenden Zustand immer wieder dieselben Pfade.²⁴ Von einem wiederkehrenden Zustand gehen immer wieder dieselben wiederkehrenden möglichen zukünftigen Geschichtsverläufe ab. Die Welt liefere im Kreis. Wenn in einer Welt in *CTL** ein Pfad aus denselben wiederkehrenden Zuständen besteht und alle möglichen Pfade, die mit demselben Welt-Zustands-Paar beginnen, bei jeder Auswertung berücksichtigt werden, sind in den wiederkehrenden Welt-Zustands-Paaren immer wieder dieselben Formeln erfüllt. Dies schließt sogar Formeln ein, die temporale Operatoren enthalten. Dies ist unplausibel. Es sollte Modelle einer Logik von Handlungen und Überzeugungen von Akteuren geben, die zwar die Möglichkeit bietet darzustellen, dass es wiederkehrende Geschichtsverläufe gibt. Die Modelle sollten jedoch nicht auf Beschreibungen von solchen Geschichtsverläufen beschränkt werden.

Darstellung der temporalen Operatoren X , U und der Pfadoperatoren A , E der *BDICTL in *bdi-stit***

In diesem Unterabschnitt soll festgehalten werden, dass durch die Erweiterung der *bdi-stit* Logik um temporale Operatoren eine ausdrucksstärkere Logik als *BDICTL** erhalten wird. In Abschnitt 5.3 ist bereits gezeigt worden, dass es sich bei der *bdi-stit* Logik um eine Generalisierung der *bdi*-Logik handelt. Es sind lediglich Zustandsfor-

²⁴Eine Annahme von unendlichen vielen Zuständen, die in einer Welt wiederkehren können, hat keinerlei Vorteile zu einer abzählbaren BT Struktur mit unendlicher Momentmenge.

meln betrachtet worden. Wenn nun temporale Operatoren zur *bdi-stit* Logik hinzuge-
nommen werden, zeigt sich, dass sich diese Art der Generalisierung auf die Erweite-
rung mit temporalen Operatoren erstreckt. Dafür gilt es, die Formeln der *BDICTL**
in *bdi-stit* einzubetten, die sich entlang der Pfade bzw. Geschichten auswerten las-
sen. Ich werde mich auf die Zukunftsoperatoren X, U und abgeleitet daraus auf F, G
beschränken und Vergangenheitsoperatoren außer Acht lassen.

Es lässt sich zu gegebenem *BDICTL** Modell \mathcal{M}_e ein *bdi-stit* Modell \mathcal{M}_h finden,
so dass es für jedes Welt-Zustands-Paar (w, s) bzw. jeden Pfad $(w, s), (w, s_1), \dots$ in \mathcal{M}_e
eine Situation (m, h) in dem noch zu konstruierenden *bdi-stit* Modell \mathcal{M}_h gibt, so dass
für alle Formeln φ gemäß Definition 5. (S1), (S2), (S4) gilt, dass

$$\mathcal{M}_e, (w, s) \models \varphi \quad \text{gdw.} \quad \mathcal{M}_h, (m, h) \models \varphi,$$

siehe Theorem 9 in Abschnitt 5.3. Es scheint offensichtlich, dass die Pfadoperatoren
 A und E aus der Definition der *BDICTL** in Abschnitt 5.1 sich durch die historische
Notwendigkeit und Möglichkeit in *bdi-stit* wiederfinden lassen. Für alle Formeln der
Gestalt $\varphi = A\psi$ und $\varphi = E\psi$ nach Definition 5. (S3) gilt daher

$$\mathcal{M}_e, (w, s) \models A\psi \quad \text{gdw.} \quad \mathcal{M}_h, (m, h) \models \Box\psi,$$

$$\mathcal{M}_e, (w, s) \models E\psi \quad \text{gdw.} \quad \mathcal{M}_h, (m, h) \models \Diamond\psi.$$

Die Pfade des *BDICTL** Modells werden mit den Geschichten des zu konstruierenden
bdi-stit Modells identifiziert. Für alle Pfadformeln gemäß Definition 5.(P1)–(P3) gilt

$$\mathcal{M}_e, (w, s), (w, s_1), \dots \models \varphi \quad \text{gdw.} \quad \mathcal{M}_h, (m, h) \models \varphi.$$

Konstruktion des zugehörigen *bdi-stit* Modells \mathcal{M}_h

Jede Welt $w \in W$ in *BDICTL** beinhaltet eine eigene Nachfolgerrelation R_w auf
den Zuständen S_w der Welt. Diese Relationen sind seriell, aber nicht deterministisch.
Jeder Zustand $m \in S_w$ hat mindestens einen möglichen Nachfolger. Somit sind in
verschiedenen Welten verschiedene Pfade also Folgen von Zuständen denkbar. Die
möglichen Pfade in jeder Welt und jedem Zustand ergeben sich aus dieser Relation.
Da lediglich zukunftsgerichtete Operatoren ausgewertet werden, wird die Menge aller
möglichen maximalen Pfade mit konkretem Anfangszustand, aber unendlich offener
Zukunft betrachtet. Die vorangegangenen Zustände innerhalb des Pfades sind quasi
uninteressant.

Zu einer Welt w wird eine BT Struktur wie folgt konstruiert. Sei S_w die Menge aller
Zustände in w , dann ist P_w^s für ein $s \in S_w$ die Menge aller Pfade $(w, s)(w, s_1) \dots$, die in

6. Vollständigkeit und Erweiterungen von *bdi-stit*

s beginnen und maximal fortgesetzt werden. Die Menge P_w sei die Menge aller Pfade P_w^s mit $s \in S_w$. Somit ist P_w die Menge aller Pfade, die in einem beliebigen Zustand der Welt w beginnen und maximal fortgesetzt sind. Die Menge aller Folgezustände in einem solchen Pfad sind abzählbar unendlich, auch wenn möglicherweise die Menge der Zustände in S_w nicht abzählbar ist. Wenn die Welt w fixiert ist, schreibe ich abkürzend für einen Pfad $p = (w, s_0)(w, s_1)(w, s_2) \dots$ in P_w auch $p = s_0 s_1 s_2 \dots$.

Für alle $n \in \mathbb{N}$ seien folgende Äquivalenzrelationen pre_n auf den Pfaden in P_w definiert. Zwei Pfade aus P_w^s sind pre_n -äquivalent, wenn sie mit denselben n Zuständen beginnen. Mit p_n werden die ersten n aufeinander folgenden Zustände des Pfades p bezeichnet. Für zwei pre_n -äquivalente Pfade p, p' gilt $p_n = p'_n$. Für $n \leq n'$ gilt demzufolge, falls p und p' $pre_{n'}$ -äquivalent sind, dass sie auch pre_n -äquivalent sind. Die Menge aller Äquivalenzklassen für pre_n bilden zusammen mit w die Menge der Momente M_w der gesuchten BT Struktur, die mit der Welt w assoziiert werden, $M_w = \{[p]_{pre_n} \mid n \in \mathbb{N}, p \in P_w\} \cup \{w\}$. Die Ordnung \leq_w auf dieser Menge ergibt sich von selbst. Für alle $m, m' \in M_w$ gilt

$$m \leq_w m' \text{ gdw. entweder } m = w \text{ oder } (m' \neq w \text{ und für alle } p \in m' \text{ gilt } p \in m).$$

Sei m eine pre_n -Klasse und m' eine $pre_{n'}$ -Klasse, dann folgt aus $m \leq m'$, dass $n \leq n'$. Es bleibt zu zeigen, dass (M_w, \leq_w) eine BT Struktur ist, vgl. Definition 1. Die Relation \leq_w auf M_w ist offensichtlich reflexiv, transitiv und antisymmetrisch. Die Konnexivität gilt ebenfalls trivialerweise, da $w \leq_w m$ für alle $m \in M_w$.

Die BT Struktur, die zu einem beliebigen *BDICTL** Modell konstruiert wird, sieht daher wie folgt aus. Zu jeder Welt w existiert eine BT-Struktur (M_w, \leq_w) . Somit ist die Menge aller Momente des gesuchten *bdi-stit* Modells \mathcal{M}_n die Menge

$$M = \{m \in M_w \mid w \in W\} \cup \{w_0\}$$

und die Relation \leq ergibt sich aus allen Relationen \leq_w mit

$$\leq = \text{trcl} \left(\{(w_0, w_0)\} \cup \{(w_0, w) \mid w \in W\} \cup \bigcup_{w \in W} \leq_w \right).$$

Dies ist ebenfalls eine BT-Struktur. Jede Geschichte enthält genau ein $w \in W$. Die Menge aller Geschichten H_w , die durch den Moment w verlaufen, kann in die Menge P_w eineindeutig abgebildet werden, $f(p) = h$ gdw. $p \in m'$ für alle $m' \neq w, w_0$ und $m' \in h$. Der Nachweis der Injektivität ist offensichtlich. Falls zwei $f(p) = f(p')$, dann gilt für alle $n \in \mathbb{N}$, dass $p_n = p'_n$, und somit, dass $p = p'$. Die Menge aller Geschichten H_{w_0} ist auf die Menge aller Pfade abbildbar, die in einer beliebigen Welt w

in einem bestimmten Zustand s der Welt w beginnen und maximal fortgesetzt werden. Die einzelnen Welt-Zustands-Paare lassen sich möglicherweise unendlich oft in der BT Struktur wiederfinden.

Sei $m \in M_w$ und eine Äquivalenzklasse zu pre_n , dann existiert genau eine Folge $s_1 s_2 \dots s_n$ von n Zuständen, so dass für alle $p \in m$ gilt, dass $p_n = s_1 s_2 \dots s_n$. Somit ist folgende Abbildung $g : M_w \rightarrow S_w$ mit $g(m) = s_n$ wohldefiniert. Die Abbildung g kann auf M fortgesetzt werden, indem die Welt w mit $g : M \times W \rightarrow S$ und $g(m, w) = s_n$ beigeordnet wird. Die Abbildung ist wohldefiniert, aber weder injektiv noch surjektiv. Jedem maximal fortgesetzten Pfad p in w , der in beliebigem Zustand s_0 beginnt, kann mindestens eine Geschichte h zugeordnet werden, so dass für ein $m \in h$ gilt, dass $g(m, w) = s_0$ und für alle Momente mit $m \leq m'$ und $m' \in h$ gilt, dass $p \in m'$. Mittels der Funktion g werden die restlichen Elemente des *bdi-stit* Modells $M_h = (M, \leq, \mathcal{A}', C', N', B', D', I', v')$ definiert. Sie ergeben sich aus dem *BDICTL** Modell $M_e = (W, \{S_w\}_{w \in W}, \{R_w\}_{w \in W}, \mathcal{A}, B, D, I, v)$ wie folgt

$$\begin{aligned}
 \mathcal{A}' &= \mathcal{A}, \\
 B'(\alpha, (m, h)) &= \{(m', h') \mid w \in h, (w, g(m, w), w') \in B(\alpha), w' \in h'\}, \\
 D'(\alpha, (m, h)) &= \{(m', h') \mid w \in h, (w, g(m, w), w') \in D(\alpha), w' \in h'\}, \\
 I'(\alpha, (m, h)) &= \{(m', h') \mid w \in h, (w, g(m, w), w') \in I(\alpha), w' \in h'\}, \\
 N'(m, h) &= \{U \in B'(\alpha, (m, h)), V \in D'(\alpha, (m, h)), I'(\alpha, (m, h)) \\
 &\quad \mid \alpha \in \mathcal{A}, h \in H_w, m \in M_w, w \in W\}, \\
 v'(\varphi) &= \{(m, h) \mid h \in H_w, g(m, w) = s, (w, s) \in v(\varphi)\}.
 \end{aligned}$$

Die Funktion C' sei geeignet aber frei gewählt. Da die Anzahl der Zustände eines Pfades abzählbar ist, lässt sich die *succ*-Funktion auf der BT Struktur ablesen. Sei $m \in M_w$ und m ist in der Äquivalenzklasse pre_n , so dass $g(m, w) = s_n$, wobei für gegebene Geschichte h genau ein Pfad p mit $p = p_{n-1} s_n s_{n+1} \dots$ existiert. Dann ist $succ(m, h) = m'$, wobei m' eine Äquivalenzklasse zu pre_{n+1} und für alle $p' \in m'$ gilt, dass $p'_{n+1} = p_{n-1} s_n s_{n+1}$. Insbesondere ist $p \in m'$ und $m' \in h$. Setze $succ(w_0, h) = w$, falls $w \in h$, und $succ(w, h) = (m, h)$ für $m = \min\{m' \in h \mid w < m'\}$ für alle $h \in H_{w_0}$. Die *succ*-Funktion, die durch die *BDICTL** Struktur erzeugt wird, ist somit eindeutig und total auf der Menge M und für jedes $m \in M \setminus \{w_0\}$ existiert ein Urbild unter *succ*.

Ich definiere eine Relation ι , die die Menge der Welt-Zustands-Paare (w, s) bzw. die Menge der Pfade $p = (w, s)(w, s_1) \dots$ zu den Situationen in dieser BT Struktur in Relation setzt. Dabei sind in allen zu p bzw. (w, s) in Relation stehenden Situationen über der BT Struktur dieselben Formeln wie in p bzw. (w, s) erfüllt. Und es gibt

6. Vollständigkeit und Erweiterungen von *bdi-stit*

mindestens eine Situation in $\iota(z)$. Die Menge $\iota(z)$ symbolisiert die zu z in Relation ι stehenden Situationen der BT Struktur, wobei z ein Pfad bzw. Welt-Zustands-Paar ist,

$$\begin{aligned}\iota((w, s_0)(w, s_1) \dots) &= \{ (m, h) \mid w \in h \text{ und für alle } i \in \mathbb{N}, g(\text{succ}^i(m, h), w) = s_i \}, \\ \iota(w, s) &= \{ (m, h) \mid w \in h \text{ und für } m \text{ gilt } g(m, w) = s \}.\end{aligned}$$

Daher gilt, falls $(m, h) \in \iota((w, s_0)(w, s_1) \dots)$, dass $(\text{succ}^i(m, h), h) \in \iota((w, s_i), \dots)$ für alle $i \in \mathbb{N}$. Zu jedem maximal fortsetzbaren Pfad, der in beliebigem Zustand beginnt, existiert mindestens eine Geschichte, die die gewünschte Bedingung erfüllt, siehe Definition von f . Somit ist jede Menge $\iota(z)$ nicht-leer.

Insbesondere gilt, falls $(m, h) \in \iota(w, s)$, dass ein maximaler Pfad $((w, s) \dots)$ mit $(m, h) \in \iota((w, s) \dots)$ existiert. Ebenfalls gilt, dass es für jeden Pfad $(w, s)(w, s_1) \dots$ und $(m, h) \in \iota(w, s)$ eine Geschichte h' mit $(m, h') \in \iota((w, s)(w, s_1) \dots)$ gibt, obwohl möglicherweise $(m, h) \notin \iota((w, s)(w, s_1) \dots)$. Umgekehrt gilt, falls $(m, h) \in \iota(w, s)$, dass $(m, h') \in \iota(w, s)$ für alle $h' \in H_m$.

In der Übersetzung der Zustandsformeln in die *bdi-stit* Logik gilt, dass Zustandsformeln notwendigerweise in einer Situation wahr sind. Die Menge der Situationen, die einem Zustand (w, s) zugeordnet sind, ist somit die Menge aller Situationen, denen die Pfade, die in diesem Zustand (w, s) starten, ebenfalls zugeordnet sind. Desgleichen gilt für $T \in \{B^\alpha, D^\alpha, N', I^\alpha\}$, falls $(m, h), (m', h') \in \iota(w, s)$, dass $T(m, h) = T(m', h')$. Dann gilt für jedes z und $(m, h) \in \iota(z)$ sowie alle Formeln φ , deren Hauptjunktur nach Definition 5.(S1), (S2), (S4), (P1) – (P3) gebildet wird, vgl. Theorem 9,

$$\mathcal{M}_e, z \models \varphi \text{ gdw. } \mathcal{M}_h, (m, h) \models \varphi, \quad (*)$$

und für Formeln der Form $E\varphi$ gilt, Definition 5.(S3),

$$\mathcal{M}_e, z \models E\varphi \text{ gdw. } \mathcal{M}_h, (m, h) \models \diamond\varphi. \quad (*)$$

Dies lässt sich induktiv über den Formelaufbau zeigen. Da $A\varphi \equiv \neg E\neg\varphi$ und $\Box\varphi \equiv \neg\diamond\neg\varphi$ gilt (*) für alle Formeln.

Beweis der Erfüllbarkeit der *BDICTL Formeln in dem konstruierten *bdi-stit* Modell \mathcal{M}_h**

Für atomare Formeln ergibt sich die Behauptung sowohl für Zustands- als auch für Pfadformeln aus der Definition von v und v' . Für negierte Formeln $\neg\varphi$ und Konjunktionen $\varphi \wedge \psi$ erhält man die Behauptung trivialerweise aus der Induktionshypothese. Es werden daher nur die *BDICTL**-Operatoren sowie die temporalen Operatoren und die Pfadoperatoren betrachtet.

6.3. Temporale Operatoren in der bdi-stit Logik

Sei (*) die Induktionsbehauptung und $\varphi = \alpha \text{ bel} : \psi$. Sei $z = (w, s)$, dann gilt

$$\begin{aligned}
 \mathcal{M}_e, (w, s) \models \alpha \text{ bel} : \psi & \text{ gdw } \text{für alle } w' \in W \text{ mit } w' \in B_s^w(\alpha) : \mathcal{M}_e, (w', s) \models \psi \\
 & \stackrel{(*)}{\text{gdw}} \text{für alle } w' \in B_s^w(\alpha) \text{ und alle } (m', h') \in \iota(w', s) \text{ gilt} \\
 & \quad \mathcal{M}_h, (m', h') \models \psi \\
 & \stackrel{\text{Def. } B'}{\text{gdw}} \text{für alle } (m', h') \in U, U \in B'(\alpha, (m, h)) \text{ gilt} \\
 & \quad (m', h') \in \iota(w', s), \mathcal{M}_h, (m', h') \models \psi \\
 & \text{gdw } \mathcal{M}_h, (m, h) \models \alpha \text{ bel} : \psi.
 \end{aligned}$$

Dies gilt für beliebiges $(m, h) \in \iota(w, s)$, da für alle $(m', h') \in \iota(z)$ $B_{(m,h)}^{\alpha} = B_{(m',h')}^{\alpha}$. Desgleichen gilt für Formeln der Form $\alpha \text{ des} : \psi$. Da in einem bdi-stit Modell der *int*-Operator anders interpretiert wird, wird dieser Fall $\varphi = \alpha \text{ int} : \psi$ gesondert betrachtet.

$$\begin{aligned}
 \mathcal{M}_e, (w, s) \models \alpha \text{ int} : \psi & \text{ gdw } \text{für alle } (w', s) \text{ mit } (w, s, w') \in I(\alpha) : \mathcal{M}_e, (w', s) \models \psi \\
 & \stackrel{\text{Def. } I'}{\text{gdw}} \text{für alle } (m', h') \in \iota(w', s) \text{ und } (m', h') \in I_{(m,h)}^{\alpha} \text{ gilt} \\
 & \quad \mathcal{M}_h, (m', h') \models \psi \\
 & \text{gdw } \mathcal{M}_h, (m, h) \models \alpha \text{ int} : \psi
 \end{aligned}$$

Wiederum gilt dies für beliebiges $(m, h) \in \iota(z)$. Bei der Interpretation der Pfadoperatoren X und U sei $z = ((w, s_0), (w, s_1) \dots)$ und (m, h) beliebig aus $\iota(z)$.

$$\begin{aligned}
 \mathcal{M}_e, z \models X\varphi & \text{ gdw. } \mathcal{M}_e, ((w, s_1), \dots) \models \varphi \\
 & \stackrel{(*)}{\text{gdw}} \text{für alle } (m', h') \in \iota((w, s_1) \dots) \text{ gilt } \mathcal{M}_h, (m', h') \models \varphi \\
 & \stackrel{\text{Def. } \iota}{\text{gdw}} \text{es gibt ein } m'' \text{ mit } (m'', h) \in \iota((w, s_1), \dots) \text{ und} \\
 & \quad \mathcal{M}_h, (m'', h) \models \varphi \\
 & \stackrel{(m,h) \in \text{succ}^{-1}(m'')}{\text{gdw}} \mathcal{M}_h, (m, h) \models X\varphi.
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{M}_e, z \models (\psi U \varphi) & \text{ gdw. } \mathcal{M}_e, (w, s), (w, s_1), \dots \models (\psi U \varphi) \\
 & \text{gdw. es gibt } k \geq 0 \text{ mit } \mathcal{M}_e, (w, s_k), \dots \models \varphi \\
 & \quad \text{und für } 0 \leq j < k \text{ gilt } \mathcal{M}_e, (w, s_j), \dots \models \psi \\
 & \text{gdw. es gibt } k \geq 0 \text{ und für alle } (m', h') \in \iota((w, s_k) \dots) \text{ gilt} \\
 & \quad \mathcal{M}_h, (m', h') \models \varphi \text{ sowie für } 0 \leq j < k \text{ und für alle} \\
 & \quad (m'', h'') \in \iota((w, s_j) \dots) \text{ gilt, dass } \mathcal{M}_h, (m'', h'') \models \psi \\
 & \stackrel{(**)}{\text{gdw}} \text{es gibt } k \geq 0 \text{ mit } \mathcal{M}_h, (\text{succ}^k(m, h), h) \models \varphi \text{ und} \\
 & \quad \text{für alle } 0 \leq j < k \text{ gilt } \mathcal{M}_h, (\text{succ}^j(m, h), h) \models \psi \\
 & \text{gdw. } \mathcal{M}_h, (m, h) \models (\psi U \varphi).
 \end{aligned}$$

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Zur Begründung von (**) sei gesagt, dass die Induktionsbehauptung einschließt, dass, wenn in einer Situation aus $\iota(z)$ eine Formel erfüllt ist, diese auch in allen anderen Situationen aus $\iota(z)$ erfüllt ist. Aufgrund der Definition von ι gilt, dass es ein $m''' \in h$ gibt, so dass $(m''', h) \in \iota((w, s_k), \dots)$ und somit $m''' = succ^k(m, h)$. Zudem gilt für jedes m'''' mit $m \leq m'''' < m'''$, dass $(m'''', h) \in \iota((w, s_j), \dots)$, so dass es immer ein m'''' für alle $0 \leq j < k$ mit $m'''' = succ^j(m, h)$ gibt.

Bezeichnet z einen Zustand (w, s) , dann gilt ebenso für beliebiges $(m, h) \in \iota(z)$ und für Formeln der Form $\varphi = E\psi$ die Behauptung (*).

$\mathcal{M}_e, z \models E\varphi$ gdw. es gibt einen maximalen Pfad $(w, s), \dots$ mit $\mathcal{M}_e, (w, s), \dots \models \varphi$
 gdw. es gibt einen maximalen Pfad $(w, s), \dots$ und für alle
 $(m', h') \in \iota((w, s), \dots)$ gilt $\mathcal{M}_h, (m', h') \models \varphi$
 gdw. es gibt $h'' \in H_m$ mit $\mathcal{M}_h, (m, h'') \models \varphi$
 gdw. $\mathcal{M}_h, (m, h) \models \diamond\varphi$.

Somit ist es nicht nur möglich, Wahrheitsbedingungen für temporale Operatoren ohne Erweiterung der *bdi-stit* Modelle anzugeben. Es lassen sich dieselben Formeln erfüllen, die sich in *BDICTL** durch die Verwendung der Operatoren X, U, A, E ergeben. Zusammen mit der Aussage aus Theorem 9 und der Wahl der Momente und der Ordnung auf der BT Struktur im Beweis zu dieser Behauptung wird deutlich, dass für jede Formelmenge, die in einem *BDICTL** Modell erfüllbar ist, die entsprechende Formelmenge in einem *bdi-stit* Modell erfüllbar ist. Somit handelt es sich bei *bdi-stit* tatsächlich um eine Generalisierung im Sinne von Abschnitt 5.3.

Es gibt Formeln, die in *BDICTL** gültig und in *bdi-stit* nicht gültig sind, z.B. $\neg X\varphi \supset X\neg\varphi$. Dies resultiert daraus, dass die Nachfolgerfunktion *succ* in *bdi-stit* Logik partiell sein kann, wohingegen in *CTL** angenommen wird, dass es in jedem Zustand eines beliebigen Pfades immer einen nächsten Zustand gibt. Dass *CTL** entscheidbar ist, wurde in [53] nachgewiesen. Dass die *BDICTL** entscheidbar bleibt, haben Rao und Georgeff in [66] gezeigt. Nimmt man an, dass die Nachfolgerfunktion *succ* tatsächlich eine totale Funktion ist, dann resultiert aus [115], dass das Segment der Sprache *bdi-stit*, welches nur aussagenlogische Junktoren und temporale Operatoren enthält, entscheidbar ist.

Ein Grund, temporale Erweiterungen der *bdi-stit* Logik zu betrachten, ist, dass sich durch die Kombination von Stit-Operatoren und temporalen Operatoren weitere Möglichkeiten aufzeigen, Handlungen von einem und mehreren Akteuren zu beschreiben. Broersen in [28, 29] beschäftigt sich mit Stit-Operatoren, die im nächsten Moment ausgewertet werden. Für einen Akteur bzw. eine Gruppe von Akteuren

führt er einen $xstit$ -Operator ein. Broersen benennt als Vorteil dieser Operatoren, dass es möglich ist auszudrücken, dass α dafür sorgt, dass β eine bestimmte Handlung ausführt. Da die Bewertung von α 's Handlung und die Bewertung vom Ausführen von β 's Handlung in verschiedenen Momenten erfolgt, ist eine Formel $\alpha xstit : \beta xstit : \varphi$ erfüllbar, ohne dass aus der Unabhängigkeit der Akteure folgen würde, dass es eine Notwendigkeit wie im Falle des $cstit$ -Operators bzw. eine Unmöglichkeit wie im Falle des $dstit$ -Operators ist. Bei Broersen ist der $xstit$ -Operator ohne Negativ-Bedingung aufgeführt.

Zuvor in [30, 31, 32] führten Broersen, Herzig und Troquard bereits unabhängig von den Stit-Operatoren temporale Operatoren X und U in die Stit-Logik ein. Dies diente, um zu zeigen, dass die jeweils erhaltenen, logischen Systeme eine Einbettung von ATL [6] und Coalition Logic [105] in die Stit Theorie ermöglichen. Somit erweitert eine Hinzunahme von temporalen Operatoren wie X und U die Anwendungsbereiche der Stit-Theorie. Da eine $bdi-stit$ Logik die Stit-Theorie ihrerseits erweitert, ist eine Einbettung dieser Logiken in $bdi-stit$ Logik ebenfalls möglich.

6.4 $bdi-stit_{dint}$ — Interaktionen zwischen Stit- und Intensionsoperatoren

Dieser Abschnitt beschäftigt sich mit einer Erweiterung der Axiomatisierung von $bdi-stit$ Logik um die bereits in Abschnitt 4.3 motivierten Interaktionen zwischen dem Intensionsoperator und den Handlungsoperatoren. Im Gegensatz zu dem bereits in $bdi-stit$ eingeführten int -Operator hatte ich dafür argumentiert, dass Intentionen sich auf Handlungen richten, so dass der intendierte Sachverhalt durch eine Handlung des Akteurs hervorgebracht wird. Daher liegen zwei Gründe dafür vor, dass es sich bei dem intendierten Sachverhalt nicht um einen notwendigerweise bestehenden Sachverhalt handeln kann. Eine Handlung eines Akteurs wird mittels des $dstit$ -Operators beschrieben. Dieser weist eine Negativ-Bedingung auf, so dass der durch die Handlung hervorgebrachte Sachverhalt, der nach der Handlung vorliegt, nicht bereits historisch notwendigerweise vorliegen kann. Wenn eine Intention sich auf eine Handlung richtet und somit auf einen Sachverhalt, der durch eine Handlung hervorgebracht wird, sollte dieser Sachverhalt nicht bereits historisch notwendigerweise bestehen.

Der zweite Grund besteht darin, dass Intentionen eine Abwägung durch den Akteur, was er tun wird, inhärent ist. Wenn man sich Bratmans Motivation des Begriffes der Intention und die Notwendigkeit der Intention hinsichtlich der Durchführung bzw.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Erstellung von Plänen in Erinnerung ruft, ist es nur natürlich anzunehmen, dass Intentionen sich auf Sachverhalte richten, die nicht bereits historisch notwendigerweise bestehen, sondern eingeplant, berücksichtigt, hervorgebracht etc. werden müssen. Dies sollte sich in der Wahrheitsbedingung eines Intentionsoperators wiederfinden lassen. Daher werde ich zu dem in Abschnitt 5.2 eingeführten *int*-Operator einen weiteren, sogenannten deliberativen Intentionsoperator einführen $\alpha \text{dint}:\varphi = \alpha \text{int}:\varphi \wedge \neg \Box\varphi$.

Die Wahrheitsbedingung des *dint*-Operators beinhaltet die Wahrheitsbedingung des *int*-Operators und wird durch eine Negativ-Bedingung ergänzt. Durch die Negativ-Bedingung wird sichergestellt, dass der intendierte Sachverhalt nicht historisch notwendigerweise besteht. Diese Definition ist in Analogie zur Unterscheidung zwischen *cstit*- und *dstit*-Operator für Handlungen gewählt worden, vgl. Definition 3 S.18.

Definition 23 Sei \mathcal{M} ein *bdi-stit* Modell, so wird der *dint*-Operator interpretiert durch

$$\begin{aligned} \mathcal{M}, (m, h) \models \alpha \text{dint}:\varphi \text{ gdw. } (+) I_{(m,h)}^\alpha &\subseteq \{s' \mid \mathcal{M}, s' \models \varphi\}, \\ (-) \text{ es gibt } h'' \in H_m \text{ mit } \mathcal{M}, (m, h'') &\not\models \varphi. \end{aligned}$$

Für eine einfachere Handhabung der Axiomatisierung und des Vollständigkeitsbeweises werde ich analog zu der Darstellung des *dstit*-Operators mittels des *cstit*-Operators den *int*-Operator verwenden. In Tabelle 6.9 sind die Axiome für den *int*-Operator festgehalten. Die gültigen Theoreme für den *dint*-Operator sind im Folgenden aufgelistet. Es gilt $\alpha \text{int}:\varphi = \alpha \text{dint}:\varphi \vee \Box\varphi$.

Der Gehalt einer Intention sollte bereits, eine Handlung auszuführen, beinhalten. Wenn φ den Sachverhalt symbolisiert, den der Akteur beabsichtigt, stellt die Formel $\alpha \text{dint}:\varphi$ dar, dass der Akteur beabsichtigt, für den Sachverhalt, dass φ zu sorgen. Die Formel $\alpha \text{dint}:\varphi$ hat daher als logische Konsequenz $\alpha \text{dint}:\alpha \text{dstit}:\varphi$. Dies spiegelt das Interaktionsaxiom (*IA*) wieder.

Jedoch kann ein Akteur nur beabsichtigen etwas hervorzubringen, was im Bereich seiner Fähigkeiten liegt. Wenn der Akteur beabsichtigt, den Sachverhalt φ zu bewirken, dann ist dies nur rational möglich, wenn der Sachverhalt von ihm bewirkt werden kann. Somit ist eine logische Konsequenz aus $\alpha \text{dint}:\varphi$, dass $\Diamond\alpha \text{dstit}:\varphi$. Für die Motivation des Interaktionsaxioms (*PA*), siehe Abschnitt 4.3 S.200.

Das Agglomerationsprinzip (A_i) für den *int*-Operator garantiert, dass die Gehalte aller Intentionen eines Akteurs zu gegebenem Zeitpunkt konsistent sind. Dieses Prinzip kann für den *dint*-Operator als Theorem $(\alpha \text{dint}:\varphi \wedge \alpha \text{dint}:\psi) \supset \alpha \text{dint}:(\varphi \wedge \psi)$ ebenfalls festgehalten werden. Somit gilt ebenso für den *dint*-Operator, dass die Gehalte konsistent sind. Aufgrund von (*PA*) sind in einer der möglichen Wahlzellen sind

6.4. $bdi-stit_{dint}$ — Interaktionen zwischen *Stit*- und *Intentionsoperatoren*

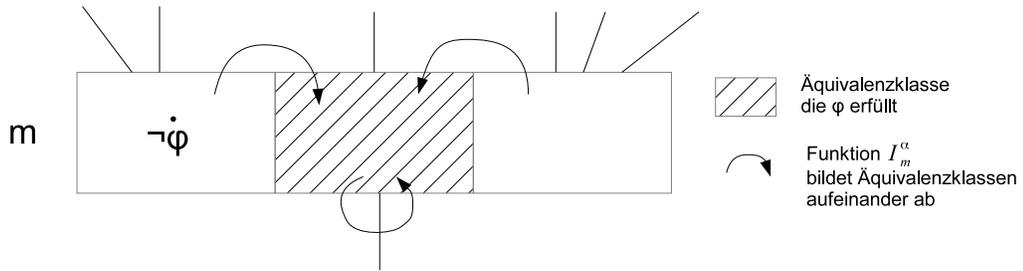


Abbildung 6.1: $\Box\alpha dint:\varphi \wedge \neg\Box\varphi$

alle Gehalte von Intentionen erfüllt, die ein Akteur in einer Situation hat. Für die Motivation von (A_i) und (K_i) , siehe Abschnitt 4.2 S.174.

Um die eben genannten Forderungen an die Interaktionen des *dint*-Operators mit den Handlungsoperatoren zu gewährleisten, muss die Menge an Situationen, die kompatibel mit dem sind, was der Akteur in der aktuellen Situation intendiert, weiter eingrenzt, als dies im allgemeinen *bdi-stit* Modell der Fall ist. Dafür wird die Funktion I eingeführt. Die Funktion $I : \mathcal{A} \times S \rightarrow H$ bildet jedes Akteur-Situations-Paar $(\alpha, (m, h))$ auf eine Geschichte $h' \in H_m$ ab. Die Menge $I_{(m,h)}^\alpha$ an Situationen, die kompatibel mit dem sind, was α in der Situation (m, h) intendiert, sind die zu der Situation $(m, I(\alpha, (m, h)))$ wahläquivalenten Situationen $I_{(m,h)}^\alpha = C_{(m,I(\alpha,(m,h)))}^\alpha$.

Im Rahmen der *bdi-stit_{dint}* Logik wird die Relation I durch die Funktion I ersetzt, wobei sich die Relation I über die Funktion I definieren lässt. Deswegen bezeichne ich beide mit I . An der Indizierung bzw. den Argumenten von I ist eindeutig erkennbar, ob die Relation oder die Funktion betrachtet wird. Zudem ist es sinnvoll zu fordern, dass zwischen wahläquivalenten Situationen keine verschiedenen Absichten vorliegen. Falls zwei Situationen s, s' im Moment m für einen Akteur α wahläquivalent sind, gilt $(m, I(\alpha, s)) = C_{(m,I(\alpha,s'))}^\alpha$ bzw. $I_s^\alpha = I_{s'}^\alpha$. Ich hatte in Abschnitt 4.3 dafür argumentiert, dass aufgrund der Kontrolle, die der Akteur über das Fassen von Intentionen hat, das Fassen einer bestimmten Intention unter einen mentalen Akt fällt. Der Akteur entscheidet sich für seine Absichten, so dass die Menge an Situationen, die kompatibel mit dem sind, was der Akteur beabsichtigt, abgeschlossen unter wahläquivalenten Situationen ist. Dies spiegelt das Interaktionsaxiom (AI) wider. Einem Zustand, dass der Akteur beabsichtigt, dass φ , geht ein mentaler Akt voraus, dass diese Intention von dem Akteur gefasst wurde $\alpha dint:\varphi \supset \alpha cstit:\alpha dint:\varphi$. Dies folgt direkt aus dem Axiom (AI) , siehe Tabelle 6.9. In Abschnitt 4.3 ist festgehalten, dass der mentale Akt des Fassens der Intention nur mit dem *cstit*-Operator beschrieben werden kann. Für eine Intention ist es möglich, dass sie historisch notwendigerweise vorliegt, vgl. Theorem

6. Vollständigkeit und Erweiterungen von *bdi-stit*

($\diamond NI$) in Abb. 6.1. Dennoch basiert die Intention auf einem mentalen Akt des Akteurs, vgl. für die Motivation Abschnitt 4.3 S.193.

Durch diese Festlegung der Situationen, die zur Bewertung des *int*-Operators und somit des *dint*-Operators herangezogen werden, ist sichergestellt, dass die Gehalte der Intentionen in allen Situationen einer möglichen Wahlzelle des Akteurs wahr sind. Damit sind die Gehalte historisch möglich und konsistent, d.h. $\alpha \text{dint}:\varphi \supset \diamond\varphi$, $\alpha \text{dint}:\varphi \supset \diamond\alpha \text{dstit}:\varphi$ und $(\alpha \text{dint}:\varphi \wedge \alpha \text{dint}:\psi) \supset \diamond(\varphi \wedge \psi)$ sind allgemeingültig über den mit der Funktion *I* spezifizierten *bdi-stit_{dint}* Rahmen.

Der *dint*-Operator erfüllt so die Axiome (K_i) und (D_i). Aufgrund der Negativ-Bedingung ist die Regel der Monotonie (RM_i) und die Regel der Necessitation (RN_i) für den *dint*-Operator im Gegensatz zum *int*-Operator nicht gültig. Dass dies rational ist zu fordern, wurde in Abschnitt 4.2 S.180 gezeigt. Dies unterstreicht die Analogie der *int*- und *dint*-Operatoren zu den *cstit*- und *dstit*-Operatoren. Im Gegensatz zum *cstit*-Operator, der die Axiomenschemata *T* und 4 erfüllt, sind diese für beide Intentionsoperatoren nicht allgemeingültig. Nichtsdestotrotz gilt das Theorem (A_i) (Agglomerationsprinzip) für beide Intentionsoperatoren, vgl. Abschnitt 6.1 S.245. Ein zentraler Unterschied zwischen dem *dint*-Operator und dem *dstit*-Operator ist die Interaktion mit dem Notwendigkeitsoperator. In Abbildung 6.1 wird deutlich, dass es möglich ist, dass der Akteur die Intention deliberativ und notwendigerweise bildet, aber der Gehalt der Intention nicht notwendigerweise wahr ist. Dies ist für den *dstit*-Operator nicht möglich. Der hervorgebrachte Sachverhalt kann nicht notwendigerweise bestehen, wenn er das Resultat einer Handlung des Akteurs ist. Demzufolge kann das Hervorbringen eines Sachverhaltes, der in der aktuellen Situation der Handlung immer besteht, nicht notwendigerweise bestehen. Sonst müsste der Sachverhalt in jeder möglichen Situation bestehen und damit selbst notwendigerweise der Fall sein. Bei einer Intention ist dies anders. Da das Vorliegen einer Intention den intendierten Sachverhalt nicht als bestehend impliziert, sollte es möglich sein, dass eine Intention notwendigerweise vorliegen kann, ohne dass dies impliziert, dass der Sachverhalt notwendigerweise vorliegt.

Da jede Intention eine Proeinstellung ist, beinhaltet jede intendierte Handlung bzw. jeder intendierte Sachverhalt eines Akteurs, dass dieser den Wunsch zu dieser Handlung bzw. diesem Sachverhalt in sich trägt. Da Wünsche eines Akteurs im Gegensatz zu seinen Intentionen durchaus konfliktieren können, impliziert die Annahme, dass α wünscht, dass φ , und beabsichtigt, dass $\neg\varphi$, keinen Widerspruch. Obgleich Intentionen Proeinstellungen sind, kann es möglicherweise Wünsche des Akteurs geben, die mit diesen Intentionen in Konflikt stehen. Die Gehalte der Intentionen eines rationa-

6.4. $bdi-stit_{dint}$ — Interaktionen zwischen $Stit$ - und $Intentions$ operatoren

(IA)	$\alpha int:\varphi \supset \alpha int:\alpha cstit:\varphi$
(NIA)	$\neg\alpha int:\varphi \supset \alpha int:\neg\alpha cstit:\varphi$
(AI)	$\alpha int:\varphi \supset \alpha cstit:\alpha int:\varphi$
(NAI)	$\neg\alpha int:\varphi \supset \alpha cstit:\neg\alpha int:\varphi$
(PA)	$\alpha int:\varphi \supset \diamond\alpha cstit:\varphi$
(NI)	$\Box\varphi \supset \alpha int:\varphi$
(DI)	$\alpha int:\varphi \supset \alpha des:\varphi$

Tabelle 6.9: Axiomatisierung des int -Operators in $bdi-stit_{dint}$ Logik

len Akteurs sind dennoch nicht konfligierend. Der $bdi-stit_{dint}$ Rahmen wird um eine Bedingung ergänzt, die sicher stellt, dass es sich bei der Intention, dass p , um einen Wunsch handelt, vgl. Schema (DI). Die Erläuterungen für dieses Schema finden sich in Abschnitt 4.2 S.170. Für alle Situationen s und Akteure α gilt, dass $I_s^\alpha \in D_s^\alpha$.

Da ich den Nachweis der Vollständigkeit der so definierten $bdi-stit_{dint}$ Logik mit $dint$ -Operator der Einfachheit halber mittels der $cstit$ - und int -Operatoren führe, werde ich die Axiome und Regeln ebenfalls mit den $cstit$ - und int -Operatoren formulieren. Für eine vollständige Axiomatisierung der eben eingeführten $bdi-stit_{dint}$ Logik mit deliberativem Intentionoperator stellen die in Abschnitt 6.1 genannten Axiome und Regeln zusammen mit den neu hinzukommenden Axiomen in Tabelle 6.9 eine vollständige Axiomatisierung dar. Folglich ergeben sich die eben genannten Theoreme aus $\alpha dint:\varphi = \alpha int:\varphi \wedge \neg\Box\varphi$. Sie lassen sich für den $dint$ -Operator aus der Axiomatisierung der $bdi-stit_{dint}$ Logik herleiten. Die wichtigsten sind in Tabelle 6.10 zusammengefasst.

Zu beabsichtigen, dass φ , für einen Akteur impliziert, dass der Akteur intendiert, dass er für das sorgt, was er beabsichtigt. Kurz gesagt, eine Absicht richtet sich auf eine gegenwärtige oder geplante Handlung des Akteurs.

$$(TIA) \quad \alpha dint:\varphi \supset \alpha dint:\alpha dstit:\varphi^{25}$$

Das Schema (NIA), vgl. Tabelle 6.9, ist jedoch nicht auf den $dint$ -Operator übertragbar. Dies ist auch nicht gewollt. Wenn ein Akteur keine Absicht zu p geformt

²⁵Den $dstit$ -Operator anstelle des $cstit$ -Operators zu schreiben, ist korrekt, da $\alpha dint:\varphi \supset \neg\Box\varphi$ gilt.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

(IA)	$\alpha \text{ dint} : \varphi \supset \alpha \text{ dint} : \alpha \text{ dstit} : \varphi$
(NIA)	$\not\vdash \neg \alpha \text{ dint} : \varphi \supset \alpha \text{ dint} : \neg \alpha \text{ dstit} : \varphi$
(AI)	$\alpha \text{ dint} : \varphi \supset \alpha \text{ cstit} : \alpha \text{ dint} : \varphi$
(NAI)	$\not\vdash \neg \alpha \text{ dint} : \varphi \supset \alpha \text{ dstit} : \neg \alpha \text{ dint} : \varphi$
(PA)	$\alpha \text{ dint} : \varphi \supset \Diamond \alpha \text{ dstit} : \varphi$
(DI)	$\alpha \text{ dint} : \varphi \supset \alpha \text{ des} : \varphi$
(MEC)	$(\Box(\varphi \supset \psi) \wedge \Box(\psi \supset \alpha \text{ dint} : \psi)) \supset (\alpha \text{ dint} : \varphi \supset \alpha \text{ dint} : \psi)$
(\Diamond NI)	$\not\vdash \Box \alpha \text{ dint} : \varphi \supset \perp$
(TII)	$\alpha \text{ dint} : \beta \text{ dint} : \varphi \supset \perp \quad \alpha \text{ dstit} : \beta \text{ dint} : \varphi \supset \perp \quad \alpha \text{ dint} : \beta \text{ dstit} : \varphi \supset \perp$
(TII _k)	$(\Diamond \alpha_1 \text{ dint} : \varphi_1 \wedge \dots \wedge \Diamond \alpha_k \text{ dint} : \varphi_k) \supset \Diamond (\alpha_1 \text{ dint} : \varphi_1 \wedge \dots \wedge \alpha_k \text{ dint} : \varphi_k)$

Tabelle 6.10: Theoreme für den Intentionsoperator in *bdi-stit_{dint}* - Darstellung der Interaktionsschemata zwischen Handlungen und Absichten

hat, ist es nicht notwendigerweise der Fall, dass er beabsichtigt, dass p , nicht bewirkt zu haben. Keinerlei Absicht hinsichtlich p zu haben, ist jederzeit für einen beliebigen Akteur möglich. Aufgrund der Negativ-Bedingung ist das Schema $\neg \alpha \text{ dint} : \varphi \supset \alpha \text{ dint} : \neg \alpha \text{ cstit} : \varphi$ für den *dint*-Operator nicht allgemeingültig, wohingegen es für den *int*-Operator gültig ist. Dies ist ein weiterer Grund, warum eine Negativ-Bedingung für einen Intentionsoperator angemessen ist. Wenn jemand nicht beabsichtigt, dass p , besteht keinerlei Notwendigkeit, dass er beabsichtigt, es nicht zustande zu bringen.

Desgleichen wie für das Schema (NIA) gilt für das Schema (NAI). Etwas nicht zu beabsichtigen, bedeutet nicht, dass man in irgendeiner Weise es bewirkt hat, dass man es nicht beabsichtigt. Wenn der Akteur keinerlei Einstellung zu p hat, wäre die Folgerung unsinnig, dass er in irgendeiner Weise auf die Nicht-Beabsichtigung von p eingewirkt hätte. Die zum Schema (NAI) analoge Formulierung für den *dint*-Operator ist daher nicht allgemeingültig. Dies macht deutlich, dass der *dint*-Operator einen mentalen Zustand wie den der Intention geeigneter als der *int*-Operator beschreiben kann, und begründet somit, warum die Erweiterung um den *dint*-Operator und die Spezifizierung der Relation I im Rahmen der *bdi-stit_{dint}* Logik sinnvoll ist. Das Schema (NI) ist ebenso wie im Falle des *dstit*-Operators für den *dint*-Operator nicht gültig.

6.4. $bdi-stit_{dint}$ — Interaktionen zwischen $Stit$ - und $Intentions$ operatoren

Da ein Akteur Kontrolle über seine Absichten hat, muss der Gehalt einer Absicht im Rahmen seiner Wahlmöglichkeiten liegen, so dass sich folgendes Theorem aus dem Axiom (AI) ableiten lässt.

$$(TAI) \quad \alpha \text{ dint}:\varphi \supset \alpha \text{ cstit}:\alpha \text{ dint}:\varphi$$

Das Axiom (PA) garantiert, dass ein idealisierter, rationaler Akteur nur das intendiert, was er hervorbringen kann. Der beabsichtigte Sachverhalt p muss im Rahmen seiner Möglichkeiten bzw. Fähigkeiten liegen.

$$(TPA) \quad \alpha \text{ dint}:\varphi \supset \diamond \alpha \text{ dstit}:\varphi$$

Die Proeinstellung wird durch den $dint$ -Operator ausgedrückt.

$$(TDI) \quad \alpha \text{ dint}:\varphi \supset \alpha \text{ des}:\varphi$$

Ein wichtiger Punkt bei der Motivation der Axiome für den Intentionsoperator war die Kohärenz der beabsichtigten Ziele und den dafür notwendigen Mitteln. In Abschnitt 4.3 hatte ich mich dagegen ausgesprochen, dass Intentionen eine bestimmte Überzeugung zugrunde liegt. Dennoch ist es rational zu fordern, dass zu einem beabsichtigten Ziel die dafür notwendigen Mittel beabsichtigt werden. Ein notwendiges Mittel sollte beabsichtigt werden, wenn es der Fall wäre, dass, ohne diese Absicht, das Mittel nicht eintreten würde. Bratman hatte deshalb gefordert, dass, wenn der Akteur eine Mittel-Zweck-Überzeugung hat, er demgemäß eine Kohärenz zwischen der Intention für das notwendig geglaubte Mittel und dem beabsichtigten Zweck sicher stellt. Ich hatte in Abschnitt 4.3 S.193f argumentiert, dass die Überzeugungen eines Akteurs bei einer solchen rationalen Forderung keine Rolle spielen. Um sicherzustellen, dass der Akteur tatsächlich erkennen kann, dass es notwendig ist, das Mittel zu beabsichtigen, spielt es keine Rolle, was er glaubt, sondern welche Mittel tatsächlich notwendig sind zu beabsichtigen. Daher wurde Theorem (MEC) für Intentionen als sinnvoll motiviert und ein Schema ($BMEC$) abgelehnt, siehe Abschnitt 4.3 S.215.

Dass das Theorem (MEC) über einem $bdi-stit_{dint}$ Rahmen gültig ist, lässt sich leicht zeigen. Angenommen es gilt $\alpha \text{ dint}:\varphi$. Falls $\Box(\varphi \supset \psi)$ in der Situation erfüllt ist, gelten $\alpha \text{ int}:\psi$ und somit die Positiv-Bedingung von $\alpha \text{ dint}:\psi$. Damit ist ψ möglich in der Situation. Aus dem zweiten Teil des Theorems, dass der Akteur für das Bestehen von ψ beabsichtigen muss, dass ψ , folgt, dass es eine mögliche Situation gibt, in der ψ besteht. Daher ist $\diamond \alpha \text{ dint}:\psi$ in der aktuellen Situation erfüllt. Die Negativ-Bedingung $\neg \Box \psi$ ist erfüllt, da $\neg \Box \varphi$ und $\Box(\varphi \supset \psi)$ in der aktuellen Situation beide erfüllt sind. Damit steht fest, dass das Theorem (MEC) allgemeingültig ist.

Die Betrachtung von Interaktionen zwischen den Intentionsoperatoren verschiedener Akteure ist interessant, aber nicht Teil dieser Arbeit. Es bleibt nur die Unabhängig-

6. Vollständigkeit und Erweiterungen von *bdi-stit*

keit der Akteure festzuhalten. Da verschiedene Akteure bezüglich ihrer Handlungen als unabhängig angenommen werden, sollten sie erst recht bezüglich ihrer propositionalen Einstellungen als unabhängig betrachtet werden. Das folgende Theorem ergibt sich aus dem Axiom der Unabhängigkeit von k Akteuren (AIA_k) und der Interaktion des *int*-Operators mit dem *cstit*-Operator, vgl. Axiom (AI) in Tabelle 6.9.

$$(TII_k) \quad (\Diamond \alpha_1 \text{dint} : \varphi_1 \wedge \dots \wedge \Diamond \alpha_k \text{dint} : \varphi_k) \supset \Diamond (\alpha_1 \text{dint} : \varphi_1 \wedge \dots \wedge \alpha_k \text{dint} : \varphi_k)$$

Möglicherweise ist die Behauptung, dass es sich hierbei um ein Theorem handelt, nicht offensichtlich. Daher skizziere ich die Beweisidee. Die Regel der Monotonie gilt für den Möglichkeitsoperator $\varphi \supset \psi / \Diamond \varphi \supset \Diamond \psi$. Da zudem aus dem Axiom (AI) das Theorem (TAI) folgt, gilt $\Diamond \alpha_i \text{dint} : \varphi_i \supset \Diamond \alpha_i \text{cstit} : \alpha_i \text{dint} : \varphi_i$ für $i \leq k$. Somit erhält man die Konjunktion der Formeln $\Diamond \alpha_i \text{cstit} : \alpha_i \text{dint} : \varphi_i$ für alle $i \leq k$. Auf diese Konjunktion wird das Axiom (AIA_k) angewandt. Man erhält $\Diamond (\bigwedge_{i \leq k} \alpha_i \text{cstit} : \alpha_i \text{dint} : \varphi_i)$. Aus der Faktivität des *cstit*-Operators und der Regel der Monotonie folgt das Theorem (TII_k).

Die Axiome (TII) folgen ebenfalls aus der Unabhängigkeit der Akteure, d.h. aus der Unabhängigkeitsbedingung der BT+AC Struktur Abschnitt 1.2 S.16. Für je zwei Wahläquivalenzklassen der Akteure α und β gilt, dass ihre Schnittmenge nicht leer ist. Sobald ein Akteur α dafür sorgen würde, dass in jeder Situation einer seiner Äquivalenzklassen $\beta \text{dint} : \varphi$ gelte, gibt es somit in jeder Äquivalenzklasse von β eine Situation, in der $\beta \text{dint} : \varphi$ wahr ist. Somit gilt in jeder Äquivalenzklasse von β , die ein Bild unter I^β ist, dass φ in jeder Situation der Bildklasse wahr ist. Daher ist in jeder möglichen Situation $\Box \beta \text{dint} : \varphi$ wahr. Die Negativ-Bedingung ist weder für $\alpha \text{dint} : \beta \text{dint} : \varphi$ noch für $\alpha \text{cstit} : \beta \text{dint} : \varphi$ erfüllbar. Folgende Formeln sind daher ebenfalls Theoreme $\alpha \text{cstit} : \beta \text{dint} : \varphi \supset \Box \beta \text{dint} : \varphi$ und $\alpha \text{int} : \beta \text{dint} : \varphi \supset \Box \beta \text{dint} : \varphi$. Die mit (TII) benannten Theoreme in Tabelle 6.9 sind allgemeingültig in *bdi-stit_{dint}*.

Die Gültigkeit der in Tabelle 6.9 benannten Axiome ist nun offensichtlich, so dass die folgende Behauptung festgehalten werden kann.

Proposition 24 *Die Axiome und Regeln der bdi-stit Logik, vgl. Abschnitt 6.1, sowie die Axiome in Tabelle 6.9 sind gültig über dem spezifizierten bdi-stit_{dint} Rahmen mit erweitertem Intensionsoperator.*

Nachdem festgehalten wurde, dass die Regeln und Axiome korrekt sind, möchte ich mich nun dem Nachweis der Vollständigkeit zuwenden. Dafür wird wie in Abschnitt 6.1 ein kanonisches Modell konstruiert. Als Basis für dieses Modell dient das bereits definierte, kanonische *bdi-stit* Modell, das geeignet verändert wird. Die Menge aller

Formeln über der Sprache der $bdi\text{-}stit_{dint}$ Logik sei mit \mathcal{L}_{dint} bezeichnet.²⁶

Sei W die Menge aller maximal konsistenten Formelmengen über \mathcal{L}_{dint} . Für jede Teilmenge X von W ist eine Relation $wR_X w'$ gdw. $\{\varphi \mid \Box\varphi \in w\} \subseteq w'$ auf der Teilmenge X definiert. Dies ist eine Äquivalenzrelation, vgl. Lemma 10.3. Sei $\{X_i\}_{i \in I}$ die Menge aller Äquivalenzklassen zu R_W und $X \in \{X_i\}_{i \in I}$. Dann ist für alle $\alpha \in \mathcal{A}$ mit $wR_\alpha w'$ gdw. $\{\varphi \mid \alpha cstit:\varphi \in w\} \subseteq w'$ ebenfalls eine Äquivalenzrelation auf X gegeben, vgl. Lemma 10. Die Relation R_α hängt von der Wahl von X ab. Die Menge X sei für die folgenden Ausführungen festgewählt. Mit E_α wird die Menge aller Äquivalenzklassen der Relation R_α auf X bezeichnet.

Das folgende Lemma lässt sich anhand der Axiomatisierung beweisen bzw. die einzelnen Punkte sind bereits in Lemma 11 festgehalten worden.

Lemma 25 Für alle $\varphi \in \mathcal{L}_{dint}$, $\alpha \in \mathcal{A} = \{\alpha_j\}_{j \in J}$

- (i) $\Box\varphi \in w$ gdw. $\varphi \in w'$ für alle $w' \in X$
gdw. $\Box\varphi \in w'$ für alle $w' \in X$.
- (ii) $\alpha cstit:\varphi \in w$ gdw. $\varphi \in w'$ für alle $w' \in X$ mit $wR_\alpha w'$
gdw. $\alpha cstit:\varphi \in w'$ für alle $w' \in X$ mit $wR_\alpha w'$.
- (iii) $\alpha dstit:\varphi \in w$ gdw. $\varphi \in w'$ für alle $w' \in X$ mit $wR_\alpha w'$
und es existiert $w'' \in X$ mit $\neg\varphi \in w''$.
- (iv) Für jede beliebige Funktion f von $\{\alpha_j\}_{j \in J}$ in die Menge aller E_{α_j} mit $j \in J$, wobei $f(\alpha_j) \in E_{\alpha_j}$, gilt

$$\bigcap_{j \in J} f(\alpha_j) \neq \emptyset.$$

BEWEIS: Vgl. [19] bzw. Beweis von Lemma 11. □

Im folgenden Lemma weicht die Konstruktion des kanonischen $bdi\text{-}stit_{dint}$ Rahmens für die Erweiterung von der Konstruktion des allgemeinen $bdi\text{-}stit$ Modells ab. Es resultiert aus der veränderten Interpretation des Intensionsoperators.

²⁶Der Einfachheit halber verzichte ich auf die Berücksichtigung der Axiome APC_n , da dies nichts an dem Nachweis der Korrektheit und Vollständigkeit geändert hätte. Lediglich der Wertebereich der im kanonischen Modell zu definierenden Funktion I hätte sich auf höchstens n Äquivalenzklassen eingeschränkt.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

Lemma 26 Sei X, R, R_α, E_α wie eben definiert, dann existiert eine Funktion $f : X \rightarrow X$, so dass für alle $\alpha \in \mathcal{A}$, $w, w' \in X$ und $\varphi \in \mathcal{L}_{\text{int}}$ gilt

$$\alpha \text{ int} : \varphi \in w \text{ gdw. für alle } w'' \in X \text{ gilt, falls } f(w)R_\alpha w'', \text{ dann } \varphi \in w''. \quad (*)$$

BEWEIS: Ich gebe die Konstruktion der Funktion f an und zeige, dass sie die gewünschte Bedingung (*) erfüllt. Die Funktion bildet jeden Repräsentanten einer Äquivalenzklasse aus E_α auf einen Repräsentanten der intendierten Äquivalenzklasse ab. Da jede Absicht impliziert, dass es möglich ist, für das Beabsichtigte zu sorgen, Axiom (PA), gehört der Repräsentant der intendierten Äquivalenzklasse zur selben Klasse X bezüglich R . Die Relation R_α ist eine Äquivalenzrelation auf X . Das Axiom (PA) sorgt zudem dafür, dass eine solche Äquivalenzklasse existiert. Die Funktion f muss dabei weder surjektiv sein, noch muss jede Äquivalenzklasse im Bild der Funktion vertreten sein. Somit ist es weiterhin möglich, dass nicht jede mögliche Wahlzelle intendiert wird.

Für jedes $w \in X$ wird mittels des *int*-Operators eine Menge von Formeln wie folgt definiert.

$$I_w^\alpha = \{ \varphi \mid \alpha \text{ int} : \varphi \in w \}$$

Aufgrund von (Di) ist die Menge konsistent und es gibt mindestens eine maximal konsistente Menge w'' mit $I_w^\alpha \subseteq w''$. Die Relation I^α ist auf X durch $wI^\alpha w'$ gdw. $\{ \varphi \mid \alpha \text{ int} : \varphi \in w \} \subseteq w'$ definiert. Die Funktion f wird mit $f(w) = w'$ für ein beliebiges Element w' mit $wI^\alpha w'$ gewählt. Da es mindestens ein w'' mit $I_w^\alpha \subseteq w''$ gibt, ist die Funktion wohldefiniert. Somit gilt

- (i) $I_w^\alpha \cup \{ \alpha \text{ cstit} : \varphi \mid \varphi \in I_w^\alpha \} \subseteq I_w^\alpha$,
- (ii) falls $wR_\alpha w'$, dann $I_w^\alpha = I_{w'}^\alpha$,
- (iii) falls $wR_\alpha w'$, dann $f(w)R_\alpha f(w')$.

Die Behauptung (i) folgt offensichtlich aus Axiom (IA). Für den Nachweis von (ii) werde angenommen, dass $wR_\alpha w'$ und $\alpha \text{ int} : \varphi \in w$ sowie $\alpha \text{ int} : \varphi \notin w'$. Die umgekehrte Richtung gilt ebenfalls, da R_α symmetrisch ist. Mit Axiom (AI) gilt, dass $\alpha \text{ cstit} : \alpha \text{ int} : \varphi \in w$. Nach Lemma 25.(ii) ist die Annahme widersprüchlich. Somit ist (ii) bewiesen.

Angenommen es gelte $wR_\alpha w'$ und es ist nicht der Fall, dass $f(w)R_\alpha f(w')$. Mit Lemma 25.(ii) folgt, dass $\alpha \text{ cstit} : \varphi \in w$ gdw. $\alpha \text{ cstit} : \varphi \in w'$ für alle φ . Die Annahme führt dazu, dass ein ψ mit $\alpha \text{ cstit} : \psi \in f(w)$ und $\neg \alpha \text{ cstit} : \psi \in f(w')$ aufgrund der Maximalität von w, w' existiert. Da jedoch $I_w^\alpha = I_{w'}^\alpha$, ist $\psi \notin I_w^\alpha$ und daher $\neg \alpha \text{ int} : \psi \in w, w'$. Aus

6.4. $bdi\text{-}stit_{dint}$ — Interaktionen zwischen Stit- und Intensionsoperatoren

der Anwendung von Axiom (NIA) resultiert, dass $\alpha int:\neg\alpha cstit:\psi \in w, w'$. Folglich ist $\neg\alpha cstit:\psi \in I_w^\alpha, I_{w'}^\alpha$. Da $I_w^\alpha \subseteq f(w)$, widerspricht dies $\alpha cstit:\psi \in f(w)$. Infolgedessen ist (i) - (iii) für f bewiesen.

Es ist einfach zu sehen, dass eine so konstruierte Funktion f die Bedingung (*) erfüllt. Falls $\alpha int:\varphi \in w$, folgt aus (i), dass $I_w^\alpha \subseteq \{\psi \mid \alpha cstit:\psi \in f(w)\}$. Somit gilt für alle $w'' \in W$, falls $f(w)R_\alpha w''$, dann $\varphi \in w''$. Für die andere Richtung sei angenommen, dass für alle $w'' \in W$ gelte, falls $f(w)R_\alpha w''$, dann $\varphi \in w''$. So lassen sich vier Fälle betrachten.

1. $\alpha int:\varphi, \neg\alpha int:\neg\varphi \in w$
2. $\alpha int:\varphi, \alpha int:\neg\varphi \in w$
3. $\neg\alpha int:\varphi, \alpha int:\neg\varphi \in w$
4. $\neg\alpha int:\varphi, \neg\alpha int:\neg\varphi \in w$

Aus 2. folgt, dass $\varphi, \neg\varphi \in I_w^\alpha$. Dies widerspricht Axiom (D_i). Aus 3. resultiert $\neg\varphi \in I_w^\alpha$. Da $f(w)R_\alpha f(w)$ aufgrund der Reflexivität von R_α und $I_w^\alpha \subseteq f(w)$, erhält man ebenfalls einen Widerspruch. Aus 4. folgt, dass $\alpha int:\neg\alpha cstit:\neg\psi \in w$ und $\alpha int:\neg\alpha cstit:\psi \in w$. Dies steht im Widerspruch zu (iii). Ergo ist der einzige nicht kontradiktorische Fall 1. $\alpha int:\varphi \in w$ und $\neg\alpha int:\neg\varphi \in w$.

□

Das folgende Theorem hält die Vollständigkeit dieser Axiomatisierung fest. Der Beweis gibt an, wie ein kanonisches Modell für jede konsistente Formelmenge konstruiert werden kann.

Theorem 27 (Vollständigkeit) *Jede konsistente Menge Φ über \mathcal{L}_{dint} ist durch ein $bdi\text{-}stit_{dint}$ Modell mit spezifiziertem Rahmen erfüllbar.*

BEWEIS: Sei $A = \{\alpha \mid \text{Akteursvariable } \alpha \text{ kommt vor in } \Phi\}$ ²⁷ die Menge aller Akteursvariablen in Φ . Sei $\{X_i\}_{i \in I}$ die Menge aller Äquivalenzklassen von R auf W . Der Rahmen des $bdi\text{-}stit_{dint}$ Modells $\mathcal{F} = (Tree, \leq, \mathcal{A}, C, B, D, I)$ besteht aus einer BT Struktur $(Tree, \leq)$, der Menge der Akteure \mathcal{A} , den Funktionen C, B, D und I . Diese sind wie folgt definiert.²⁸

²⁷Ich werde wieder die Menge der Akteursvariablen mit der Menge der Akteure identifizieren und Identitäten zwischen Akteuren unberücksichtigt lassen. Dass dies o.E.d.A. erfolgt, hat sich für den $bdi\text{-}stit$ Rahmen schon erwiesen.

²⁸Die Konstruktion der BT Struktur weicht nicht von der des kanonischen $bdi\text{-}stit$ Modells ab, vgl. Abschnitt 6.1 bzw. [127]. Die Funktion N und der \diamond -Operator wurden hierbei nicht berücksichtigt, da keine Interaktion mit dem Intensionsoperator eingeführt wurde und er sich gegenüber dem $bdi\text{-}stit$ Modell nicht anders verhält.

6. Vollständigkeit und Erweiterungen von *bdi-stit*

- $Tree := \{w \mid w \in W\} \cup \{X_i \mid i \in I\} \cup \{W\}$
- $\leq := rtrcl(\{(\tilde{W}, \tilde{X}_i), (\tilde{X}_i, w) \mid w \in \tilde{X}_i, i \in I\})$
- $\mathcal{A} := A$
- Für jeden Akteur $\alpha \in \mathcal{A}$ und jeden Moment $m \in Tree$ seien die Wahläquivalenzen wie folgt festgelegt.
 - $C(\alpha, w) := \{h_w\}$, wobei h_w die einzige Geschichte ist, die durch w verläuft, $h_w = \{w, X_i, W\}$ und X_i ist die Äquivalenzklasse, die w enthält.
 - $C(\alpha, W) := \{h_w \mid w \in W\}$
 - Für jede Äquivalenzklasse X_i wird die Relation R_α^i definiert und die Menge aller Äquivalenzklassen von R_α^i auf X_i mit E_α^i notiert. Dann setze $C(\alpha, X_i) := \{H \mid \exists e : e \in E_\alpha^i \text{ und } H = \{h_w \mid w \in e\}\}$.

Die Menge der Situationen S und die Mengen $C_{(m, h_w)}^\alpha$ ergeben sich auf die übliche Weise mit $C_{(m, h_w)}^\alpha = \{(m, h) \mid h \in C(\alpha, m)(h_w)\}$ für jede Situation $(m, h_w) \in S$ und $m \in \{W, X_i, w \mid w \in X_i\}$. Außerdem gilt, dass $wR_\alpha^i w'$ gdw. $w, w' \in e$ für ein $e \in E_\alpha^i$, und dass $wR_\alpha^i w'$ gdw. $h_{w'} \in C(\alpha, X_i)(h_w)$ gdw. $h_w \in C(\alpha, X_i)(h'_w)$.

- $|\varphi| \in B_{(m, h_w)}^\alpha$ gdw. $\alpha \text{ bel} : \varphi \in w$ ²⁹
- $|\varphi| \in D_{(m, h_w)}^\alpha$ gdw. $\alpha \text{ des} : \varphi \vee \alpha \text{ int} : \varphi \in w$ ²⁹
- Die Funktion $I : \mathcal{A} \times S \rightarrow H$ sei wie folgt definiert. Nach Lemma 26 gibt es für jedes $i \in I$ und jeden Akteur $\alpha \in \mathcal{A}$ eine Funktion $f_i^\alpha : X_i \rightarrow X_i$, die die Bedingung (*) erfüllt. Für jeden Akteur $\alpha \in \mathcal{A}$ und jede Situation $s = (m, h_w) \in S$ sei

$$\begin{aligned} I(\alpha, (w, h_w)) &:= h_w, & I(\alpha, (W, h_w)) &:= h_w, \\ I(\alpha, (X_i, h_w)) &:= h_{f_i^\alpha(w)}, & \text{wobei } w &\in X_i. \end{aligned}$$

In Analogie zum *dstit* Rahmen, vgl. bspw. [19], und oben definierten allgemeinen *bdi-stit* Rahmen, vgl. Abschnitt 5.2, erfüllt dieser *bdi-stit*_{dint} Rahmen die Unabhängigkeitsbedingung nach Lemma 25.(iv). Die Argumentation ist analog zu dem Nachweis in Abschnitt 6.1 für den allgemeinen *bdi-stit* Rahmen.

Somit lässt sich ein Modell definieren. Sei v die Bewertungsfunktion, welche jede Akteursvariable in die Menge der Akteure, in diesem Fall auf sich selbst abbildet,

²⁹Die Menge $|\varphi|$ ist durch $|\varphi| := \bigcup_{i \in I} \{(X_i, h_{w'}) \mid \varphi \in w', X_i \in h_{w'}\}$ definiert.

6.4. $bdi-stit_{dint}$ — Interaktionen zwischen Stit- und Intentionsoperatoren

und jeder atomaren Formel p die Menge an Situationen S_p zuordnet, für die für alle $m \in Tree$ und $w \in W$ gilt, falls $s = (m, h_w) \in S_p$, dann $p \in w$. Offensichtlich gelten $v(\top) = S$ und $v(\perp) = \emptyset$.

Mittels Induktion lässt sich zeigen, dass $\mathcal{M}, (X_i, h_w) \models \varphi$ gdw. $\varphi \in w$, für jede Formel $\varphi \in \mathcal{L}_{dint}$ in jeder Situation (X_i, h_w) . Die Induktionsschritte sind für atomare Formeln, Negation, Konjunktion, ebenso für die Stit-Operatoren und den Überzeugungsoperator identisch zum Beweis der Vollständigkeit in der allgemeinen $bdi-stit$ Logik in Abschnitt 6.1. Daher werde ich die Induktionsschritte nur für den Intentions- und Wunschoperator angeben.

$$\begin{aligned} (X_i, h_w) \models \alpha des : \varphi & \text{ gdw. } \text{es gibt } U \in D_{(X_i, h_w)}^\alpha \text{ und } \emptyset \neq U \subseteq \|\varphi\| \\ & \text{(**)} \\ & \text{gdw. } \alpha des : \varphi \in w. \end{aligned}$$

Zu zeigen ist (**). Falls $\alpha des : \varphi \in w$, dann $|\varphi| \in D_{(X_i, h_w)}^\alpha$ mit $w \in X_i$. Nach Induktion gilt $|\varphi| \subseteq \|\varphi\|$. Angenommen es gibt ψ mit $U = |\psi| \neq \emptyset$ und $|\psi| \subseteq \|\varphi\|$, wobei $\alpha des : \psi \vee \alpha int : \psi \in w$. Da $|\psi| = \bigcup_{i \in I} \{(X_i, h_{w'}) \mid \psi \in w', X_i \in h_{w'}\}$, gilt für alle $w' \in W$ und $i \in I$, falls $\mathcal{M}, (X_i, h_{w'}) \models \psi$, dann $\mathcal{M}, (X_i, h_{w'}) \models \varphi$. Nach Induktionsbehauptung gilt für alle $w' \in W$: falls $\psi \in w'$, dann $\varphi \in w'$. Folglich ist $(\psi \supset \varphi) \in w'$ für alle $w' \in W$. Gemäß Regel (RM_d) ist $(\alpha des : \psi \supset \alpha des : \varphi) \in w'$ für alle $w' \in W$. Da $\alpha des : \psi \vee \alpha int : \psi \in w$, gilt entweder $\alpha des : \psi \in w$ und $\alpha des : \varphi \in w$ oder $\alpha int : \psi \in w$. Im zweiten Fall folgt aus Axiom (DI) , dass $\alpha des : \psi \in w$ und somit $\alpha des : \varphi \in w$.

$$\begin{aligned} (X_i, h_w) \models \alpha int : \varphi & \text{ gdw. } C_{(X_i, I(\alpha, (X_i, h_w)))}^\alpha \subseteq \|\varphi\| \\ & \text{gdw. } (X_i, h_{w'}) \models \varphi \text{ für alle } (X_i, h_{w'}) \in C_{(X_i, h_{f_i^\alpha(w)})}^\alpha \\ & \text{Def.C} \\ & \text{gdw. } \text{für alle } w' \in X_i \text{ gilt, falls } w' R_{\alpha J_i}^i f_i^\alpha(w), \\ & \text{dann } (X_i, h_{w'}) \models \varphi \\ & \text{Ind.} \\ & \text{gdw. } \text{für alle } w' \in X_i \text{ gilt, falls } f_i^\alpha(w) R_{\alpha}^i w', \text{ dann } \varphi \in w' \\ & \text{Lem.26(*)} \\ & \text{gdw. } \alpha int : \varphi \in w. \end{aligned}$$

Es gibt mindestens eine maximal konsistente Menge $w_0 \in W$ mit $\Phi \subseteq w_0$. Da w_0 zu einer Äquivalenzklasse X_{i_0} von R gehört, gilt für alle $\varphi \in \Phi$, dass $\mathcal{M}, (X_{i_0}, h_{w_0}) \models \varphi$. Folglich ist jede konsistente Menge Φ erfüllbar. □

Somit ist mit den Erweiterungen um die Interaktionsaxiome zwischen dem Intentionsoperator und dem Handlungsoperator $bdi-stit_{dint}$ Logik eine vollständig und endlich axiomatisierbare Logik.

7 Zusammenfassung

Die vorliegende Arbeit stellte sich das Ziel, eine Logik von Handlungen und Überzeugungen zu erstellen, die wesentliche Aspekte der Begriffe 'Handlung' und 'Überzeugung' widerspiegelt. Dazu wurden zu Beginn dieser Arbeit drei Fragen aufgeworfen: Welcher Begriff von Handlung bzw. Handlungsbeschreibung sollte zugrunde gelegt werden? Welchen Bedingungen sollte die Zuschreibung von Überzeugungen bei rationalen Akteuren genügen? Wie wird beides einheitlich in einem logischen System beschrieben?

Die erste Frage wurde bereits im ersten Kapitel beantwortet. Eine modale Aussagenlogik mit einem deliberativen Stit-Operator ist am ehesten geeignet, Handlungen von unabhängigen Akteuren zu beschreiben. Diese Entscheidung zur Beschreibung von Handlungen mittels Stit Theorie wurde in den Kapiteln 2 und 3 untermauert. So wurde festgehalten, dass eine Handlung beschrieben werden kann, als das Hervorbringen eines nicht notwendigerweise bestehenden Sachverhaltes durch einen Akteur. Entgegen der Aussage von Belnap ist es trotz einer solchen Zuschreibung möglich, dass der Handlung ein Ereignis zugrunde liegt, in dem Sinne wie Davidson Handlungen definiert. Ob ein Ereignis eine Handlung eines Akteurs ist, richtet sich nach den *Beschreibungen* des Ereignisses. Bei Davidson ist ein Ereignis eine Handlung des Akteurs, wenn es unter einer Beschreibung als absichtlich vom Akteur hervorgebracht ist. Im Gegensatz dazu wurde gezeigt, dass der Begriff der Absichten und damit die Absichtlichkeit einer Handlung dem Begriff der Handlung nicht vorgelagert sein muss.

Um mittels Beschreibungen eines Ereignisses ein solches als Handlung eines Akteurs zu bestimmen, musste zuerst der Begriff der Beschreibung näher betrachtet werden. Unter einer Beschreibung eines Ereignisses wurde ein Sachverhalt verstanden, der das Ereignis charakterisiert. Einem Satz korrespondiert ein nicht-transienter Sachverhalt, so dass ein Sachverhalt als Beschreibung eines Ereignisses gewertet werden kann, wenn Zeit und Geschichtsverlauf durch das Zeitintervall des Ereignisses fixiert wurden. Anhand der Beschreibungen eines Ereignisses leiten sich die Bezeichnungen dessen ab. Danach wurde die sogenannte Stit-Paraphrase für Ereignisse festgehalten. Ein Ereignis ist eine Handlung des Akteurs, wenn jede Stit-paraphrasierte Bezeich-

7. Zusammenfassung

nung ebenfalls eine Bezeichnung des Ereignisses ist.

Es ist unbestritten, dass eine Stit-Paraphrase eine Handlung des Akteurs ausdrückt, ebenso dass die nicht-paraphrasierte Aussage das Ereignis charakterisiert. Bezweifelt wird die Möglichkeit, dass diese Aussage das Ereignis als Handlung des Akteurs kennzeichnet. Ereignisse sind schwer oder fast immer unmöglich voneinander abzugrenzen. Welcher Sachverhalt, dieses Ereignis und welcher nur ein Teilereignis und welcher bereits das nächste Ereignis charakterisiert, ist meist nicht zu ermitteln bzw. steht nicht fest. Es ist möglich, dass ein Sachverhalt mehrere Ereignisse charakterisiert bzw. eine unbestimmte Menge von Sachverhalten ein Ereignis charakterisiert. Eine Zuschreibung einer Handlung zu einem Akteur ist somit nicht durch die Berücksichtigung aller zur Zeit des Ereignisses bestehender Sachverhalte definierbar. Um zu beschreiben, dass eine Handlung des Akteurs vorliegt, genügt es jedoch einen Sachverhalt, der das Ereignis oder den mit dem Ereignis der Handlung erreichten Zustand charakterisiert, mittels einer Aussage zu erfassen und deren Stit-Paraphrase zu verwenden. Wenn eine Situation zu einem Zeitpunkt während oder nach dem Ereignis betrachtet wird, ist die paraphrasierte Aussage wahr. Der Akteur hat für das Bestehen des genannten, nunmehr nicht-transienten Sachverhaltes gesorgt. Dies drückt aus, dass er den Sachverhalt durch eine Handlung hervorgebracht hat, ohne die Handlung im Einzelnen zu beschreiben. Die zusammengefasste Aussage ist somit, dass einer Handlung immer ein Ereignis in der Welt zugrunde liegt, aber es durch eine Stit-Paraphrase einem Akteur zugeschrieben werden kann.

Von den Ereignissen unterscheiden sich die Zustände. Während Ereignisse Zeiträume einnehmen, sind Zustände punktuell. D.h. ein Zustand existiert zu einem Zeitpunkt. Das Vorliegen eines Zustandstyps, der durch eine Menge an Sachverhalten charakterisiert wird, ist wiederum ein Ereignis. Dies unterscheidet Handlungen von Überzeugungen. Einer Handlung liegt ein Ereignis zugrunde, einer Überzeugung ein mentaler Zustand. Jedoch erfolgt eine Zuschreibung einer Handlung oder einer Überzeugung zu einem Akteur immer in einer Situation. Zu einem festen Zeitpunkt in einem Geschichtsverlauf steht fest, dass der Akteur glaubt, dass p , im Falle der Überzeugung bzw. dass der Akteur dafür gesorgt hat, dass p , im Falle der Handlung. Für letzteres muss in der Situation p wahr sein. Der das Ereignis charakterisierende Sachverhalt ist durch den Akteur hervorgebracht worden.

Um zu verdeutlichen, dass es tatsächlich eine Einwirkung des Akteurs gegeben hat, wird der Handlungsbeschreibung in der deliberativen Stit Theorie zur Positiv-Bedingung, dass der Akteur mit seiner Wahlmöglichkeit den Sachverhalt sichergestellt hat, noch eine Negativ-Bedingung hinzugefügt. Es muss einen möglichen Geschichts-

verlauf zu dem Zeitpunkt geben, in dem der Sachverhalt ohne Zutun des Akteurs nicht eingetreten wäre. Diese aus Akteurssicht äußere Bedingung an die Umstände unterscheidet eine Zuschreibung einer Handlung von der eines mentalen Aktes. Für einen mentalen Akt ist die Diskussion einer Negativ-Bedingung, ob das mentale Vorkommnis ohne Zutun des Akteurs nicht zustande gekommen wäre, unsinnig. Die Antwort ist, dass, wenn der durch einen mentalen Akt hervorgebrachte Sachverhalt notwendigerweise besteht, dies ein Akt des Akteurs wäre, da das mentale Vorkommnis allein im Akteur ist und nicht an äußere Umstände gebunden ist. Definitorisch ist für einen mentalen Akt von Bedeutung, ob der Akteur sich für diesen Akt entschieden hat oder nicht. D.h. im Rahmen seiner Wahlmöglichkeit muss der den Akt charakterisierende Sachverhalt eingetreten sein. Die Positiv-Bedingung des deliberativen Stit-Operators genügt, um ein mentales Vorkommnis einem Akteur als mentalen Akt zuzuschreiben. Demzufolge lassen sich mentale Akte ebenfalls durch einen aussagenlogischen Operator darstellen.

Als somit geklärt war, wie eine Handlungsbeschreibung in einer verzweigenden Zeitlogik gefasst werden sollte und in welchem Verhältnis das Ereignis der Handlung bzw. des mentalen Aktes zu ihren jeweiligen Beschreibungen steht, wurde der zweiten Frage nachgegangen. Welchen Bedingungen sollte die Zuschreibung von Überzeugungen bei rationalen Akteuren genügen? Begonnen wurde damit, ob es sich bei einer Bildung einer Überzeugung um einen mentalen Akt handelt oder nicht.

Mentale Akte können vielfältig sein. Rechnen und Urteilen sind Beispiele dafür. Weitere Beispiele für mentale Akte sind das Bilden, Verwerfen und Beibehalten bestimmter mentaler Zustände. Jedoch wird nicht jeder mentale Zustand durch einen mentalen Akt allein gebildet. Es gibt Zustände, die diese Möglichkeit quasi ausschließen, da der Akteur keine direkte, instantane Kontrolle über deren Vorliegen hat. Ein Beispiel für einen solchen Zustand ist der der Überzeugung. Es ist einem Akteur nicht möglich, für eine konkrete Proposition p in der aktuellen Situation zu entscheiden, die Überzeugung, dass p , zu haben oder nicht.

Ein Beispiel für einen mentalen Zustand, der durch einen mentalen Akt gebildet wird, ist der der Intention. Als Akteur kann ich mich entscheiden, dass ich beabsichtige, dass p , oder dass ich nicht länger beabsichtige, dass p . Diese Entscheidung kann ich direkt, instantan und nur im Hinblick auf p treffen. Diese Kontrolle, die ein Akteur über seine Intentionen hat, war ein Grund dafür, dass Intentionen sich weder als Konglomerat von Wünschen und Überzeugungen verstehen noch aus ihnen eindeutig herzuleiten sind. Die drei mentalen Zustandstypen Wunsch, Überzeugung und Intention sind unabhängig voneinander aus mehreren Gründen. Zum einen ist es die Kontrolle,

7. Zusammenfassung

die ein Akteur über letztere hat, über Überzeugungen nicht hat und in der Regel über seine Wünsche nicht ausübt. Des Weiteren sind es die Funktionen, die der jeweilige Zustand im Akteur übernimmt. Während die Überzeugungen, die Welt dem Akteur präsentieren, wie sie ist, stellen die Wünsche, die Welt dem Akteur als etwas Hervorzubringendes dar. Aber erst der Zustand der Intentionen, der zum Gehalt hat, wie die Welt sein wird, und dessen Gehalt als möglich durch den Akteur hervorzubringend vom Akteur verstanden wird, tritt in Interaktionen mit den Handlungen des Akteurs. Daher wurde aus der Frage, wie eine Überzeugung dem Akteur zugeschrieben werden kann, die Fragen: Welche mentalen Zustände neben der Überzeugung spielen im Zusammenhang mit Handlungen eine Rolle? Und welche Zustände gehen mit Handlungen Interaktionen ein, die sich in logischen Gesetzmäßigkeiten festhalten lassen?

Es wurde gezeigt, dass die Wünsche eines Akteurs Aufschluß darüber geben, warum ein Akteur eine bestimmte Handlung ausführt. Jedoch sind Wünsche auch von idealisierten, rationalen Akteuren nicht notwendigerweise widerspruchsfrei. Es ist einem idealisierten, rationalen Akteur sicher nicht zuschreibbar, sich etwas Paradoxes zu wünschen. Das kann er nicht als hervorbringend begreifen. Aber er kann sich wünschen, dass p , und ebenfalls wünschen, dass $\neg p$. Diese Antagonismen bei Wünschen sind nicht sofort irrational. Desgleichen gilt für Überzeugungen. Zu einem gegebenen Zeitpunkt ist es aufgrund der Umstände, wie sich die Welt dem Akteur präsentiert, möglich, dass er die Überzeugungen, dass p und dass $\neg p$, hat. Irrational wäre es lediglich, wenn eine inkonsistente Überzeugung vorliegen würde. Das Vorliegen konfligierender Überzeugungen in einem Moment ist nicht notwendigerweise irrational. Aufgrund der fehlenden Kontrolle und der Funktion von Überzeugungen, ist es nicht sinnvoll, eine Agglomeration von Überzeugungsgehalten zu fordern. Somit ist es möglich, zwischen der Rationalität von inkonsistenten und konfligierenden Überzeugungen und von paradoxen und antagonistischen Wünschen zu unterscheiden.

Für rationale Akteure schien in der Diskussion dieser Arbeit diese Unterscheidung im Hinblick auf die Rationalität von Intentionen weniger sinnvoll. Für den Gehalt einer Intention entscheidet sich der Akteur. Es ist ihm nicht durch die Umgebung vorgegeben. Auf der anderen Seite ist eine Intention der mentale Zustand, der im Akteur vorliegt, wenn er sich vorstellt, was er hervorbringen möchte. Um in der Welt etwas Hervorzubringen, stehen dem Akteur Handlungen zur Verfügung. Der Gehalt einer Intention richtet sich daher immer auf das Ausführen einer Handlung. Wenn ein Wunsch die Antwort auf die Warum-Frage bei einer Handlung geben kann, so beantworten die Intentionen die Frage, wie ein Akteur einen Sachverhalt umsetzen möchte. Einer Intention, nach Paris zu fahren, folgt, die nächste. Für das Ziel nach Paris zu fahren, ist

es nötig, eine Fahrkarte zu kaufen und weitere Handlungen zu planen und Intentionen zu bilden, so dass von einem rationalen Akteur eine Kohärenz zwischen der Intention, bestimmte Ziele zu erreichen, und dem Intendieren dafür notwendiger Mittel gefordert werden kann.

Somit wurde zwar in dieser Arbeit festgehalten, dass sich keine Interaktionen zwischen Überzeugungen und Handlungen begründen lassen, die in logischen Gesetzmäßigkeiten dargestellt werden können, auch dann nicht, wenn auf Rationalitätsforderungen zurückgegriffen wird. Es wurden jedoch Interaktionen zwischen dem Vorliegen von Intentionen und dem Durchführen von Handlungen begründet, die sich sowohl auf begriffliche als auch rationale Forderungen stützen. Darauf aufbauend wurde in den letzten beiden Kapiteln eine Antwort auf die dritte Frage gegeben: Wie werden diese mentalen Zustände, die Handlungen und Interaktionen einheitlich in einem logischen System beschrieben?

Eine Intention, dass p , ist ein mentaler Zustand, der jedoch durch den Sachverhalt charakterisiert wird, der der Aussage korrespondiert, dass der Akteur in diesem Zustand ist. Somit ist das Vorliegen einer Intention in einem Moment durch einen aussagenlogischen Operator beschreibbar. Nun könnte man einwenden, dass der Gehalt von Intentionen sich auf Handlungen richtet. Dies ist korrekt. Aber wie ich bei der Beschreibung von Handlungen festgehalten habe, genügt ein die Handlung charakterisierender Sachverhalt, für den der Akteur sorgt, um die Handlung zu beschreiben. Eine Intention $\alpha int:p$ beinhaltet, dass der Akteur α beabsichtigt, für p zu sorgen, $\alpha int:dstit:p$. Da eine Handlung ebenfalls durch einen aussagenlogischen Operator beschrieben werden kann, wird auch eine Intention durch einen solchen instantanen, aussagenlogischen Operator beschrieben.

Gilt dies für die mentalen Zustände wie Überzeugung und Wunsch ebenso? Der Gehalt einer Überzeugung ist eine Proposition und korrespondiert einem Sachverhalt. Des Weiteren ist das Vorliegen des mentalen Zustandes der Überzeugung durch den Sachverhalt beschrieben, dass der Akteur instantan in diesem Zustand ist. Somit lässt sich auch eine Überzeugung in einer aussagenlogischen Modallogik beschreiben. Für den Gehalt eines Wunsches ist es sicherlich strittig, ob der Gehalt propositional sein muss. Ich würde annehmen, dass der Gehalt nicht notwendigerweise propositional ist, dass aber bei der Zuschreibung eines Wunsches zu einem Akteur zur Vereinfachung der Gehalt des Wunsches als propositional dargestellt wird. So wurden auch für die mentalen Zustände aussagenlogische Operatoren eingeführt. Den Schluss der Arbeit bildete ein Vorschlag wie diese Operatoren in einer verzweigenden Zeitlogik interpretiert werden sollten, um die in den vorangegangenen Kapiteln motivierten Ei-

7. Zusammenfassung

genschaften zu haben und die genannten Interaktionen einzugehen. Dies wurde in der *bdi-stit* Logik umgesetzt.

Hierbei ist festzuhalten, dass die *bdi-stit* Logik so allgemein gehalten werden konnte, dass sich andere Logiken wie *BDI*-Logiken durch diese Logik generalisieren lassen. So ist es möglich, durch bestimmte Bedingungen an die Modelle der *bdi-stit* Logik konkrete Interaktionsaxiome zu validieren, die in *BDI*-Logiken postuliert werden. Die Axiome für die Überzeugungs-, Wunsch- und Intentionoperatoren, die in *BDI*-Logiken diskutiert werden, ließen sich durch entsprechende Forderungen an die Modelle der *bdi-stit* Logik ebenso validieren. Dies ist möglich. Es war jedoch nicht Teil dieser Arbeit, sondern lässt sich als ein Anknüpfungspunkt für weitere Anwendungen der *bdi-stit* Logik verstehen.

Es war unter anderem deswegen nicht Teil dieser Arbeit, weil bei der Untersuchung der in *BDI*-Logiken vertretenen Axiome für Überzeugung-, Wunsch- und Intentionoperatoren sich herausstellte, dass sich diese nicht bzw. nur ungenügend motivieren lassen. Es finden sich weder deskriptive noch normative Aspekte des Begriffes ‘Wunsch’, die eine Beschreibung dieses mentalen Zustandes in einer relationale Semantik rechtfertigen oder erklären würden. Im Falle der Überzeugung wird häufig erklärt, dass die normativen Aspekte erfordern, dass es nur rational ist, Überzeugungen zu haben, die konsistent sind und agglomerieren. Dem wurde jedoch widersprochen. Die Axiome A_b und D_b wurden abgelehnt. Die normativen Aspekte des Begriffes ‘Überzeugung’ sind auf Wahrheit ausgerichtet. Es ist nicht rational zu fordern, möglichst viele Überzeugungen anzuhäufen, sondern möglichst korrekte Überzeugungen zu haben. Das Aufgeben einer Überzeugung kann nicht deswegen verlangt werden, weil eine dazu konfligierende Überzeugung vorliegt. Der Akteur ist nicht rational, wenn er eine der beiden Überzeugungen instantan aufgibt. Daher ist es in *bdi-stit* Logik möglich, $\alpha \text{ bel}:\varphi \wedge \alpha \text{ bel}:\neg\varphi$ zu erfüllen. Ein Akteur wäre rational, wenn er sich der Aufgabe stellt herauszufinden, welche seiner Überzeugungen korrekt ist. Dies lässt sich jedoch nicht in den Schemata eines instantanen Überzeugungsoperators ausdrücken.

Dass es keine in logischen Schemata zu beschreibenden Zusammenhänge zwischen Handlungen auf der einen Seite und Wünsche und Überzeugungen des Akteurs auf der anderen Seite gibt, führt dazu, dass keine Interaktionsaxiome zwischen den jeweiligen Operatoren in *bdi-stit* Logik festgehalten worden sind. Auch hier besteht die Möglichkeit in weiteren Untersuchungen nachzuforschen, unter welchen Bedingungen solche Interaktionen sinnvoll sind. Es wäre eine Herangehensweise der folgenden Art denkbar: Wenn ein bestimmtes Axiom hinzugenommen wird, welche weiteren Theoreme

ergeben sich aus der Bedingung, die dieses Axiom an die Modelle der *bdi-stit* Logik stellt?

Im Gegensatz zu Wünschen und Überzeugungen sind die Intentionen eines Akteurs auf seine Handlungen gerichtet. Diese enge Verbindung von Handlungen und Intentionen hat ergeben, dass diese nicht nur durch analoge Operatoren zu beschreiben sind, sondern auch, dass sich vielfältige Interaktionsschemata zwischen diesen Operatoren zusammentragen lassen. Einige wurden in dieser Zusammenfassung verbalisiert, für die anderen vgl. *bdi-stit_{dint}* Logik in Abschnitt 6.4.

Einen Ansatzpunkt für mögliche Erweiterungen bietet die darunter liegende Zeitlogik, die in dieser Arbeit nur am Rande berücksichtigt wurde. Es ist möglich, Operatoren bzw. Interaktionsschemata nicht nur instantan in einem Zeitpunkt zu bewerten, sondern entlang eines Geschichtsverlaufs. Dies würde Beschreibungen zulassen, dass Akteure sich gegenseitig beeinflussen, obwohl sie zu den jeweiligen Momenten unabhängig sind. Es wäre ebenso möglich, bestimmte Interaktionen von Operatoren über Zeitpunkte hinweg zu beschreiben. So z.B. dürften konfligierende Überzeugungen nicht über mehrere Zeitpunkte innerhalb einer Geschichte hinweg vorliegen. Oder möglicherweise würde gefordert werden, dass der Akteur, wenn er konfligierende Überzeugungen hat, diese solange hat, bis er dafür gesorgt hat, dass er diese aufhebt.

Dem Leser, der sich befremdete, dass es rational sein kann, konfligierende Überzeugungen zu haben, und dass es nicht möglich ist, in einem Moment für eine bestimmte Überzeugung zu sorgen, möchte ich einen Ausweg aufzeigen. Entlang eines Geschichtsverlaufs kann ein Akteur indirekt und eigenschaftsbezogen Einfluß auf seine Überzeugungen nehmen, indem er Handlungen ausführt, die die Qualität, also die Wahrheitsgerichtetheit, seiner Überzeugungen beeinflussen können. Es ist eine Erweiterung der *bdi-stit* Logik denkbar, die Interaktionen nicht instantan beschreibt, sondern mittels temporalen Operatoren entlang von Geschichtsverläufen. Diese nicht-instantane Beschreibung von Interaktionsschemata stellt einen interessanten Ausblick auf weitere Untersuchungen dar.

So ist die in dieser Arbeit vorgestellte *bdi-stit* Logik als eine Logik von Handlungen und Überzeugungen zu verstehen, die aus der Zustandsperspektive beschreibt, in welchen logischen Beziehungen Handlungen und Überzeugungen zu einem festen Zeitpunkt stehen. Die Erweiterungsmöglichkeiten, diese logischen Beziehungen über Zeiträume hinweg festzuhalten, sind per se in der darunter liegenden Zeitlogik schon mitgegeben und bieten jede Menge Anregungen, sich weiter mit logischen Beziehungen zwischen Handlungen, Überzeugungen und Intentionen zu beschäftigen.

Abbildungsverzeichnis

1.1	Branching Time Struktur	15
2.1	Sachverhalt und Situation	45
2.2	Zustand und Ereignis als konkrete Zeitpunkte bzw. Zeitintervalle . . .	55
6.1	$\Box \alpha \text{ dint} : \varphi \wedge \neg \Box \varphi$	297

Tabellenverzeichnis

2.1	Zustand und Ereignis - Sachverhalte in Situationen und Geschichten . . .	57
3.1	Vendlers Unterscheidungen in Bezug auf Handlung und Vorgang . . .	77
3.2	Einstellungen als mentale Zustände und Ereignisse	93
5.1	Axiome des Überzeugungsoperators in BDI_{CTL}^K , in $B^{KD45}D^{KD}I_{CTL}^{KD}$ und in <i>bdi-stit</i> Logik	237
5.2	Axiome des Wunschoperators in BDI_{CTL}^K , in $B^{KD45}D^{KD}I_{CTL}^{KD}$ und in <i>bdi-stit</i> Logik	238
6.1	Identitätsregeln für Akteursvariablen	260
6.2	Strukturregeln für die BT+AC Struktur I	261
6.3	Strukturregeln für die BT+AC Struktur II	262
6.4	Operatorregeln für eine Stit-Logik mit <i>dstit</i> -Operator	265
6.5	Operatorregeln für die mentalen Zustandsoperatoren in <i>bdi-stit</i> Logik	266
6.6	Ein Beispiel für ein unendliches, vollständiges, offenes Tableau	280
6.7	Handlungsunabhängigkeit der Akteure	281
6.8	Offenes Tableau der konfigrierenden Überzeugungen	283
6.9	Axiomatisierung des <i>int</i> -Operators in <i>bdi-stit</i> _{<i>dint</i>} Logik	299
6.10	Theoreme für den Intensionsoperator in <i>bdi-stit</i> _{<i>dint</i>} - Darstellung der Interaktionsschemata zwischen Handlungen und Absichten	300

Literaturverzeichnis

- [1] C. Adam, A. Herzig, und D. Longin. A logical formalization of the OCC theory of emotions. *Synthese*, 168(2):201–248, 2009.
- [2] R.M. Adams. Theories of actuality. *Noûs*, 8(3):211–231, 1974.
- [3] K. Ajdukiewicz. Das Weltbild and die Begriffsapparatur. *Erkenntnis*, 4:259–287, 1934.
- [4] W.P. Alston. *Epistemic Justification - Essays in the Theory of Knowledge*. Cornell University Press, Ithaca, London, 1989.
- [5] W.P. Alston. Epistemic desiderata. *Philosophy and Phenomenological Research*, 53(3):527–551, 1993.
- [6] R. Alur, T.A. Henzinger, und O. Kupferman. Alternating-time temporal logic. *J. ACM*, 49(5):672–713, 2002.
- [7] G.E.M. Anscombe. *Absicht*. Alber Verlag, Freiburg, 1986. übers. von J.M. Connolly und T. Keutner (*Intention*, 1963).
- [8] R. Audi. Intending, intentional action, and desire. In J. Marks, Hrsg., *The Ways of Desire: New Essays in Philosophical Psychology on the Concept of Wanting*, S. 17–38. Transaction Publishers, 1986.
- [9] R. Audi. Belief, faith, and acceptance. *International Journal of Philosophy of Religion*, 63(1/3):87–102, 2008.
- [10] J.L. Austin. How to talk: Some simple ways. *Proc. of the Aristotelian Society*, 53:227–246, 1953.
- [11] A.C. Baier. Act and intent. *Journal of Philosophy*, 67(19):648–658, 1970.
- [12] A.C. Baier. The intentionality of intentions. *The Review of Metaphysics*, 30(3):389–414, 1977.

Literaturverzeichnis

- [13] J. Barwise. Information and impossibilities. *Notre Dame Journal of Formal Logic*, 38(4):488–515, 1997.
- [14] J. Barwise und J. Perry. *Situations and Attitudes*. MIT Press, Cambridge MA, 1988.
- [15] N.D. Belnap. Backwards and forwards in the modal logic of agency. *Philosophy and Phenomenological Research*, 51(4):777–807, 1991.
- [16] N.D. Belnap. Branching space-time. *Synthese*, 92(3):385–434, 1992.
- [17] N.D. Belnap und J.F. Horty. The deliberative stit: A study of action, omission, ability and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.
- [18] N.D. Belnap und M. Perloff. Seeing to it that: a canonical form for agentives. *Theoria*, 54(3):175–199, 1988.
- [19] N.D. Belnap, M. Perloff, und M. Xu. *Facing the Future: Agents and Choices in our Indeterminist World*. Oxford University Press, New York, 2001.
- [20] H. Ben-Yami. Against characterizing mental states as propositional attitudes. *The Philosophical Quarterly*, 47(186):84–89, 1997.
- [21] J. Bennett. *Events and Their Names*. Clarendon Press, Oxford, 1988.
- [22] P.A. Boghossian. The normativity of content. *Philosophical Issues*, 13(1):31–45, 2003.
- [23] M. Bratman. *Structures of agency: essays*. Oxford University Press, 2007.
- [24] M.E. Bratman. Two faces of intention. *The philosophical Review*, 93(3):375–405, 1984.
- [25] M.E. Bratman. *Intentions, Plans, and Practical Reason*. Harvard University Press, Cambridge MA, 1987.
- [26] M.E. Bratman. Practical reasoning and acceptance in a context. *Mind*, 101(401):1–16, 1992.
- [27] M.E. Bratman. Intention, belief, practical, theoretical reasoning. In S. Robertson, Hrsg., *Spheres of Reason*, S. 29–61. Oxford University Press, Oxford, 2009.

- [28] J. Broersen. A complete stit logic for knowledge and action, and some of its applications. In M. Baldoni, T.C. Son, M. Birna van Riemsdijk, und M. Wini-koff, Hrsg., *Declarative Agent Languages and Technologies VI, 6th International Workshop, DALT 2008, Estoril, Portugal, May 12, 2008*, Volume 5397 von *Lecture Notes in Computer Science*, S. 47–59, 2009.
- [29] J. Broersen. Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of applied logic*, 9(2):127–152, 2011.
- [30] J. Broersen, A. Herzig, und N. Troquard. A normal simulation of coalition logic and an epistemic extension. In D. Samet, Hrsg., *Proceedings theoretical aspects Rationality and Knowledge, TARK XI*, S. 92–101. ACM Digital Library, 2007.
- [31] J.M. Broersen, A. Herzig, und N. Troquard. Embedding Alternating-time Tem-poral Logic in strategic STIT logic of agency. *Journal of Logic and Computa-tion*, 16(5):559–578, 2006.
- [32] J.M. Broersen, A. Herzig, und N. Troquard. From coalition logic to STIT. In *Proceedings LCMAS 2005*, Volume 157 von *Electronic Notes in Theoretical Computer Science*, S. 23–35. Elsevier, 2006.
- [33] J. Broome. Normative requirements. *Ratio*, 12(4):398–419, 1999.
- [34] J. Broome. Practical reasoning. In J.L. Bermúdez und A. Millar, Hrsg., *Reason and Nature: Essays in the Theory of Rationality*, S. 85–113. Oxford University Press, 2002.
- [35] J. Broome. Is rationality normative? *Disputatio*, 11:153–171, 2008.
- [36] A.A. Buckareff. How (not) to think about mental action. *Philosophical Explorations*, 8(1):83–89, 2005.
- [37] K. Bykvist und A. Hattiangadi. Does thought imply ought? *Analysis*, 67(4):277–8211, 2007.
- [38] B. Chellas. *The Logical Form of Imperatives*. PhD thesis, Stanford University, Perry Lane Press, California, 1969.
- [39] B. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, Cam-bridge, 1980.

Literaturverzeichnis

- [40] R.M. Chisholm. Freedom and action. In Keith Lehrer, Hrsg., *Freedom and Determinism*. Random House, 1966.
- [41] R.M. Chisholm. Events and propositions. *Noûs*, 4(1):15–24, 1970.
- [42] R.M. Chisholm. The structure of intention. *Journal of Philosophy*, 67(19):633–647, 1970.
- [43] R.M. Chisholm. *Person and Object: a Metaphysical Study*. La Salle: Open Court, Hudson, Illinois, 1976.
- [44] R.M. Chisholm. Events without times: An essay on ontology. *Noûs*, 24(3):413–427, 1990.
- [45] L.J. Cohen. Belief and acceptance. *Mind*, 98(391):367–389, 1989.
- [46] L.J. Cohen. *An Essay on Belief and Acceptance*. Clarendon Press, Oxford, 1992.
- [47] P.R. Cohen und H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261, 1990.
- [48] D. Davidson. *Essays on Action and Events*. Oxford University Press, New York, 1980.
- [49] H. de Swart. Aspect shift and coercion. *Natural Language and Linguistics Theory*, 16(2):347–358, 1998.
- [50] E.A. Emerson. Temporal and modal logic. In Jan van Leeuwen, Hrsg., *Handbook of theoretical computer science (vol. B)*, S. 995–1072. MIT Press, Cambridge, MA, USA, 1990.
- [51] E.A. Emerson und J.Y. Halpern. Decision procedures and expressiveness in the temporal logic of branching time. In *STOC*, S. 169–180, 1982.
- [52] E.A. Emerson und J.Y. Halpern. “Sometimes” and “not never” revisited: on branching versus linear time temporal logic. *J. ACM*, 33(1):151–178, 1986.
- [53] E.A. Emerson und A.P. Sistla. Deciding branching time logic. In *Proc. 16th ACM Symposium on Theory of Computing*, S. 14–24. ACM, New York, NY, USA, 1984.

- [54] P. Engel. Truth and the aim of belief. In D. Gillies, Hrsg., *Laws and Models in Science*, S. 77–97. College Publication, London, 2004.
- [55] R. Fagin und J.Y. Halpern. Belief, awareness and limited reasoning. *Artificial Intelligence*, 34(1):39–76, 1988.
- [56] R. Fagin, J.Y. Halpern, Y. Moses, und M. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge MA, 1995.
- [57] M. Fara. Masked abilities and compatibilism. *Mind*, 117(468):843–865, 2008.
- [58] R. Feldman. The ethics of belief. *Philosophy and Phenomenological Research*, 60(3):667–695, 2000.
- [59] S. Finlay. Motivation to the means. In D.K. Chan, Hrsg., *Moral Psychology Today: Values, Rational Choice, and the Will*, S. 173–191. Springer, 2008.
- [60] H. Frankfurt. Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66(3):829–839, 1969.
- [61] H. Frankfurt. Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1):5–20, 1971.
- [62] D. Gabbay, I. Hodkinson, und M. Reynolds. *Temporal Logic - Mathematical Foundations and Computational Aspects, Volume I*. Oxford University Press, Oxford, 1994.
- [63] D. Gabbay, M. Reynolds, und M. Finger. *Temporal Logic - Mathematical Foundations and Computational Aspects, Volume II*. Oxford University Press, Oxford, 2000.
- [64] M.P. Georgeff und A.S. Rao. Modeling rational agents within a bdi-architecture. In R. Fikes J. Allen und E. Sandewall, Hrsg., *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, S. 473–484. Morgan Kaufmann, San Mateo, 1991.
- [65] M.P. Georgeff und A.S. Rao. Formal models and decision procedures in multi-agent systems. *Technical Note*, 61:293–342, 1995.
- [66] M.P. Georgeff und A.S. Rao. Decision procedures for bdi logics. *Journal of Logic and Computation*, 8(3):293–342, 1998.

Literaturverzeichnis

- [67] B. Gertler. *Self-Knowledge*. Routledge and Kegan Paul, London, 2008.
- [68] B. Gertler. Self-knowledge. In *Stanford Encyclopedia of Philosophy*. 2008. <http://plato.stanford.edu/entries/self-knowledge>, 28.10.2008.
- [69] C. Ginet. Deciding to believe. In M. Steup, Hrsg., *Knowledge, Truth and Duty*, S. 63–76. Oxford University Press, 2001.
- [70] H. Götz. *Lateinisch-Althochdeutsch-Neuhochdeutsches Wörterbuch*. Akademischer Verlag, Berlin, 1999.
- [71] S. Hampshire. *Thought and action*. Chatto & Windus, London, 1959.
- [72] G. Harman. *Change in View*. MIT Press, 1986.
- [73] G. Harman. Practical reasoning. In *Reasoning, Meaning, and Mind*, S. 46–74. Oxford University Press, 1999.
- [74] G. Harman. Practical aspects of theoretical reasoning. In A.R. Mele und P. Rawling, Hrsg., *The Oxford Handbook of Rationality*, S. 45–56. Oxford University Press, 2004.
- [75] J. Heil. Privileged access. *Mind*, 97(386):238–251, 1988.
- [76] D.P. Henry. *The logic of Saint Anselm*. Oxford University Press, Oxford, 1967.
- [77] E. Hentschel und H. Weydt. *Handbuch der deutschen Grammatik*. Walter de Gruyter, Berlin, 2003.
- [78] A. Herzig und D. Longin. A logic of intention with cooperation principles and with assertive speech acts as communication primitives. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2, AAMAS '02*, S. 920–927. ACM, New York, NY, USA, 2002.
- [79] A. Herzig und E. Lorini. A dynamic logic of agency i: Stit, capabilities and powers. *Journal of Logic, Language and Information*, 19(1):89–121, 2010.
- [80] J. Higginbotham. The logic of perceptual reports: an extensional alternative to situation semantics. *The Journal of Philosophy*, 80(February):100–127, 1983.
- [81] J. Hintikka. *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca NY, 1962.

- [82] Jaakko Hintikka. Impossible possible worlds vindicated. *Journal of Philosophical Logic*, 4(3):475–484, 1975.
- [83] J. Horty. An alternative stit operator. *Manuscript*, unpublished, 1989.
- [84] F. Howard-Snyder. „Cannot“ implies „Not Ought“. *Philosophical Studies*, 130(2):233–246, 2006.
- [85] I.L. Humberstone. Direction of fit. *Mind*, 101(401):59–83, 1992.
- [86] D. Hume. A treatise of human nature: Being an attempt to introduce the experimental method of reasoning into moral subjects. <http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3ll3/hume/treat.html>, 26. Mai 1998.
- [87] U. Hustadt, C. Dixon, R.A. Schmidt, M.Fisher, J.-J. Ch. Meyer, und W. van der Hoek. Reasoning about agents in the KARO framework. In *Proceedings. Eighth International Symposium on Temporal Representation and Reasoning*, S. 206–213, 2001.
- [88] D. Kaplan. Demonstratives. In J. Almog, J. Perry, und H. Wettstein, Hrsg., *Themes from Kaplan*, S. 481–564. Oxford University Press, Oxford, 1989.
- [89] G. Katz. A semantic account of the stative adverb gap. In C. Fabricius-Hansen, E. Lang, und C. Maienborn, Hrsg., *Approaching the Grammar of Adjuncts*, ZAS papers in Linguistics 17, S. 135–151. Blackwell, Berlin, 2000.
- [90] T.S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962.
- [91] L. Lamport. 'Sometime' is sometimes 'Not Never' — but not always. In *Proc. 7th ACM Symposium on the Principles of Programming Languages*, S. 174–185. Association for Computing Machinery, Las Vegas, 1980.
- [92] J.H. Levesque. A logic of implicit and explicit belief. In *Proceedings of the Fourth National Conference on Artificial Intelligence*, S. 198–202. American Association for Artificial Intelligence, 1984.
- [93] D. Lewis. *Counterfactuals*. Harvard University Press, Cambridge Ma, 1973.
- [94] B. Linsky und E.N. Zalta. In defense of the simplest quantified modal logic. In *Philosophical Perspectives 8: Mally, E., 1912, Gegenstandstheoretische Grundlagen der Logik und Logistik, Leipzig:Barth*, S. 431–458, 1994.

Literaturverzeichnis

- [95] B. Linsky und E.N. Zalta. In defense of the contingently concrete. *Philosophical Studies*, 84(2-3):283–294, 1996.
- [96] S. Löbner. Ansätze zu einer integralen semantischen Theorie von Tempus, Aspekt und Aktionsarten. In V. Ehrlich und H. Vater, Hrsg., *Temporalsemantik: Beiträge zur Linguistik der Zeitreferenz*, S. 163–191. Niemeyer, Tübingen, 1988.
- [97] E. Lorini und A. Herzig. A logic of intention and attempt. *Synthese*, 163(1):45–77, 2008.
- [98] E. Mally. *Gegenstandstheoretische Grundlagen der Logik und Logistik*. Barth, Leipzig, 1912.
- [99] M. Moens und M. Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28, 1988.
- [100] R. Montague. Universal grammar. *Theoria*, 36(3):373–398, 1970.
- [101] T. Nagel. *The Possibility of Altruism*. Oxford Clarendon Press, 1970.
- [102] N. Nottelmann. *Blameworthy Belief: A Study in Epistemic Deontologism*. Springer Verlag, Berlin, 2007.
- [103] T. Parsons. *Events in the semantics of English: a study of subatomic semantics*. MIT Press, Cambridge MA, 1990.
- [104] M. Pauen. *Illusion Freiheit?: Mögliche und unmögliche Konsequenzen der Hirnforschung*. Fischer Verlag, Frankfurt, 2006.
- [105] M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.
- [106] A. Plantinga. On existentialism. *Philosophical Studies*, 44(1):1–20, 1983.
- [107] A. Plantinga. *Essays in the Metaphysics of Modality*. In M. Davidson, Hrsg., *Essays in the Metaphysics of Modality*. Oxford University Press, London, 2003.
- [108] M. Platts. *Ways of Meaning*. Routledge and Kegan Paul, London, 1979.
- [109] J.L. Pollock. Plantinga on possible worlds. In J. Tomberlin und P. van Inwagen, Hrsg., *Alvin Plantinga*, S. 121–144. D. Reidel Press, Dordrecht, 1985.

- [110] P. Priest. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge MA, 2001.
- [111] A. Prior. *Past, Present, and Future*. Oxford University Press, Oxford, 1967.
- [112] J. Proust. A plea for mental acts. *Synthese*, 129(1):105–128, 2001.
- [113] D. Radcliffe. Scott-Kakures on believing at will. *Philosophy and Phenomenological Research*, 57(1):145–51, 1997.
- [114] A.S. Rao und M.P. Georgeff. Intentions and rational commitment. *Technical Note*, 8:293–342, 1993.
- [115] M. Reynolds. An axiomatization of full computation tree logic. *Journal of Symbolic Logic*, 66(3):1011–1057, 2001.
- [116] M. Reynolds. Axioms for branching time. *Journal of Logic and Computation*, 12(4):679–697, 2002.
- [117] G. Ryle. *The Concept of Mind*. Hutchinson, London, 1949.
- [118] R.M. Sainsbury. *Paradoxien*. Philipp Reclam jun., Stuttgart, 2010. übers. von V.C. Müller und Volker Ellerbeck (*Paradoxes*, 2009).
- [119] J. Sakalauskaite. A sequent calculus for propositional dynamic logic for agents with interactions. *Lithuanian Mathematical Journal*, 45(2):217–224, 2005.
- [120] U. Scheffler. *Ereignis und Zeit: Ontologische Grundlagen der Kausalrelation*. Logos Verlag, Berlin, 2001.
- [121] D. Scott. Advice on modal logic. In K. Lambert, Hrsg., *Philosophical Problems in Logic: Some Recent Developments*, S. 143–173. D. Reidel, Dordrecht, 1970.
- [122] D. Scott-Kakures. On belief and captivity of the will. *Philosophy and Phenomenological Research*, 54(1):77–103, 1994.
- [123] J.R. Searle. *Intentionalität. Eine Abhandlung zur Philosophie des Geistes*. Suhrkamp Verlag, Frankfurt am Main, 1991. übers. von H.P. Gavagai (*Intentionality. An essay in the philosophy of mind*, 1983).
- [124] K. Segerberg. Bringing it about. *Journal of Philosophical Logic*, 18(4):327–347, 1989.

Literaturverzeichnis

- [125] K. Segerberg. Getting started: Beginnings in the logic of action. *Studia logica*, 51(3-4):347–378, 1992.
- [126] C. Semmling und H. Wansing. From *bdi* and *stit* to *bdi-stit* logic. *Logic and Logical Philosophy*, 17:185–207, 2008.
- [127] C. Semmling und H. Wansing. A sound and complete axiomatic system of *bdi-stit* logic. In M. Pelis, Hrsg., *The Logica Yearbook 2008*, S. 193–211. College Publications, London, 2009.
- [128] C. Semmling und H. Wansing. Reasoning about belief revision. In E.J. Olsson & S. Enqvist, Hrsg., *Belief Revision Meets Philosophy of Science*, S. 303–328. Springer Verlag, Berlin, 2011.
- [129] M. Smith. The humean theory of motivation. *Mind*, 96(381):36–61, 1987.
- [130] D. Sobel und D. Copp. Against direction of fit accounts of belief and desire. *Analysis*, 61(1):44–53, 2001.
- [131] Roy A. Sorensen. *Blindspots*. Oxford University Press, 1988.
- [132] R.C. Stalnaker. A theory of conditionals. In N. Rescher, Hrsg., *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series 2, S. 98–112. Blackwell, Oxford, 1968.
- [133] R.C. Stalnaker. Possible worlds. *Noûs*, 10(1):65–75, 1976.
- [134] R.C. Stalnaker. *Inquiry*. Bradford Books, MIT Press, 1985.
- [135] R.C. Stalnaker. The problem of logical omniscience I. *Synthese*, 89(3):425–440, 1991.
- [136] R.C. Stalnaker. *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Oxford University Press, MIT, 2003.
- [137] R.C. Stalnaker und R.H. Thomason. A semantic theory of adverbs. *Linguistic inquiry*, 4(2):195–220, 1973.
- [138] G. Strawson. Mental ballistics or the involuntariness of spontaneity. *Proceedings of the Aristotelian Society*, 76(3):227–256, 2003.

- [139] K. Su, A. Sattar, H. Lin, und M. Reynolds. A modal logic for beliefs and pro attitudes. In *Proceedings of the 22nd national conference on Artificial intelligence*, Volume 1 von *AAAI'07*, S. 496–501. AAAI Press, 2007.
- [140] P. Thagard. Desires are not propositional attitudes. *Dialogue*, 45(1):151–156, 2006.
- [141] R. Thomason. Combinations of tense and modality. In D. Gabbay und F. Guethner, Hrsg., *Handbook of Philosophical Logic, Volume II: Extensions of Classical Logic*, S. 135–165. Reidel, Dordrecht, 1984.
- [142] M. Thompson. *Life and Action*. Harvard University Press, Cambridge MA, 2008.
- [143] R. Tuomela. Belief versus acceptance. *Philosophical Explorations*, 3(2):122–137, 2000.
- [144] W. van der Hoek und M. Wooldridge. Towards a logic of reational agency. *Logic Journal of the IGPL*, 11(2):133–157, 2003.
- [145] B. van Linder, W. van der Hoek, und J.-J. Ch. Meyer. Formalising abilities and opportunities of agents. *Fundamentae Informaticae*, 34(1-2):53–101, 1998.
- [146] M. Vardi. On epistemic logic and logical omniscience. In J.Y. Halpern, Hrsg., *Theoretical Aspects of Reasoning about Knowledge. Proceedings of the 1986 Conference*, S. 293–305. Morgan Kaufmann Publishers, Los Altos, 1986.
- [147] J.D. Velleman. Practical reflection. *The Philosophical Review*, 94(1):33–61, 1985.
- [148] J.D. Velleman. The guise of the good. *Noûs*, 26(1):3–26, 1992.
- [149] J.D. Velleman. *The Possibility of Practical Reason*. Oxford University Press, Oxford, 2000.
- [150] J.D. Velleman. What good is a will? In *Action in Context*. de Gruyter/Mouton, 2007.
- [151] Z. Vendler. Verbs and times. *The Philosophical Review*, 66(2):143–160, 1957.
- [152] Z. Vendler. Facts and events. In Z. Vendler, Hrsg., *Linguistics in Philosophy*, S. 143–160. Cornell University Press, Ithaca, 1967.

Literaturverzeichnis

- [153] F. von Kutschera. Bewirken. *Erkenntnis*, 24(3):253–281, 1986.
- [154] G.H. von Wright. *Norm and Action. A Logical Inquiry*. Routledge and Kegan Paul, London, 1963.
- [155] R.J. Wallace. Three conceptions of rational agency. *Ethical Theory and Moral Practice*, 2(3):217–242, 1999.
- [156] R.J. Wallace. Normativity, commitment, and instrumental reason. *Philosophers' Imprint*, 1(4):1–26, 2001.
- [157] R.J. Wallace. Addiction as defect of the will. In *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason*, S. 165–189. Oxford University Press, 2006.
- [158] H. Wansing. A general possible worlds framework for reasoning about knowledge and belief. *Studia Logica*, 49(4):523–539, 1990.
- [159] H. Wansing. Tableaux for multi-agent deliberative-stit logic. In I. Hodkinson G. Governatori und Y. Venema, Hrsg., *Advances in Modal Logic. Vol. 6*, S. 503–520. College Publications, London, 2006.
- [160] R. Wedgwood. The aim of belief. *Philosophical Perspectives*, 16:267–97, 2002.
- [161] E.P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematic*, 13:1–14, 1960.
- [162] B. Williams. Deciding to believe. In *Problems of the Self*, S. 136–151. Cambridge University Press, New York, 1973.
- [163] T. Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.
- [164] B. Winters. Believing at will. *Journal of Philosophy*, 76(5):243–256, 1994.
- [165] L. Wittgenstein. Tractatus logico-philosophicus. In *Werkausgabe Band I*. Suhrkamp Verlag, Frankfurt am Main, 1984.
- [166] M. Wooldridge. *Reasoning about Rational Agents*. MIT Press, Cambridge MA, 2000.

- [167] M. Xu. Decidability of deliberative stit theories with multiple agents. In D.M. Gabbay & H.J. Ohlbach, Hrsg., *Temporal logic, First International Conference, ICTL'94*, S. 332–348. Springer Verlag, Berlin, 1994.
- [168] M. Xu. Axioms for deliberative stit. *Journal of Philosophical Logic*, 27(5):505–552, 1998.
- [169] E.N. Zalta. Twenty-five basic theorems of situation and world theory. *Journal of Philosophical Logic*, 22(4):385–428, 1993.
- [170] A. Zanardo. A finite axiomatization of the set of strongly valid ockhamist formulas. *Journal of Philosophical Logic*, 14(4):447–468, 1985.
- [171] A. Zanardo. Branching-time logic with quantification over branches: The point of view of modal logic. *Journal of Symbolic Logic*, 61(1):1–39, 1996.
- [172] A. Zanardo. Undivided and indistinguishable histories in branching time logics. *Journal of Logic, Language and Information*, 7(3):297–315, 1998.
- [173] N. Zangwill. Direction of fit and normative functionalism. *Philosophical Studies*, 91(2):173–203, 1998.

Danksagung

Mein herzlicher Dank für die anregende und motivierende Unterstützung bei der Erstellung dieser Doktorarbeit gilt vor allem meinem Doktorvater Prof. Dr. Heinrich Wansing. Er verfolgte den Fortschritt der Arbeit, baute mich auf, wenn es Rückschläge zu verkraften galt, und stand stets mit konstruktiver Kritik und Ideen, die neuen Schwung gaben, an meiner Seite.

Weiterhin danken möchte ich Dr. Uwe Scheffler, der durch hilfreiche Kommentare und wertvolle Ratschläge zum Gelingen der Arbeit beitrug. Und danken möchte ich Andrea Kruse, die mehrere Kapitel dieser Arbeit Korrektur gelesen hat. Ihrer ehrlichen Kritik, Diskussionfreude und anschließenden Motivation haben diese Arbeit und ich viel zu verdanken.

Mein Dank gilt auch den Mitarbeitern am Institut für Philosophie I und II an der Ruhr-Universität-Bochum und am Institut für Philosophie an der TU Dresden, die mich während dieser Arbeit tatkräftig unterstützten. Besonders erwähnen möchte ich dabei Judith Hecker, die mir bei allen organisatorischen Problemen half.

Eine Logik von Handlungen betrachten zu wollen, hat eine lange Tradition von Anselm von Canterbury bis in die Gegenwart. Was kennzeichnet eine Handlung? Wann beschreibt ein Satz eine Handlung? Und wie lässt sich dies in einem logischen System beschreiben? Die Stit-Theorie gibt eine einfache Antwort auf die beiden zuletzt gestellten Fragen. Ein Satz φ drückt eine Handlung aus, wenn es einen Akteur gibt, so dass sich der Wahrheitswert des Satzes nicht ändert, wenn er durch „der Akteur sorgt dafür, dass φ “ paraphrasiert wird.

In diesem Buch wird gezeigt, warum diese Ansicht nicht nur einfach und praktisch ist, um eine solche Handlungslogik um Operatoren für Überzeugungen, Wünsche und Intentionen zu erweitern, sondern auch warum sich diese Beschreibung mit der Ansicht Donald Davidsons vereinen lässt, dass Handlungen Ereignisse zugrunde liegen. So wird versucht, eine Brücke zwischen der Stit-Theorie und Davidsonscher Handlungstheorie zu schlagen, bei der der Begriff der Intention eine Rolle spielt. Dabei wird der Intentionsoperator in dem logischen System *bdi-stit_{dint}*, bestehend aus Handlungsoperatoren und den Operatoren für die mentalen Zustände, alternativ axiomatisiert.

Logos Verlag Berlin

ISBN 978-3-8325-3484-4