FernUniversität Hagen

Lehrgebiet Praktische Informatik VIII

Diplomarbeit

Konditionale Indifferenz als Basis von Default-Logiken

Andrea Kunzi

Reutlingen, den 29. Dezember 2004

Inhaltsverzeichnis

Kapitel 1

Einleitung

Auf dem Gebiet der Default-Logiken versucht man, das allgemeine menschliche Schlussfolgern bei unsicherem Wissen nachzubilden. Ein Default ist eine Regel wie z.B. „Vögel können fliegen“, zu der es Ausnahmen wie z.B. Pinguine geben kann.

Die klassische Logik geht davon aus, dass das vorhandene Wissen, das aus Fakten und Regeln besteht, immer wahr ist. In der klassischen Logik wird aus dem Wissen „Vögel können fliegen“ und dem Fakt „Tweety ist ein Vogel“ grundsätzlich geschlossen, dass Tweety fliegen kann. Diese Schlussfolgerung kann nicht mehr rückgängig gemacht werden, wenn man zusätzlich erfährt, dass Tweety ein Pinguin ist, der nicht fliegen kann. Man kann die Struktur des Wissens in der klassischen Logik daher als „statisch“ bezeichnen, während das menschliche Denken eher eine dynamische Struktur hat.

Default-Logiken sind eher geeignet, diese dynamische Natur des menschlichen Denkens zu formalisieren. Das Schließen in Default-Logiken geht davon aus, dass alle Regeln im Allgemeinen gelten, aber Ausnahmen haben können. Diese Ausnahmen müssen nicht von vornherein alle bekannt sein. Logiken aus dieser Gruppe heissen „nichtmonoton“, da das Wissen nicht automatisch durch die Hinzunahme neuer Axiome (Fakten oder Regeln) wächst. In dem oben genannten Beispiel kann man in nicht-monotonen Systemen das vorhandene Wissen um den Fakt „Tweety ist ein Pinguin“ und die Regel „Pinguine können nicht fliegen“ erweitern und kann danach den davor gezogenen Schluss „Tweety kann fliegen“ revidieren.

Der Begriff „Default“ wurde 1980 von R.Reiter [12] eingeführt. Aber auch schon vor Reiter wurde auf diesem Gebiet geforscht. Der Philosoph E.W.Adams hat 1966 konditionale Regeln unter einem wahrscheinlichkeitstheoretischen Aspekt untersucht, indem er allen Regeln eine Wahrscheinlichkeitsfunktion $1 - \varepsilon$ zuordnete, wobei ε beliebig klein sein konnte ([13],[14]). 1967 enwickelte A.Dempster die sogenannte „Dempster'sche Kombinationsregel“, die es erlaubt, verschiedene Evidenzen wie „Vögel können fliegen“, „Pinguine sind Vögel“ und „Pinguine können nicht fliegen“ mathematisch zu kombinieren [15]. Mit der Dempster'schen Kombinationsregel erhält man ein Basismaß, das der Kombination verschiedener Evidenzen zugeordnet wird. Aus diesem Basismaß kann man Belief- und Plausibilitätsfunktionen bilden. Der Ansatz von Dempster wurde 1976 von G.Shafer zur „Dempster-Shafer-Theorie“ weiterentwickelt [9].

Alle Default-Logiken, die in dieser Arbeit vorgestellt werden, basieren auf einer Präferenzrelation zwischen allen möglichen Welten. Eine Welt ist eine vollständige Zuordnung aller vorgegebenen Variablen zu einem Wahrheitswert. Wenn man z.B. von den Variablen „Vogel“ und „Pinguin“ ausgeht, wobei diese Variablen in diesem Fall die Eigenschaften eines Objekts namens „Tweety“ repräsentieren sollen, dann gibt es die vier möglichen Welten „Vogel und Pinguin“, „Vogel und kein Pinguin“, „kein Vogel und Pinguin“ und „kein Vogel und kein Pinguin“.

Das Prinzip der „konditionalen Indifferenz“ [1] wurde etwa im Jahr 1999 von G.Kern-Isberner aus wahrscheinlichkeitstheoretischen Methoden abgeleitet. Die „konditionale Indifferenz“ basiert auf der „konditionalen Struktur“ aller möglichen Welten. Zwei verschiedene Welten haben dieselbe konditionale Struktur bezüglich einer Default-Regel-Menge, wenn sie genau dieselben Regeln erfüllen und genau denselben Regeln widersprechen. Wenn man z.B. die o.g. vier Welten auf die konditionale Struktur bezüglich der Default-Regel

„Pinguine sind Vögel“ untersucht, dann haben die Welten „kein Pinguin und kein Vogel“ und „kein Pinguin und Vogel“ dieselbe konditionale Struktur, da die Regel „Pinguine sind Vögel“ auf beide Welten nicht angewandt werden kann.

Man sieht an diesem Beispiel, dass es für eine Welt bezüglich aller Defaults immer drei Möglichkeiten gibt: Eine Welt kann einen Default wie z.B. den Default „Pinguine sind Vögel“ verifizieren („Tweety ist ein Pinguin und ein Vogel“), einem Default widersprechen („Tweety ist ein Pinguin und kein Vogel“), oder der Default ist gar nicht auf die Welt anwendbar, weil die Voraussetzung der Default-Regel nicht erfüllt ist („Pinguine sind Vögel“, aber „Tweety ist gar kein Pinguin“). Diese dreiwertige Struktur der Default-Logiken macht die Handhabung dieser Logiken, auch im Vergleich zur klassischen Logik, bei der alle Aussagen entweder wahr oder falsch sind, relativ kompliziert.

Wenn man davon ausgeht, dass Default-Logiken einen wichtigen Aspekt des komplexen menschlichen Denkens repräsentieren, dann ist es einleuchtend, dass im menschlichen Denken alle Default-Regeln einen Einfluss auf die Präferenzen haben, die man allen möglichen Welten zuordnet. Zum Beispiel würde man in dem obigen Modell die Welt „Tweety ist Pinguin und Vogel“, die die Regel „Pinguine sind Vögel“ erfüllt, gegenüber der Welt „Tweety ist Pinguin und kein Vogel“, die der Regel „Pinguine sind Vögel“ widerspricht, bevorzugen.

Das Prinzip der „konditionalen Indifferenz“ bedeutet, dass alle möglichen Welten und Produkte dieser Welten, die bezüglich einer Menge von Default-Regeln dieselbe konditionale Struktur haben, auch dieselbe Wahrscheinlichkeit oder Plausibilität haben sollten. Das bedeutet z.B., dass man von zwei Welten, die dieselbe konditionale Struktur haben, nicht die eine Welt gegenüber der anderen bevorzugen sollte.

Default-Logiken erfreuen sich großer Beliebtheit in der Künstlichen Intelligenz. Sie können beispielsweise in wissensbasierten Systemen, insbesondere in Expertensystemen, eingesetzt werden.

In dieser Arbeit werden verschiedene Default-Logiken vorgestellt und unter dem Aspekt der „konditionalen Indifferenz" miteinander verglichen. Es wird dabei hauptsächlich geprüft, inwieweit konditional indifferente Repräsentationsfunktionen passende Semantiken für die verschiedenen Logiken liefern. Der Aufbau dieser Arbeit ist wie folgt. Im Kapitel 2 werden zunächst einige Begriffe und Grundkonzepte von Default-Logiken erläutert. In Kapitel 3 wird das Prinzip der konditionalen Indifferenz vogestellt. In den darauffolgenden Kapiteln 4 bis 7 werden die Default-Logiken „Penalty-Logik" [4] (von 1991), „lexikographisches System" [5] (von 1993), „Brewkas bevorzugte Subtheorien" [6] (von 1989) und „Geffners conditional entailment" [7] (von 1992) vorgestellt, und es wird untersucht, inwieweit diese Logiken dem Prinzip der konditionalen Indifferenz entsprechen. Kapitel 8 enthält eine Einführung in die Dempster-Shafer-Theorie und untersucht den Zusammenhang zwischen den Belief- und Plausibilitätsfunktionen der Dempster-Shafer-Theorie und dem Prinzip der konditionalen Indifferenz.

Die Ansätze von Adams und Dempster/Shafer inspirierten 2000 die Autoren S. Benferhat, A. Saffiotti und P. Smets zu ihrem Artikel „Belief functions and default reasoning", auf dem wesentliche Teile dieser Arbeit basieren und der in Kapitel 9 vorgestellt wird. Dieser Ansatz, der hier BSS-Ansatz [2] genannt wird, untersucht infinitesimale Belief- und Plausibilitätsfunktionen. Belief-Funktionen ordnen jeder Welt ein „Glaubensmaß" zu, und Plausibilitätsfunktionen ordnen jeder Welt ein Plausibilitätsmaß zu. Dieses Plausibilitätsmaß für einen Zustand wird umso kleiner, je größer der Glaube an das Gegenteil der Welt ist. Am Ende von Kapitel 9 wird der Zusammenhang zwischen den infinitesimalen Belief- und Plausibilitätsfunktionen des BSS-Ansatzes und dem Prinzip der konditionalen Indifferenz überprüft.

Zuletzt wird in Kapitel 10 noch die neuere Logik „LCD" aus [2] (von 2000) vorgestellt und ebenfalls auf konditionale Indifferenz überprüft, und in Kapitel 11 findet sich eine kurze Zusammenfassung dieser Arbeit.

Ich bedanke mich sehr herzlich bei Frau Prof. Kern-Isberner für die freundliche Betreuung meiner Arbeit.

Kapitel 2

Begriffe und Grundkonzepte der Default-Logik

2.1 Klassische Logik

Die Grundlage aller Logiken bildet ein *Alphabet* $\mathcal{A}$={a,b,...,}, das aus verschiedenen Symbolen besteht. Aus diesem Alphabet kann man *Formeln* bilden, indem man die Symbole durch die verschiedenen klassischen Junktoren $\vee$, $\wedge$, $\neg$ verknüpft. Im folgenden Text werden Formeln immer durch griechische Buchstaben $(\alpha, \beta, \ldots)$ oder durch Großbuchstaben $(A_1, A_2, \ldots, B_1, B_2 \ldots)$ dargestellt.

Eine logische *Sprache* $\mathcal{L}$ besteht aus allen Formeln über dem vorgegebenen Alphabet $\mathcal{A}$.

Die *Implikation* der klassischen Logik wird durch den Pfeil „$\Rightarrow$“ dargestellt. $v \Rightarrow f$ bedeutet also „Vögel können (immer, ohne Ausnahme) fliegen“. Da man aus einer wahren Aussage keine falsche Aussage folgern können sollte, ist $\alpha \Rightarrow \beta$ in der klassischen Logik genau dann wahr, wenn α falsch oder β wahr ist. $\alpha\overline{\beta}$ bedeutet $\alpha \wedge \neg\beta$.

Eine *Welt* ω ordnet jedem Symbol des Alphabets einen Wahrheitswert $\in \{\top, \bot\}$ zu. $\Omega := \{\omega_i, i = 1 \ldots 2^{|\mathcal{A}|}\}$ ist die Menge aller möglichen Welten.

Beispiel 2.1.1 (Mögliche Welten)
Wenn das Alphabet $\mathcal{A}$ aus den drei Symbolen v,f und p besteht, dann gibt es in $\mathcal{L}$ die folgenden $2^3 = 8$ möglichen Welten:

ω_1	$v \wedge p \wedge f$	ω_4	$v \wedge \neg p \wedge \neg f$	ω_7	$\neg v \wedge \neg p \wedge f$
ω_2	$v \wedge p \wedge \neg f$	ω_5	$\neg v \wedge p \wedge f$	ω_8	$\neg v \wedge \neg p \wedge \neg f$
ω_3	$v \wedge \neg p \wedge f$	ω_6	$\neg v \wedge p \wedge \neg f$		

Tabelle 1: Mögliche Welten

2.2 Konditionale und Defaults

2.2.1 Definition und Vergleich von Konditionalen und Defaults

Die *nichtmonotone Regelbeziehung* wird durch das Symbol $\longrightarrow$ dargestellt. Diese nichtmonotone Regel unterscheidet sich von den *klassischen Konditionalen* $\Rightarrow$ dadurch, dass nichtmonotone Regeln im Allgemeinen gelten und Ausnahmen beinhalten können, während die klassische Folgerungsrelation $\Rightarrow$ grundsätzlich immer ohne Ausnahme gültig ist.

Ein *Default* ist eine Regel $A \rightarrow B$, wobei A und B wieder Formeln aus der Sprache $\mathcal{L}$ über dem Alphabet $\mathcal{A}$ sein müssen. Der Default $v \rightarrow f$ kann z.B. interpretiert werden als „Vögel können (normalerweise) fliegen“.

In der klassischen Logik ist $v \Rightarrow f$ gleichbedeutend mit $\neg v \vee f$, was bedeutet, dass alle zu betrachtenden Objekte keine Vögel sind oder fliegen können (oder beides). Diese Entsprechung $A \Rightarrow B \equiv \neg A \vee B$ wird als *materiale Implikation* bezeichnet. Die materiale Implikation wird im folgenden Text immer durch eines der Symbole $\Rightarrow$ oder φ notiert.

Im Gegensatz zu der klassischen Implikation sind Defaults nicht äquivalent zu ihrer entsprechenden materialen Implikation. Die materiale Entsprechung zu dem Default $v \rightarrow f$ wäre z.B. die Regel $v \Rightarrow f = \neg v \vee f$. Da alle Ausnahmen zu der Regel $v \rightarrow f$ wie z.B. Pinguine weder keine Vögel sind noch fliegen können, gilt $v \rightarrow f \not\equiv \neg v \vee f$ im Gegensatz zu $v \Rightarrow f \equiv \neg v \vee f$.

Da die klassische Logik keine impliziten Ausnahmen zu den vorhandenen Regeln zulässt, schließt man aus den Regeln $v \Rightarrow f$ und $p \Rightarrow v$ („Vögel können fliegen“ und „Pinguine sind Vögel“), dass Pinguine fliegen können. Wenn man in der klassischen Logik ausdrücken möchte, dass Pinguine zwar Vögel sind, aber nicht fliegen können, dann muss man die Regel $v \Rightarrow f$ modifizieren, z.B. durch $v \wedge \neg p \Rightarrow f$, während die Default-Regel $v \rightarrow f$ diese Ausnahme schon implizit enthalten kann. Daher muss diese Regel bei der Erweiterung des Wissens durch die Kenntnis weiterer Ausnahmen nicht geändert werden. $v \rightarrow f$ bedeutet also: solange man von einem Tier weiss, dass es ein Vogel ist und aus dem vorhandenen Wissen nichts dagegen spricht, dass es sich um eine Ausnahme bezüglich des Fliegen-Könnens handelt, dann schließt man, dass dieses Tier fliegen kann.

2.2.2 Begriffe der Default-Logik

Die folgende Begriffsbildung orientiert sich an [1] und [2].

Eine Welt ω *verifiziert* einen Default d=$A \rightarrow B$, wenn $\omega \models A \wedge B$ gilt, wobei $\models$ die klassische Folgerungsrelation ist. Wenn $\omega \models A \wedge \neg B$ gilt, dann *falsifiziert* ω den Default d, und wenn $\omega \models \neg A$ gilt, dann kann der Default d nicht auf ω angewandt werden. Eine Welt ω *genügt* einem Default d, wenn

ω den Default d nicht falsifiziert. Δ sei im folgenden immer eine Menge von Defaults.

Für alle Formeln F wird mit $[F]$ die Menge aller Welten ω_j bezeichnet, die F verifizieren, also gilt z.B. $[\neg v \vee f] = \{\omega_j \in \Omega \mid \omega_j \models \neg v \vee f\}$. $[\neg v \vee f]$ besteht also aus der Menge aller Welten, in denen es keine Vögel gibt, die nicht fliegen können. Wenn aus dem Kontext ersichtlich ist, dass die Menge $[A]$ gemeint ist, dann werden die Klammern teilweise auch weggelassen.

Die *materiale Entsprechung* φ_d zu einem Default $d := A \rightarrow B$ wird aus der klassischen Logik übernommen: $\varphi_d := \neg A \vee B$.

Eine Menge von Defaults Δ *toleriert* einen Default d', wenn es eine Welt ω gibt, die keinen der Defaults aus Δ falsifiziert und d' verifiziert ([1]).

Eine *Stratifikation* einer Defaultbasis Δ ist eine Partitionierung von Δ in verschiedene Ebenen bzw. Prioritätsklassen, $\Delta := \Delta_1 \cup \ldots \cup \Delta_h$.

Mit $\Delta_i^{sat}(\omega)$ wird im weiteren Text die Menge aller Defaults aus Δ_i bezeichnet, denen ω genügt. Entprechend bezeichnet $\Delta_i^{fals}(\omega)$ die Menge aller Defaults aus Δ_i, die von ω falsifiziert werden.

Mit $|M|$ wird wie allgemein üblich die *Kardinalität* der Menge M bezeichnet, so dass $|\Delta_i^{fals}(\omega)|$ die Kardinalität der Menge $\Delta_i^{fals}(\omega)$ ist, also die Anzahl der Defaults in Δ_i, die ω falsifiziert.

In den Beweisen ab Kapitel 5 soll

$$\begin{aligned} & \utilde{A} + C + \ldots \\ = \quad & B + C + \ldots \end{aligned}$$

bedeuten, dass der Term A durch B ersetzt wurde.

2.3 Wahrscheinlichkeitsfunktionen

Eine *Wahrscheinlichkeitsfunktion* V_W ordnet den Formeln aus $\mathcal{L}$ einen Wahrscheinlichkeitswert aus dem reellen Intervall $[0, 1]$ zu.

Definition 2.3.1 (Wahrscheinlichkeitsfunktion)
Eine Funktion $V_W : \mathcal{L} \rightarrow [0,1]$ wird Wahrscheinlichkeitsfunktion genannt genau dann, wenn V_W die drei folgenden Bedingungen erfüllt:

1. Für alle $\alpha \in \mathcal{L} : V_W(\alpha) \geq 0$ (Positivität)

2. $V_W(\Omega) = 1$ (Normierung)

3. Für alle widersprüchlichen $\alpha, \beta \in \mathcal{L} : V_W(\alpha \vee \beta) = V_W(\alpha) + V_W(\beta)$ (endliche Additivität)

Eine Wahrscheinlichkeitsfunktion ist durch die Werte $V_W(\omega)$ auf Ω festgelegt. Für einen Default $d := \alpha \rightarrow \beta$ definiert man $V_W(\alpha \rightarrow \beta) := \frac{V_W(\alpha \wedge \beta)}{V_W(\alpha)}$ für $V_W(\alpha) > 0$.

2.4 OCFs

OCFs (ordinal conditional functions) sind Funktionen $\kappa : \Omega \rightarrow \mathbb{N} \cup \{0, \infty\}$ von der Menge Ω der möglichen Welten auf die Menge der natürlichen Zahlen, die um 0 und ∞ ergänzt wird. Sie ordnen jeder Welt ω eine positive ganze Zahl zu, die als Grad des „Nichtglaubens“ an ω interpretiert werden kann, d.h. $\kappa(\omega)$ ist sehr groß, wenn ω als sehr unplausibel eingeschätzt wird. Für aussagenlogische Formeln A, B $\in \mathcal{L}$ wird $\kappa(A)$ definiert durch $\kappa(A) := \min_{\omega \in \Omega}\{\kappa(\omega) \mid \omega \models A\}$. Für einen Default $d = A \rightarrow B$ wird $\kappa(d)$ durch $\kappa(d) := \kappa(A \wedge B) - \kappa(A)$ definiert. Man sagt, dass κ den Default $A \rightarrow B$ erfüllt, und schreibt $\kappa \models A \rightarrow B$, wenn $\kappa(A \wedge B) < \kappa(A \wedge \neg B)$ gilt, wenn also $A \wedge B$ plausibler als $A \wedge \neg B$ ist.

Eine OCF κ *repräsentiert* eine Defaultmenge Δ, wenn κ alle Defaults

$d_i \in \Delta$ erfüllt, wenn also für alle Defaults $d_i := A_i \rightarrow B_i \in \Delta$ gilt:

$$\kappa(A_i \wedge B_i) < \kappa(A_i \wedge \neg B_i)$$

In [2] wird diese Eigenschaft, dass eine Funktion eine Defaultmenge Δ repräsentiert, als *Autodeduktion* bezeichnet.

Im Allgemeinen wird eine Default-Menge bezüglich einer Logik als konsistent bezeichnet, wenn es in dieser Logik ein Modell gibt, das diese Default-Menge repräsentiert. Das bedeutet für OCFs:

Definition 2.4.1 (Konsistente Default-Menge für OCFs)

1. Eine Defaultmenge Δ ist *konsistent* genau dann, wenn es eine OCF $\kappa : \mathcal{L} \rightarrow \mathbb{N} \cup \infty$ gibt mit $\kappa(\alpha \wedge \beta) < \kappa(\alpha \wedge \neg\beta)$ für alle Defaults $d := \alpha \rightarrow \beta \in \Delta$.[1]

2. Dies ist nach ([11] Theorem 4) genau dann der Fall, wenn es in jeder nicht-leeren Teilmenge Δ' von Δ mindestens einen Default gibt, der von Δ' toleriert wird.

3. Dies ist nach ([11] Theorem 5) wiederum genau dann der Fall, wenn es eine geordnete Partition $\Delta_0, \Delta_1, \ldots \Delta_n$ von Δ gibt, so dass jeder Default in Δ_k von $\bigcup_{l=k}^{n} \Delta_l$ toleriert wird, $k \leq l \leq n$.

Begründung:
Die Äquivalenz der drei Aussagen wurde aus [11] entnommen und von J.Pearl beispielsweise in [18] bewiesen und wird hier nur kurz begründet.
$1 \Longrightarrow 3$:
Wenn es eine OCF gibt, die alle Defaults repräsentiert, für die also gilt

$$\kappa(\alpha_i \wedge \beta_i) < \kappa(\alpha_i \wedge \neg\beta_i) \text{ für alle Defaults } d_i := \alpha_i \rightarrow \beta_i \in \Delta$$

dann kann man die Defaults aufsteigend anhand ihres Ranges

$$k := \kappa(\alpha_i \wedge \beta_i) = \min_{\omega \models \alpha_i \wedge \beta_i} \kappa(\omega)$$

[1] vgl.([11] Definition 2)

in verschiedene Teilmengen Δ_k einordnen. Für alle Defaults $d \in \Delta_k$ gibt es dann eine Welt ω mit $\kappa(\omega) = k$, die d verifiziert und allen Defaults aus $\Delta_j,\ j \geq k$, genügt. Denn falls ω einen Default $d_j := \alpha_j \rightarrow \beta_j \in \Delta_j, j \geq k$, falsifizieren würde,

$$\omega \models \alpha_j \wedge \neg\beta_j,\ j \geq k$$

dann wäre $\kappa(\omega)$ wegen

$$j = \kappa(\alpha_j \wedge \beta_j) < \kappa(\alpha_j \wedge \neg\beta_j) = \min_{\omega \models \alpha_j \wedge \neg\beta_j} \kappa(\omega)$$

größer als j, und dies widerspricht der Voraussetzung $\kappa(\omega) = k \leq j$.
Die Aussage, dass es für alle Defaults $d \in \Delta_k$ eine Welt ω gibt, die d verifiziert und allen Defaults aus $\Delta_j,\ j \geq k$, genügt, bedeutet, dass d von $\bigcup_{l=k}^{n} \Delta_l$ toleriert wird.
$3 \Longrightarrow 1$:
Wenn alle Defaults $d \in \Delta_k$ von $\bigcup_{l=k}^{n} \Delta_l$ toleriert werden, dann folgt die Existenz einer OCF, die alle Defaults repräsentiert, beispielsweise aus der Existenz einer System-Z-Funktion, die in Definition 2.4.2 vorgestellt wird.

Die Äquivalenz $2 \Longleftrightarrow 3$ der Definition 2.4.1 ergibt sich aus der Prozedur Consistency-Test, die auf Seite 18 angegeben wird. Wenn es in jeder Teilmenge $\Delta' \subseteq \Delta$ einen Default gibt, der von Δ' toleriert wird, dann sortiert die Prozedur Consistency-Test die Defaults aufsteigend anhand ihrer Toleranz, und daraus ergibt sich $2 \Longrightarrow 3$. Die umgekehrte Folgerung $3 \Longrightarrow 2$ ergibt sich direkt aus der Definition.

Beispiel 2.4.1 (Konsistente Default-Basis)
Sei $\Delta = \{\Delta_1, \Delta_2\}$ eine Default-Basis mit $\Delta_1 = \{v \rightarrow f\}, \Delta_2 = \{p \rightarrow v, p \rightarrow \neg f\}$, mit $\mathcal{A} = \{p(inguin), v(ogel), f(liegt)\}$. Der Default $v \rightarrow f$ bedeutet also „Vögel können fliegen“, $p \rightarrow v$ bedeutet „Pinguine sind Vögel“, und $p \rightarrow \neg f$ bedeutet „Pinguine können nicht fliegen“. Diese Default-Basis ist konsistent, da der Default $v \rightarrow f$ in Δ_1 von Δ_2 toleriert wird und da alle Defaults aus Δ_2 von Δ_2 toleriert werden. Die Welt $\omega_1 := v \wedge \neg p \wedge f$ verifiziert den Default $v \rightarrow f$

und falsifiziert keinen der Defaults aus Δ_2. Die Welt $p \wedge v \wedge \neg f$ verifiziert beide Defaults aus Δ_2 und falsifiziert damit keinen Default aus Δ_2.

Eine mögliche OCF ist System-Z. System-Z basiert auf der „gröbstmöglichen“ konsistenten Stratifikation von Δ. Diese Stratifikation erhält man durch eine Prozedur „Consistency-Test“. [2]

Prozedur Consistency-Test

Die Prozedur Consistency-Test liefert, wie der Name schon sagt, einen Test auf Konsistenz der Default-Menge und gibt genau dann eine Stratifikation $\Delta = \Delta_1, \ldots, \Delta_n$ aus, wenn die Default-Menge konsistent ist. Die Prozedur ordnet zuerst der Teilmenge $\Delta_1 \in \Delta$ alle Defaults aus Δ zu, die von allen anderen Defaults aus Δ toleriert werden. Danach wird derselbe Test für alle Defaults aus $\Delta \backslash \Delta_1$ fortgesetzt, indem alle Defaults, die von $\Delta \backslash \Delta_1$ toleriert werden, der Menge Δ_2 zugeordnet werden. Daraufhin wird auch die Menge Δ_2 aus $\Delta \backslash \Delta_1$ entfernt, und die Testprozedur wird auf $\{\Delta \backslash \Delta_1\} \backslash \Delta_2$ angewandt usw. Der Test endet, wenn die Defaultmenge leer ist. Falls in einem Schritt der Prozedur keine Defaults gefunden werden, die von der aktuellen Defaultmenge toleriert werden, dann ist Δ inkonsistent, und die Prozedur muss abgebrochen werden. $\square$

Definition 2.4.2 (System-Z)

Sei $\Delta := \{d_1, \ldots, d_n\}$ eine Default-Menge mit der konsistenten Partitionierung $\Delta = \Delta_1, \ldots \Delta_n$, die die Prozedur „Consistency-Test“ ausgibt. Die System-Z-ranking-Funktion $\kappa^Z \colon \Omega \to \mathbb{N}$ ist definiert durch

$$\kappa^Z(\omega) = \begin{cases} 0, \text{ falls } \omega \text{ kein } d_i \text{ falsifiziert} \\ 1 + \max\limits_{\substack{1 \leq i \leq n \\ \omega \models A_i \overline{B_i}}} Z(d_i) \text{ sonst} \end{cases}$$

mit $Z(d_i) = j$ genau dann, wenn $d_i \in \Delta_j$. [3]

[2] vgl.[11] S.65

κ^Z ordnet jeder Welt ω den kleinstmöglichen Rang zu, der in bezug auf die Bedingungen in Δ erreichbar ist.

Beispiel 2.4.2 (System-Z)
Wenn man System-Z auf die Default-Basis des obigen Beispiels 2.4.1 für eine konsistente Default-Basis anwendet, dann erhält man

$$Z(v \rightarrow f) = 1 \text{ und } Z(p \rightarrow v) = Z(p \rightarrow \neg f) = 2$$

wegen

$$(v \rightarrow f) \in \Delta_1, (p \rightarrow v) \in \Delta_2 \text{ und } (p \rightarrow \neg f) \in \Delta_2$$

Deshalb ergibt sich für die möglichen Welten:

ω	$\kappa^Z(\omega)$	ω	$\kappa^Z(\omega)$
$v \wedge p \wedge f$	3	$\neg v \wedge p \wedge f$	3
$v \wedge p \wedge \neg f$	2	$\neg v \wedge p \wedge \neg f$	3
$v \wedge \neg p \wedge f$	0	$\neg v \wedge \neg p \wedge f$	0
$v \wedge \neg p \wedge \neg f$	2	$\neg v \wedge \neg p \wedge \neg f$	0

2.5 Bevorzugte Welten

Das Konzept der „bevorzugten Welten“ kann man dazu verwenden, verschiedene Logiken miteinander zu vergleichen. Viele nichtmonotone Logiken basieren auf einer Präferenzrelation R auf Ω, durch die allen möglichen Welten

[3] [11] Def.12

$\omega \in \Omega$ ein Rang in einem Ordnungssystem zugewiesen wird. Zum Beispiel sollte eine solche Präferenzrelation bei einer gegebenen Defaultmenge eine Welt ω, die alle Defaults verifiziert, gegenüber einer Welt ω', die alle Defaults falsifiziert, bevorzugen. Diese Tatsache, dass ω bezüglich der Präferenzrelation R eine „bevorzugte Welt“ gegenüber ω' ist, kann man duch $\omega <_R \omega'$ notieren.

2.5.1 Bevorzugte Welten und beliebige Präferenzrelationen

Alle in den folgenden Kapiteln 4 bis 7 vorgestellten Logiken werden durch eine Präferenzrelation R auf Ω definiert. Aufgrund dieser Präferenzrelation R erhält man die bevorzugten Welten für Formeln $\alpha \in \mathcal{L}$:

Definition 2.5.1 (Bevorzugte Welten für eine Formel α bzgl. $<_R$)
Sei $<_R$ eine beliebige Relation auf Ω und α eine Formel aus $\mathcal{L}$. Eine Welt $\omega \in \Omega$ wird als

$$\alpha_R\text{-bevorzugt}$$

bezeichnet genau dann, wenn $\omega \models \alpha$ gilt und es keine Welt $\omega' \in \Omega$ gibt mit $\omega' \models \alpha$ und $\omega' <_R \omega$.

Die Folgerungsrelation $\mid\!\sim_R$ bezüglich der Relation $<_R$ auf Ω wird dann durch die folgende Definition festgelegt:

Definition 2.5.2 (Inferenzrelation $\mid\!\sim_R$)
Es gilt

$$\alpha \mid\!\sim_R \beta$$

genau dann, wenn für alle ω, die α_R-bevorzugt sind, gilt $\omega \models \beta$.

Solche Folgerungsrelationen $\mid\!\sim_R$, die auch Inferenzrelationen genannt werden,

werden durch die Relationen $<_R$ eindeutig bestimmt.

2.5.2 Bevorzugte Welten und OCFs

Man kann auch jede OCF κ durch ein sogenanntes Präferenzmodell beschreiben.

Definition 2.5.3 (Bevorzugte Welten für OCFs)
Sei κ eine OCF auf Ω, und seien ω_1 und ω_2 zwei Welten aus Ω. ω_1 wird als $\kappa - bevorzugt$ gegenüber ω_2 bezeichnet, geschrieben

$$\omega_1 <_\kappa \omega_2$$

genau dann, wenn $\kappa(\omega_1) < \kappa(\omega_2)$.

Die strikte partielle Ordnung $<_\kappa$ wird κ-Präferenz-Relation genannt. Die Ordnung $<_\kappa$ kann dafür verwendet werden, um die Menge der bezüglich κ bevorzugten Welten für eine gegebene Formel zu definieren.

Definition 2.5.4 (Bevorzugte Welten für Formeln bzgl. OCFs)
Sei κ eine OCF auf Ω und A_i eine Formel aus $\mathcal{L}$. Eine Welt ω aus Ω wird als κ-bevorzugte Welt für A_i bezeichnet, falls

1. ω A_i erfüllt, d.h. $\omega \models A_i$, und

2. es keine andere Welt ω' gibt, die A_i erfüllt, so dass $\kappa(\omega') < \kappa(\omega)$.

Laut Definition gilt $\kappa \models A \to B$ genau dann, wenn $\kappa(AB) < \kappa(A\overline{B})$, und $\kappa(AB) = \min\{\kappa(\omega) \mid \omega \models AB\}$. Sei κ eine OCF auf Ω.

Satz 2.5.1 (Bevorzugte Welten und OCFs)
Für alle Formeln A, B aus $\mathcal{L}$ und für alle OCFs κ gilt $\kappa \models A \to B$ genau dann, wenn jede κ-bevorzugte Welt für A auch B erfüllt.

Mit diesem Satz kann man zeigen, dass eine Logik, die durch eine beliebige Präferenzrelation auf Ω definiert wurde, dieselbe Folgerungsrelation induziert wie eine Logik, die durch eine OCF definiert wurde, indem man beweist, dass die bevorzugten Welten in beiden Logiken gleich sind.

BEWEIS: $\Longrightarrow$
Angenommen, es gilt

$$\kappa \models A \to B, \text{ d.h. } \min_{\omega \models A \wedge B} \kappa(\omega) < \min_{\omega \models A \wedge \neg B} \kappa(\omega) \tag{2.1}$$

und

$$\omega_1 \text{ sei eine } A\text{-bevorzugte Welt, die nicht } B \text{ erfüllt,} \tag{2.2}$$

$$\text{also } \omega_1 \models A \wedge \neg B$$

Wegen $\min_{\omega \models A \wedge B} \kappa(\omega) < \min_{\omega \models A \wedge \neg B}$ (2.1) gibt es dann eine Welt ω_2, $\omega_2 \models A \wedge B$, mit $\kappa(\omega_2) < \kappa(\omega_1)$. Wegen $\omega_2 \models A$ und $\kappa(\omega_1) > \kappa(\omega_2)$ kann dann ω_1 nicht A-bevorzugt sein: $\kappa(\omega_1) \neq \min_{\omega \models A} \kappa(\omega)$. Widerspruch zu (2.2)
$\Longleftarrow$

$$\text{Angenommen, jede } A\text{-bevorzugte Welt erfüllt } B \tag{2.3}$$

und

$$\kappa \not\models A \to B, \text{ d.h. } \min_{\omega \models A \wedge B} \kappa(\omega) \not< \min_{\omega \models A \wedge \neg B} \kappa(\omega) \tag{2.4}$$

Sei ω_1 eine A-bevorzugte Welt, also $\kappa(\omega_1) = \min_{\omega \models A} \kappa(\omega)$. ω_1 muss dann wegen (2.3) auch B erfüllen, also $\omega_1 \models A \wedge B$. Jede Welt ω muss entweder B oder $\neg B$ erfüllen, da alle ω jedem Symbol aus $\mathcal{L}$ einen Wahrheitswert zuordnen. Wenn $\min_{\omega \models A \wedge \neg B} \kappa(\omega) \leq \min_{\omega \models A \wedge B} \kappa(\omega)$ (2.4) gilt, dann gibt es eine Welt ω_2 mit $\omega_2 \models A \wedge \neg B$ und $\kappa(\omega_2) \leq \kappa(\omega_1)$. Wegen $\kappa(\omega_2) \leq \kappa(\omega_1) = \min_{\omega \models A} \kappa(\omega)$ und $\omega_2 \models A$ muss auch ω_2 A-bevorzugt sein. Da aber ω_2 nicht B erfüllt, ergibt sich ein Widerspruch zu (2.3). □

Kapitel 3

Konditionale Indifferenz

In [1] werden OCFs κ und Wahrscheinlichkeitsfunktionen V_W auf konditionale Indifferenz untersucht.

Sei Δ eine Menge von Defaults $d_i = A_i \to B_i$. Die konditionale Struktur einer Welt ω wird aus dem Verhältnis von ω zu den n einzelnen Defaults $d_i = A_i \to B_i,\ 1 \leq i \leq n$, einer Logik abgeleitet.[1] Man assoziiert zuerst mit jedem Default d_i ein Symbol a_i^+, das bedeutet, dass der Default verifiziert wird, und ein Symbol a_i^-, das bedeutet, dass der Default falsifiziert wird, und erzeugt aus diesen Symbolen eine freie abelsche Gruppe $F_\Delta =< a_1^+, a_1^-, ..., a_n^+, a_n^- >$. Eine Gruppe muss assoziativ sein, ein neutrales Element beinhalten, und zu jedem Element muss es ein Inverses geben. In einer abelschen Gruppe muss zusätzlich das Kommunikativgesetz gelten. Das neutrale Element in der Gruppe F_Δ ist die 1, und das Inverse zu jedem Element a_i^+ ist dann $\frac{1}{a_i^+}$. Entsprechend ist $\frac{1}{a_i^-}$ das Inverse zu a_i^- wegen $a_i^- \cdot \frac{1}{a_i^-} = 1$.

Danach definiert man eine Funktion $\sigma_i = \sigma_{(A_i \to B_i)} : \Omega \to F_\Delta$, indem man

[1] vgl.[1]Abschnitt 5

festlegt

$$\sigma_i(\omega) := \begin{cases} a_i^+ \text{ wenn } \omega \models A_i \wedge B_i \\ a_i^- \text{ wenn } \omega \models A_i \wedge \neg B_i \\ 1 \text{ wenn } \omega \models \neg A_i \end{cases} \quad \text{für alle Defaults } A_i \rightarrow B_i \in \Delta$$

Diese σ_i ermöglichen die Definition einer Funktion σ_Δ:

Definition 3.0.5 (Konditionale Struktur)
[2]Die *konditionale Struktur* σ_Δ einer Welt ω ist gegeben durch

$$\sigma_\Delta(\omega) := \prod_{1 \leqslant i \leqslant n} \sigma_i(\omega) = \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \wedge B_i}} a_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \wedge \neg B_i}} a_i^- \tag{3.1}$$

Da das leere Produkt $\prod_{i=0}^{0}$ gleich 1 ist, gilt $\sigma_\Delta(\omega) = 1$, falls ω alle Defaults weder verifiziert noch falsifiziert, das bedeutet, falls $\omega \models \neg A_i$ für alle Defaults $A_i \rightarrow B_i$. σ_Δ beschreibt also den simultanen Effekt aller Defaults d_i auf ω.

Die konditionale Struktur σ_Δ ist ein nützliches Instrument, um den Einfluss der Defaults auf die Rangordnung, die man einer beliebigen Welt ω zuordnet, zu veranschaulichen. Es ist intuitiv einleuchtend, dass die Verifikation eines Defaults durch ω einen positiven und die Falsifizierung eines Defaults einen negativen Einfluss auf die Rangordnung von ω bezüglich aller Präferenzrelationen auf Ω und der von ihnen abgeleiteteten Folgerungsrelationen haben sollte. Genauso sollten alle Defaults $A \rightarrow B$, die von ω weder verifiziert noch falsifiziert werden, weil ω die Defaultvoraussetzung A nicht erfüllt, möglichst keinen Einfluss auf die Rangordnung von ω haben.

Beispiel 3.0.1 (Konditionale Struktur)
Sei $\Delta := \{d_1 = v \rightarrow f,\ d_2 = p \rightarrow v,\ d_3 = p \rightarrow \neg f\}$ *wie in Beispiel 2.4.2 eine Menge von Defaults auf dem Alphabet* $\mathcal{A} = \{v, f, p\}$. *Die Welt* $\omega_1 = v \wedge p \wedge f$

[2]vgl.[1] Gleichung(8)

verifiziert den ersten Default $d_1 = v \rightarrow f$. *Also ergibt sich* $\sigma_1(\omega_1) = a_1^+$. *Ebenso verifiziert* ω_1 *auch den zweiten Default* $d_2 = p \rightarrow v$, *so dass man* $\sigma_2(\omega_1) = a_2^+$ *erhält. Der dritte Default* d_3 *wird jedoch von* ω_1 *falsifiziert:* $p \wedge v \wedge f \models p \wedge f$, *und deshalb ergibt sich* $\sigma_3(\omega_1) = a_3^-$. *Die konditionale Struktur* $\sigma_\Delta(\omega_1)$ *ist also* $\sigma_\Delta(\omega_1) = \prod_{i=1}^{3} \sigma_i(\omega_1) = a_1^+ a_2^+ a_3^-$. *Die folgende Tabelle listet die konditionalen Strukturen aller möglichen Welten auf* $\mathcal{A}$ *auf:*

ω	$\sigma_\Delta(\omega)$
$v \wedge p \wedge f$	$a_1^+ a_2^+ a_3^-$
$v \wedge p \wedge \neg f$	$a_1^- a_2^+ a_3^+$
$v \wedge \neg p \wedge f$	a_1^+
$v \wedge \neg p \wedge \neg f$	a_1^-

ω	$\sigma_\Delta(\omega)$
$\neg v \wedge p \wedge f$	$a_2^- a_3^-$
$\neg v \wedge p \wedge \neg f$	$a_2^- a_3^+$
$\neg v \wedge \neg p \wedge f$	*1*
$\neg v \wedge \neg p \wedge \neg f$	*1*

Diese Funktion σ_Δ wird jetzt noch erweitert auf die Juxtaposition bzw. Multiplikation von Welten:

Definition 3.0.6 (Produkt $\hat{\omega}$)
Das Symbol $\hat{\omega}$ bezeichnet ein Produkt von Welten

$$\hat{\omega} := \omega_1^{r_1} \cdots \omega_n^{r_n}$$

mit $r_i \in \mathbb{Z} \; \forall \; i = 1, \ldots, n$

Diese Multiplikation von Welten ist lediglich als Juxtaposition, also „Nebeneinanderstellung“, zu interpretieren. Die Exponenten können wegen $r_i \in \mathbb{Z}$ auch negativ sein.

Es ist klar, das man mit dieser Definition eine freie abelsche Gruppe $\hat{\Omega}$ erhält, wobei $\hat{\Omega}$ aus allen Elementen $\hat{\omega}$ besteht, die sich aufgrund von Definition 3.0.6 bilden lassen. Die Multiplikation ist assoziativ und kommunikativ, das neutrale Element ist die 1, und zu jedem Element ${\omega_i}^{r_i}$ ist ${\omega_i}^{-r_i} = \frac{1}{{\omega_i}^{r_i}}$ das Inverse.

Daraufhin wird die konditionale Struktur eines Produkts $\hat{\omega}$ von Welten definiert:

Definition 3.0.7 (Konditionale Struktur von $\hat{\omega}$)

$$\sigma_\Delta(\omega_1^{r_1}\cdots\omega_n^{r_n}) := \sigma_\Delta(\omega_1)^{r_1}\cdots\sigma_\Delta(\omega_n)^{r_n} \tag{3.2}$$

Eine Wahrscheinlichkeitsfunktion V_W oder OCF κ wird entsprechend auf einen Homomorphismus auf $\hat{\Omega} = \{\hat{(\omega)}\}$ fortgesetzt. Da die Division durch 0 nicht definiert ist, schließt man alle Welten, die die Wahrscheinlichkeit 0 haben, von der Definitionsmenge aus. $\Omega^\times := \{\omega \in \Omega \mid V_W(\omega) \neq 0\}$.

Definition 3.0.8 (Fortsetzung von V_W auf $\hat{\Omega}$)
Eine Wahrscheinlichkeitsfunktion $V_W : \Omega \to [0,1]$ bzw. eine OCF $\kappa : \Omega \to \mathbb{N} \cup \infty$ wird zu einem Homomorphismus $V_W : \hat{\Omega}^\times \to \mathbb{R}^+$ bzw. $\kappa : \hat{\Omega} \to (\mathbb{Z}, +)$ fortgesetzt durch

$$V_W(\hat{\omega}) := V_W(\omega_1^{r_1}\ldots\omega_k^{r_k}) := V_W(\omega_1)^{r_1}\ldots V_W(\omega_k)^{r_k}$$

bzw.

$$\kappa(\hat{\omega}) = \kappa(\omega_1^{r_1}\cdots\omega_r^{r_n}) := r_1\kappa(\omega_1) + \ldots + r_n\kappa(\omega_n)$$

Für die Definition der konditionalen Indifferenz benötigt man noch die Relation $\equiv_\top$:

Definition 3.0.9 ($\equiv_\top$)
Seien $\hat{\omega}_1 = \omega_1^{r_1} \dots \omega_m^{r_m}$, $\hat{\omega}_2 = \nu_1^{s_1} \dots \nu_p^{s_p} \in \hat{\Omega}$ gegeben. Es gilt

$$\hat{\omega}_1 \equiv_\top \hat{\omega}_2$$

genau dann, wenn

$$\sum_{1 \le j \le m} r_j = \sum_{1 \le k \le p} s_k$$

$\hat{\omega}_1 \equiv_\top \hat{\omega}_2$ gilt also genau dann, wenn die Summe der Exponenten der Generatoren ω_i von $\hat{\omega}_1$ und $\hat{\omega}_2$ gleich ist, wobei die Exponenten $r_j,\ s_k \in \mathbb{Z}$, wie schon erwähnt wurde, auch negativ sein können. $\hat{\omega}_1 \equiv_\top \hat{\omega}_2$ gilt z.B. dann, wenn $\hat{\omega}_1$ und $\hat{\omega}_2$ aus derselben Anzahl von Welten bestehen, die multipliziert, also nebeneinandergestellt bzw. juxtapositioniert, wurden.

Die wichtigste Definition in diesem Kapitel ist nun die Definition der konditionalen Indifferenz:

Definition 3.0.10 (Konditionale Indifferenz)
[3]Sei

$$\Delta = \{(A_i \to B_i), 1 \le i \le n\} \subseteq (\mathcal{L} \to \mathcal{L})$$

eine Menge von Defaults. Eine Wahrscheinlichkeitsfunktion V bzw. eine OCF V ist indifferent in bezug auf Δ genau dann, wenn $V(\hat{\omega}_1) = V(\hat{\omega}_2)$ immer dann, wenn $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$ für $\hat{\omega}_1 \equiv_\top \hat{\omega}_2$ gilt.

Beispiel 3.0.2 (Konditionale Indifferenz von System-Z-Funktionen)
System-Z-Funktionen sind in der Regel nicht konditional indifferent. In dem obigen Beispiel 2.4.2 mit $\Delta = \{d_1 = v \to f, d_2 = p \to v, d_3 = p \to \neg f\}$ *haben die Welten* $\hat{\omega}_1 := (\neg v \wedge p \wedge f) \cdot (v \wedge p \wedge \neg f) \cdot (v \wedge \neg p \wedge f)$ *und*

[3] vgl.[1] Def.1

$\hat{\omega}_2 := (\neg v \wedge p \wedge \neg f) \cdot (v \wedge p \wedge f) \cdot (v \wedge \neg p \wedge \neg f)$ *dieselbe konditionale Struktur* $\sigma_\Delta(\hat{\omega}_1) = (a_2^- a_3^-) \cdot (a_1^- a_2^+ a_3^+) \cdot (a_1^+) = \sigma_\Delta(\hat{\omega}_2) = (a_2^- a_3^+) \cdot (a_1^+ a_2^+ a_3^-) \cdot (a_1^-)$. *Da* $\hat{\omega}_1$ *und* $\hat{\omega}_2$ *jeweils das Produkt von drei Welten sind, gilt* $\hat{\omega}_1 \equiv_\top \hat{\omega}_2$. *Wegen* $\kappa^Z(\hat{\omega}_1) = 3 + 2 + 0 = 5 \neq \kappa^Z(\hat{\omega}_2) = 3 + 3 + 2 = 8$ *ist diese System-Z-Ranking-Funktion* κ^Z *jedoch nicht konditional indifferent. (*$\kappa^Z : \Omega \to \mathbb{N}$ *wurde entsprechend der Definition 3.0.8 auf* $\hat{\Omega}$ *fortgesetzt.)*

Die konditionale Indifferenz von OCFs kann man laut [1] Theorem 1 (9) auch anhand folgender Bedingung überprüfen:

Definition 3.0.11 (Konditionale Indifferenz von OCFs)
Eine (endliche) OCF ist indifferent in bezug auf eine Menge von Defaults $\Delta = \{(A_1 \to B_n), ..., (A_n \to B_n)\}$ genau dann, wenn es rationale Zahlen $\kappa_0, \kappa_i^+, \kappa_i^- \in \mathbb{Q}, 1 \leq i \leq n$ gibt, so dass für alle $\omega \in \Omega$ gilt:

$$\kappa(\omega) = \kappa_0 + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \kappa_i^+ + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_i^- \tag{3.3}$$

Beispiel 3.0.3 (System-Z und Definition 3.0.11)
Man kann auch anhand von Definition 3.0.11 nachweisen, dass die System-Z-Funktion κ^Z *aus Beispiel 2.4.2 nicht konditional indifferent ist. Die folgende Tabelle listet für alle* $\omega \in \Omega$ *jeweils den Rang* $\kappa^Z(\omega)$ *und die konditionale Struktur* $\sigma_\Delta(\omega)$ *auf:*

ω_i		$\kappa^Z(\omega)$	$\sigma_\Delta(\omega)$
ω_1	$v \wedge p \wedge f$	3	$a_1^+ a_2^+ a_3^-$
ω_2	$v \wedge p \wedge \neg f$	2	$a_1^- a_2^+ a_3^+$
ω_3	$v \wedge \neg p \wedge f$	0	a_1^+
ω_4	$v \wedge \neg p \wedge \neg f$	2	a_1^-
ω_5	$\neg v \wedge p \wedge f$	3	$a_2^- a_3^-$
ω_6	$\neg v \wedge p \wedge \neg f$	3	$a_2^- a_3^+$
ω_7	$\neg v \wedge \neg p \wedge f$	0	1
ω_8	$\neg v \wedge \neg p \wedge \neg f$	0	1

κ_0 in Gleichung (3.3) muss wegen $\kappa^Z(\omega_8) = 0$ gleich 0 sein. Wegen $\kappa^Z(\omega_4) = 2$, und da ω_4 nur den ersten Default falsifiziert, muss κ_1^- in derselben Gleichung gleich 2 sein. Wegen $\kappa^Z(\omega_3) = 0$ muss κ_1^+ gleich 0 sein. Wegen $\kappa^Z(\omega_5) = \kappa^Z(\omega_6)$ und da sich die konditionale Struktur dieser Welten nur durch a_3^- bzw. a_3^+ unterscheidet, muss gelten $\kappa_3^- = \kappa_3^+$. κ_3^+ und κ_3^- kann man also mit κ_3 bezeichnen. Man findet unter diesen Voraussetzungen keinen Wert κ_2^+ mehr, so dass die Bedingungungen $\kappa^Z(\omega_1) = 3 = \kappa_1^+ + \kappa_2^+ + \kappa_3^{(-)}$ und $\kappa^Z(\omega_2) = 2 = \kappa_1^- + \kappa_2^+ + \kappa_3^{(+)}$ erfüllt sind, denn es ergibt sich $0 + \kappa_2^+ + \kappa_3 = 3$ und $2 + \kappa_2^+ + \kappa_3 = 2$, also $\kappa_2^+ + \kappa_3 = 0$.

Das Problem von System-Z bezüglich der konditionalen Indifferenz besteht darin, dass die Definition von System-Z-Funktionen über das Maximum nicht mit Homomorphismen und Gruppenverknüpfungen vereinbar ist. Wenn man eine konkrete OCF κ oder eine Wahrscheinlichkeitsfunktion V_W gegeben hat, die die Defaults aus Δ repräsentiert, dann wird diese Funktion κ bzw. V_W in [1] in Definition 2 als *c-Repräsentation* bezeichnet, falls sie in bezug auf Δ konditional indifferent ist:

Definition 3.0.12 (c-Repräsentation)
OCFs κ und Wahrscheinlichkeitsfunktionen V_W, die eine Menge Δ von Defaults repräsentieren und die in bezug auf Δ konditional indifferent sind, werden *c-Repräsentationen* genannt.

c-Repräsentationen sind ein mächtiges Mittel, um Logiken zu repräsentieren.

In dieser Arbeit werden nur c-Repräsentationen bzw. OCFs betrachtet, für die in Gleichung (3.3) $\kappa_0 = 0$, $\kappa_i^+ = 0$ und $\kappa_i^- \geq 0$ gilt.

Kapitel 4

Penalty-Logik

4.1 Grundlagen der Penalty-Logik

Sei Δ eine Menge von Defaults. Die Penalty-Logik [4] ordnet jedem Default $d_i \in \Delta$, $d_i = A_i \rightarrow B_i$, eine Zahl $\rho_i \in \mathbb{R}^+$ zu, die als „Strafe“ interpretiert wird, die man bezahlen muss, wenn der Default falsifiziert wird. Man geht von einer vorgegebenen Stratifikation $\Delta = \{\Delta_1, \ldots, \Delta_n\}$ der Default-Basis aus. Die Penalties ρ_i für jeden Default $d_i \in \Delta$ wurden vom Anwender der Logik zugewiesen. Alle Defaults d_i, die als wichtig erachtet wurden, erhalten größere Penalties ρ_i als andere Defaults, die als weniger wichtig eingestuft werden.

Der Rang *Vrank* einer Welt ω wird dann berechnet aus der Summe aller ρ_i derjenigen Defaults d_i bzw. ihrer materialen Entsprechung φ_i, die von ω falsifiziert werden:

Definition 4.1.1 (Penalty-Funktion Vrank)
Für alle Welten $\omega \in \Omega$ wird die Penalty-Funktion Vrank definiert durch

$$\text{Vrank} : \Omega \to \mathbb{R}^+, \ \text{Vrank}(\omega) = \sum_{\omega \not\models \varphi_i} \rho_i$$

Eine Welt ω wird als „Penalty-bevorzugt“ gegenüber einer Welt ω' bezeichnet,

$$\omega <_{pen} \omega'$$

falls $\text{Vrank}(\omega) < \text{Vrank}(\omega')$. Eine Interpretation ω wird analog zu Definition 2.5.1 für eine Formel α als „α_{pen}-bevorzugt“ bezeichnet, falls α in ω wahr ist und es keine Welt ω' gibt, die α erfüllt, so dass $\omega' <_{pen} \omega$. Wie in Definition 2.5.2 basiert die Folgerungsrelation $\mid\!\sim_{pen}$ auf der Relation Vrank auf Ω. Eine Formel β ist eine Penalty-Folgerung aus α, notiert durch

$$\alpha \mid\!\sim_{pen} \beta$$

falls jede α_{pen}-bevorzugte Welt auch β erfüllt.

Beispiel 4.1.1 (Penalty-Logik)
Sei Δ wieder die Menge der Defaults $d_1 = v \to f, d_2 = p \to v, d_3 = p \to \neg f$ mit der oben erwähnten Bedeutung „Vögel können fliegen“, „Pinguine sind Vögel“ und „Pinguine können nicht fliegen“. Man kann diesen Defaults nun z.B. die „Penalties“ $\rho_1 = 1, \rho_2 = 2$ und $\rho_3 = 3$ zuordnen. Wenn man jetzt untersuchen möchte, ob Pinguine, die Vögel sind, fliegen können, also

$$p \wedge v \to f?$$

dann muss man zunächst alle „$p \wedge v$“-bevorzugten Welten suchen. Die Welt $\omega_1 := p \wedge v \wedge f$ hat den Rang $\text{Vrank}(\omega_1) = \rho_3 = 3$, da ω_1 den dritten Default falsifiziert. Die Welt $\omega_2 := p \wedge v \wedge \neg f$ hat den Rang $\text{Vrank}(\omega_2) = \rho_1 = 1$, da diese Welt den ersten Default falsifiziert. Wegen $\text{Vrank}(\omega_2) < \text{Vrank}(\omega_1)$ ist also ω_2 eine (die einzige) $p \wedge v$-bevorzugte Interpretation. Da diese Interpretation $\neg f$ erfüllt, folgert man, dass Pinguine, die Vögel sind, nicht fliegen können.

Wenn die Stratifikation, auf der die Wert-Zuweisung der ρ_i basiert, konsistent im Sinne von Definition 2.4.1 ist, dann gibt es auch immer eine Funktion Vrank, die die Default-Menge repräsentiert. Denn wenn die Default-Menge konsistent ist, dann gibt es laut Definition 2.4.1 für jeden gegebenen Default $\alpha_i \rightarrow \beta_i \in \Delta_i$ eine Welt ω mit $\omega \models \alpha_i \wedge \beta_i$, die alle Default aus $\bigcup_{j=i}^{n} \Delta_j$ toleriert. D.h. die „Strafe" Vrank(ω), die dieser Welt ω zugeordnet wird, kann höchstens so groß wie die Summe der Strafen aller Defaults aus tieferen Ebenen Δ_k, $k < i$ sein, da ω keinen Default aus Δ_j, $j > i$, falsifiziert. Die „Strafe" Vrank, die einer Welt ω', die $\alpha_i \rightarrow \beta_i$ falsifiziert, zugeordnet wird, muss dagegen laut Definition 4.1.1 mindestens so groß wie ρ_i sein. Wenn man dann allen ρ_i Werte zuordnet, so dass für alle $1 \leq i \leq n$ gilt, dass $\rho_i > \sum_{k=1}^{i-1} \rho_k$ für alle $1 \leq i \leq n$, dann können alle ω', die $\alpha_i \rightarrow \beta_i$ falsifizieren, wegen $\mathrm{Vrank}(\omega)_{\omega \models \alpha_i \wedge \beta_i} < \mathrm{Vrank}(\omega')_{\omega \models \alpha_i \wedge \neg \beta_i}$ nicht α_i-pen-bevorzugt sein. Deshalb sind die α_i-pen-bevorzugten Welten eine Teilmenge der $\alpha_i \wedge \beta_i$-pen-bevorzugten Welten, und daraus folgt, dass jede α_i-pen-bevorzugte Welt auch β_i erfüllt für alle Defaults $\alpha_i \rightarrow \beta_i \in \Delta$.

Wenn man dagegen von einer inkonsistenten Stratifikation ausgeht, dann kann es Probleme geben, wie das folgende Beispiel zeigt.

Beispiel 4.1.2 (Penalty-Logik bei inkonsistenter Stratifikation)

Wenn man in dem vorherigen Beispiel den Defaults andere Penalties ρ_i zuweist, z.B. $\rho_1 = 3, \rho_2 = 2$ und $\rho_3 = 1$, dann hätte die Welt $\omega_1 = p \wedge v \wedge f$ den Rang $\mathrm{Vrank}(\omega_1) = 1$ *und die Welt $\omega_2 = p \wedge v \wedge \neg f$ den Rang* $\mathrm{Vrank}(\omega_2) = 3$. *Damit wäre ω_1 eine $p \wedge v$-bevorzugte Welt, und da Pinguine, die Vögel sind, in dieser Welt fliegen können, würde man folgern*

$$p \wedge v \rightarrow f$$

Diese Wertzuweisungen wären natürlich unsinnig, da Pinguine, die Vögel sind, in der Regel nicht fliegen können. Die Funktion Vrank *mit diesen Wertzuweisungen repräsentiert die gegebene Default-Menge nicht mehr, da sie dem Prinzip der Autodeduktion, das bedeutet, dass alle Defaults ableitbar sein müssen, nicht* *entspricht, denn der dritte Default ist nicht mehr ableitbar: Um abzuleiten, dass Pinguine nicht fliegen können, also $p \to \neg f$, muss man zuerst alle Pinguin-bevorzugten Welten finden. Die Werte von $\mathrm{Vrank}(\omega_1)$ und $\mathrm{Vrank}(\omega_2)$ wurden oben schon angegeben. Die zwei restlichen Welten, die „p“ verifizieren, sind $\omega_3 := p \wedge \neg v \wedge f$ und $\omega_4 := p \wedge \neg v \wedge \neg f$. Es gilt $\mathrm{Vrank}(\omega_3) = 2 + 1 = 3$ und $\mathrm{Vrank}(\omega_4) = 2$. Damit ist ω_1 wegen $\mathrm{Vrank}(\omega_1) = \min_{\omega \models p} \mathrm{Vrank}$ die einzige p(inguin)-bevorzugte Welt, und daher kann man in diesem Beispiel nur $p \to f$ ableiten, was dem dritten Default widerspricht.*

Die Inkonsistenz dieses Beispiels ergibt sich aus der Tatsache, dass der Default $d_3 = p \to \neg f$, der in diesem Beispiel die niedrigste Priorität $\rho_3 = 1$ hat, von der Defaultmenge $\{v \to f, p \to v\}$, die höhere Priorität hat, nicht toleriert wird. Die beiden Welten, die $p \to \neg f$ verifizieren, sind $\omega_1 := v \wedge p \wedge \neg f$ und $\omega_2 := \neg v \wedge p \wedge \neg f$. ω_1 falsifiziert jedoch den Default $v \to f$, und ω_2 falsifiziert den Default $p \to v$.

Da inkonsistente Default-Mengen widersprüchliche Aussagen beinhalten, kann man aus solchen Default-Mengen unplausible Schlussfolgerungen ziehen.

4.2 Konditionale Indifferenz von Penalty-Funktionen

In diesem Abschnitt soll gezeigt werden, dass alle Penalty-Funktionen konditional indifferent sind.

Satz 4.2.1 (Penalty-Logik und konditionale Indifferenz)
Sei Δ eine Default-Basis. Alle Penalty-Funktionen Vrank auf Δ sind konditional indifferent bezüglich Δ.

BEWEIS: Der Beweis für Penalty-Funktionen mit $\rho_i \in \mathbb{N}$ für alle ρ_i ergibt sich ganz einfach: Falls für alle ρ_i gilt: $\rho_i \in \mathbb{N}$, dann sind diese Penalty-Funktionen $\text{Vrank}_{\mathbb{N}}$ spezielle OCFs, die die Bedingungen aus Definition 3.0.11 auf S.28 erfüllen: $\kappa_{pen\mathbb{N}}(\omega) = \kappa_0 + \sum_{\omega \models A_i B_i} \kappa_i^+ + \sum_{\omega \models A_i \overline{B}_i} \kappa_i^-$ mit $\kappa_0 := 0$, $\kappa_i^+ := 0$ und $\kappa_i^- := \rho_i$, $\kappa_{pen\mathbb{N}}(\omega) = \text{Vrank}_{\mathbb{N}}(\omega)$ für alle $\omega \in \Omega$, und damit sind diese Funktionen konditional indifferent. Wegen $\kappa_{pen\mathbb{N}}(\omega) = \text{Vrank}_{\mathbb{N}}(\omega)$ gilt in diesem Fall $\kappa_{pen\mathbb{N}}(\omega) < \kappa_{pen\mathbb{N}}(\omega')$ genau dann, wenn $\text{Vrank}_{\mathbb{N}}(\omega) < \text{Vrank}_{\mathbb{N}}(\omega')$, und damit sind alle $\kappa_{pen\mathbb{N}}$-bevorzugten Welten auch Penalty-bevorzugt, und das bedeutet, dass beide Logiken gleich sind.

Für den allemeineren Fall $\rho_i \notin \mathbb{N}$, also $\rho_i \in \mathbb{R}^+$, kann man entweder die ρ_i so auf $\mathbb{N}$ abbilden, dass für alle ρ_i gilt: falls eine Summe $\sum_{1 \leq i \leq k} \rho_i$ von Penalties kleiner als eine andere Summe $\sum_{1 \leq j \leq l} \rho_j$ ist, dann muss auch die Summe der natürlichen Zahlen, auf die die ρ_i abgebildet werden, kleiner als die Summe der natürlichen Zahlen sein, auf die die ρ_j abgebildet werden. Falls z.B. $\rho_i \in \mathbb{Q}$, dann kann man die ρ_i durch Brüche darstellen. Wenn man diese Brüche auf den Nenner 1 erweitert, dann erhält man die entsprechenden natürlichen Zahlen.

Diese Bedingung für die obigen Summen garantiert, dass die Folgerungsrelation der Penalty-Logik bei der Abbildung der Penalites auf natürliche Zahlen nicht verfälscht wird, da die Penalty-bevorzugten Welten unter den reell-wertigen Penalties ρ_i auch die Penalty-bevorzugten Welten unter dem Bild der ρ_i sind. Die konditionale Indifferenz dieser $\mathbb{N}$-wertigen Penalty-Funktionen kann man wieder aus Definition 3.0.11 ableiten.

Falls die ρ_i irrationale reelle Zahlen sind, dann ist es schwieriger, die korrespondierenden natürlichen Zahlen zu finden, die genau dieselbe Folgerungsrelation beschreiben. Deshalb wird hier die folgende Möglichkeit gewählt, die konditionale Indifferenz der Funktion Vrank zu überprüfen. Man definiert analog zu Definition 3.0.10 :

Definition 4.2.1 (Konditionale Indifferenz von Penalty-Funktionen)
Eine Penalty-Funktion Vrank ist konditional indifferent in bezug auf Δ genau dann, wenn $\text{Vrank}(\hat{\omega}_1) = \text{Vrank}(\hat{\omega}_2)$ immer dann, wenn $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$ für $\hat{\omega}_1 \equiv_\top \hat{\omega}_2$

In diesem Fall muss man zusätzlich angeben, was $\text{Vrank}(\hat{\omega})$ bedeuten soll. Diese Definition lautet analog zu Definition 3.0.8:

Definition 4.2.2 (Fortsetzung von Vrank auf $\hat{\omega}$)
Eine Penalty-Funktion $\text{Vrank} : \Omega \longrightarrow \mathbb{R}^+$ wird zu einem Homomorphismus $\text{Vrank} : \hat{\Omega} \longrightarrow \mathbb{R}$ fortgesetzt durch

$$\text{Vrank}(\hat{\omega}) = \text{Vrank}(\omega_1^{r_1} \cdots \omega_r^{r_n}) := r_1 \text{Vrank}(\omega_1) + \ldots + r_n \text{Vrank}(\omega_n)$$

$(\mathbb{R}, +)$ ist eine additive Gruppe mit neutralem Element 0. Wenn man nun zeigen kann, dass es einen Homomorphismus $V : F_\Delta \longrightarrow \mathbb{R}$ gibt, der aus der konditionalen Struktur $\sigma(\hat{\omega})$ aller $\hat{\omega}$ die „Plausibilität" Vrank ableitet, dann folgt aus

$$\begin{aligned} \sigma(\hat{\omega}_1) &= \sigma(\hat{\omega}_2) \\ \Rightarrow \text{Vrank}(\hat{\omega}_1) = V(\sigma(\hat{\omega}_1)) &= V(\sigma(\hat{\omega}_2)) = \text{Vrank}(\hat{\omega}_2) \end{aligned}$$

Wegen

$$\sigma_\Delta(\omega) \overset{Def.\ 3.0.5}{=} \prod_{1 \leqslant i \leqslant n} \sigma_i(\omega) = \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} a_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} a_i^- \text{ und} \tag{4.1}$$

$$\text{Vrank}(\omega) \overset{Def.\ 4.1.1}{=} \sum_{\omega \not\models \varphi_i} \rho_i \tag{4.2}$$

erhält man die Plausibilität Vrank einer Interpretation ω aus der konditionalen Struktur, indem man z.B.

$$V(a_i^+) := 0,\ V(a_i^-) := \rho_i,\ V(1) := 0 \tag{4.3}$$

setzt und

$$\begin{aligned} & V((a_1^+)^{r_1}(a_1^-)^{r_2}\cdots(a_n^+)^{r_{2n-1}}(a_n^-)^{r_{2n}}) \\ := \; & r_1V(a_1^+) + r_2V(a_1^-) + \ldots + r_{2n-1}V(a_n^+) + r_{2n}V(a_n^-) \\ = \; & 0 + r_2\rho_1 + \ldots + 0 + r_{2n}\rho_n \\ = \; & \sum_{i=1}^{n} r_{2i}\rho_i \end{aligned}$$

definiert. $V(1) = 0$ in Gleichung (4.3) bedeutet, dass das neutrale Element 1 der Gruppe F_Δ auf das neutrale Element 0 der Gruppe $(\mathbb{R}, +)$ abgebildet wird. Für einzelne Welten ergibt sich

$$V(\sigma_\Delta(\omega)) = V(\prod_{\omega\models A_iB_i} a_i^+ \prod_{\omega\models A_i\overline{B}_i} a_i^-) = \sum_{\omega\models A_i\overline{B}_i} \rho_i = \sum_{\omega\not\models\varphi_i} \rho_i = \mathrm{Vrank}(\omega)$$

Für Produkte von Welten folgt: Falls

$$\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$$

also

$$\frac{\sigma_\Delta(\hat{\omega}_1)}{\sigma_\Delta(\hat{\omega}_2)} = \frac{(a_1^+)^{r_1}(a_1^-)^{r_2}\cdots(a_n^+)^{r_{2n-1}}(a_n^-)^{r_{2n}}}{(a_1^+)^{s_1}(a_1^-)^{s_2}\cdots(a_n^+)^{s_{2n-1}}(a_n^-)^{s_{2n}}} = 1$$

dann muss für alle r_i und s_i gelten:

$$r_i = s_i \; \forall \; 1 \leq i \leq 2n$$

und damit folgt

$$V(\sigma_\Delta(\hat{\omega}_1)) = V(\sigma_\Delta(\hat{\omega}_2))$$

Es bleibt noch zu zeigen, dass $V(\sigma_\Delta(\hat{\omega})) = \text{Vrank}(\hat{\omega})$:

$$\begin{aligned}
V(\sigma_\Delta(\hat{\omega})) &= V\left(\sigma_\Delta(\omega_1^{r_1}\cdots\omega_n^{r_n})\right) \\
&= V(\sigma_\Delta(\omega_1)^{r_1}\cdots\sigma_\Delta(\omega_n)^{r_n}) \\
&= V\left((\prod_{\omega_1 \models A_i B_i} a_i^+)^{r_1} (\prod_{\omega_1 \models A_i \overline{B}_i} a_i^-)^{r_1} \cdots (\prod_{\omega_n \models A_i B_i} a_i^+)^{r_n} (\prod_{\omega_n \models A_i \overline{B}_i} a_i^-)^{r_n} \right) \\
&= r_1 \sum_{\omega_1 \not\models \varphi_i} \rho_i + \ldots + r_n \sum_{\omega \not\models \varphi_i} \rho_i \\
&= r_1 \text{Vrank}(\omega_1) + \ldots + r_n \text{Vrank}(\omega_n) \\
&= \text{Vrank}(\hat{\omega})
\end{aligned}$$

Insgesamt folgt damit, dass Penalty-Funktionen konditional indifferent sind, da aus $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$ folgt: $\text{Vrank}(\hat{\omega}_1) = \text{Vrank}(\hat{\omega}_2)$. $\square$

Kapitel 5

Lexikographisches System

5.1 Grundlagen des lexikographischen Systems

Das lexikographische System wird ausführlich in [5] vorgestellt. Sei Δ wieder eine Menge von Defaults d_i. Die lexikographische Logik geht von einer vorgegebenen Stratifikation der Defaults $\Delta = \{\Delta_1, ...\Delta_n\}$ aus. Man stratifiziert also zuerst die Default-Basis Δ in geeigneter Weise, z.B. erhält man eine konsistente Stratifikation durch die Rangordnung der Regeln durch System-Z, da in System-Z alle Defaults einer Ebene alle Defaults in höheren Ebenen tolerieren.

Definition 5.1.1 (lex-bevorzugt $<_{lex}$)
Eine Interpretation ω wird als „lex-bevorzugt“ zu einer Interpretation ω' bezeichnet,

$$\omega <_{lex} \omega'$$

genau dann, wenn ein Index $1 \leq i \leq n$ existiert so dass

1. $$|\Delta_i^{sat}(\omega)| > |\Delta_i^{sat}(\omega')| \tag{5.1}$$

   und

2. falls i<n,
$$\forall n \geq j > i, |\Delta_j^{sat}(\omega)| = |\Delta_j^{sat}(\omega')| \tag{5.2}$$

Der Begriff „lex-bevorzugt“ definiert also eine Präferenzrelation auf Ω. Die Relation $<_{lex}$ hängt natürlich von der vorgegebenen Stratifikation der Defaultmenge ab. Wie schon in Abschnitt 2.2.2 erwähnt wurde, ist $\Delta_i^{sat}(\omega)$ in Definition 5.1.1 die Menge der Defaults in Δ_i (also der Defaults in der i-ten Ebene), denen ω genügt. $|\Delta_i^{sat}(\omega)|$ ist die Kardinalität dieser Menge, also die Anzahl der Defaults in Δ_i, denen ω genügt.

Eine nicht-monotone Kosequenzrelation, genannt „lexikographische Folgerung“, wird durch $\mid\!\sim_{lex}$ dargestellt und genauso wie in der Penalty-Logik definiert: Die lexikographische Folgerung $\mid\!\sim_{lex}$ basiert auf der lexikographischen Ordnungsrelation auf Ω, $<_{lex}$. Die α_{lex}-bevorzugten Welten werden analog zu Definition 2.5.1 festgelegt: Eine Welt ω ist α_{lex}-bevorzugt, wenn $\omega \models \alpha$ und es keine Welt $\omega' \in \Omega$ gibt mit $\omega' \models \alpha$ und $\omega' <_{lex} \omega$. Die Inferenzrelation $\mid\!\sim_{lex}$ ergibt sich dann aus Definition 2.5.2: Eine Formel β ist eine lexikographische Folgerung aus α, geschrieben

$$\alpha \mid\!\sim_{lex} \beta$$

wenn jede α_{lex}-bevorzugte Welt auch β erfüllt.

Wenn die vorgegebene Stratifikation konsistent ist, dann gilt für alle $d_i = A_i \rightarrow B_i \in \Delta$: $A_i \mid\!\sim_{lex} B_i$.

Alle Defaults in der i-ten Ebene der Stratifikation werden von dem lexikographischen System als gleich wichtig und wichtiger als irgendeine Menge von Defaults in tieferen Ebenen angesehen.

Beispiel 5.1.1 (Lexikographisches System)
Sei $\Delta = \{\Delta_1, \Delta_2\}$ wieder die Default-Basis $\Delta_1 = \{v \rightarrow f\}, \Delta_2 = \{p \rightarrow v, p \rightarrow \neg f\}$. Betrachtet man nun die Welten $\omega_1 := v \wedge \neg p \wedge f$ und $\omega_2 := v \wedge \neg p \wedge \neg f$, dann verifiziert ω_1 den Default $v \rightarrow f \in \Delta_1$ und genügt den Defaults $p \rightarrow v \in \Delta_2$ und $p \rightarrow \neg f \in \Delta_2$. ω_2 falsifiziert den Default $v \rightarrow f$ aus Δ_1 und genügt ebenfalls den Defaults $p \rightarrow v$ und $p \rightarrow \neg f$ aus Δ_2. Es existiert also ein Index i:=1, so dass $|\Delta_1^{sat}(\omega_1)| = 1 > 0 = |\Delta_1^{sat}(\omega_2)|$ und $|\Delta_2^{sat}(\omega_1)| = 2 = |\Delta_2^{sat}(\omega_2)|$. Deshalb ist ω_1 lex-bevorzugt zu ω_2. Damit ist ω_1 auch die einzige „$v \wedge \neg p$“-bevorzugte Welt, und da $\omega_1 \models f$ gilt, folgert man, dass Vögel, die keine Pinguine sind, fliegen können.

Das folgende Beispiel aus [5] wird im nächsten Abschnitt fortgesetzt, um den Unterschied zwischen dem lexikographischen System und Brewkas bevorzugten Subtheorien zu veranschaulichen.

Beispiel 5.1.2 (Unterschied zwischen $\mid\!\sim_{lex}$ und Brewkas Logik[1])
Sei $\mathcal{L} := \{v, f, p\}$ und $\Delta := \{\Delta_1, \Delta_2, \Delta_3, \Delta_4\}$ mit $\Delta_1 := \{\top \rightarrow \neg p\}$, $\Delta_2 := \{\top \rightarrow \neg f\}$, $\Delta_3 := \{\top \rightarrow \neg v \vee f, \top \rightarrow v \vee p\}$ und $\Delta_4 := \{\top \rightarrow v, \top \rightarrow \neg v\}$. Diese Stratifikation ist nicht konsistent, da sich die Defaults aus Δ_4 nicht gegenseitig tolerieren. Die Kardinalität aller $\Delta_i^{sat}(\omega_j)$ kann man für alle möglichen Welten ω_j aus folgender Tabelle ablesen.

ω_j	$\lvert\Delta_1^{sat}(\omega_j)\rvert$	$\lvert\Delta_2^{sat}(\omega_j)\rvert$	$\lvert\Delta_3^{sat}(\omega_j)\rvert$	$\lvert\Delta_4^{sat}(\omega_j)\rvert$
$\omega_1 = v \wedge f \wedge p$	0	0	2	1
$\omega_2 = v \wedge f \wedge \neg p$	1	0	2	1
$\omega_3 = v \wedge \neg f \wedge p$	0	1	1	1
$\omega_4 = v \wedge \neg f \wedge \neg p$	1	1	1	1
$\omega_5 = \neg v \wedge f \wedge p$	0	0	2	1
$\omega_6 = \neg v \wedge f \wedge \neg p$	1	0	1	1
$\omega_7 = \neg v \wedge \neg f \wedge p$	0	1	2	1
$\omega_8 = \neg v \wedge \neg f \wedge \neg p$	1	1	1	1

Man sieht aus dieser Tabelle sofort, dass die Welt $\omega_7 := \neg v \wedge \neg f \wedge p$ die einzige $\top$-lex-bevorzugte Welt ist, da für diese Welt die Mengen $\Delta_3^{sat}(\omega_7)$ und $\Delta_4^{sat}(\omega_7)$ die maximal mögliche Kardinalität $|\Delta_3^{sat}(\omega_7)| = 2$ und $|\Delta_4^{sat}(\omega_7)| = 1$ haben, und für alle anderen Welten ω_j mit $|\Delta_4^{sat}(\omega_j)| = 1$ und $|\Delta_3^{sat}(\omega_j)| = 2$, also für ω_1, ω_2 und ω_5, gilt $|\Delta_2^{sat}(\omega_j)| = 0$, aber $|\Delta_2^{sat}(\omega_7)| = 1$. Da die $\top$-lex-bevorzugte Welt ω_7 $\neg v$ erfüllt, folgert man in diesem Beispiel $\top \mid\!\sim_{lex} \neg v$.

5.2 Konditionale Indifferenz und lexikographisches System

Das Verhalten des lexikographischen Systems kann man durch konditional indifferente OCFs modellieren, indem man Zusatzbedingungen an die κ_i^- stellt. Man geht von der vorgegebenen Stratifikation Δ der Defaults aus. Sei d_{ki} der i-te Default in der k-ten Ebene und κ_{ki} der entsprechende Rang zu diesem Default. Zunächst kann $\kappa_1 \in \mathbb{N}$ beliebig gewählt werden. Danach ordnet man,

[1] vgl.[5] Beispiel 2

beginnend mit i:=2, den κ_i der i-ten Ebene natürliche Zahlen zu, so dass

$$\sum_{k=1}^{i-1} |\Delta_k| \kappa_k < \kappa_i \tag{5.3}$$

für i=2,...n, wobei $|\Delta_k|$ wieder die Anzahl der Defaults in Δ_k ist. Eine solche Zuweisung ist immer möglich, denn wenn die linke Seite der Gleichung (5.3) vorgegeben ist, findet man immer ein $\kappa_i \in \mathbb{N}$, das größer als die Summe auf der linken Seite ist. Jeder Default aus Δ_j erhält denselben Rang κ_j, also $\kappa_{il} = \kappa_{il'}$ $\forall l, l'$ mit $1 \leq \{l, l'\} \leq |\Delta_i|$.

Alle OCFs, die die Bedingung (5.3) erfüllen, werden mit $\kappa_{\oplus 2}$ bezeichnet. Da alle $\kappa_{\oplus 2}$ von der gewählten Stratifikation Δ abhängen, müsste man sie genauer durch $\kappa_{\oplus 2}(\Delta_1, \ldots, \Delta_n)$ bezeichnen.

Den Rang einer Welt ω bestimmt man wieder wie bei der Penalty-Logik durch

$$\kappa_{\oplus 2}(\omega) = \sum_{\omega \not\models \varphi_{ij}} \kappa_{ij}$$

Es ist leicht zu sehen, dass alle $\kappa_{\oplus 2}$ konditional indifferent bezüglich Δ sind, da sie die Bedingungen von Definition 3.0.11 erfüllen mit $\kappa_0 := 0$, $\kappa_{ij}^+ := 0$, $\kappa_{ij}^- := \kappa_i$ für beliebige j.

Definition 5.2.1 (lexikographische OCFs $\kappa_{\oplus 2}$)
Sei $K_{\oplus 2}{}^2$ die Menge aller OCFs $\kappa_{\oplus 2}$ auf Δ, die die obige Bedingung (5.3) erfüllen:
$\sum_{k=1}^{i-1} |\Delta_k| \kappa_k < \kappa_i \; \forall i = 2, \ldots, n$.
Alle Defaults aus einer Ebene Δ_i sollen denselben Rang κ_i haben, also für alle i, $1 \leq i \leq n$, und alle j,k, $1 \leq \{j, k\} \leq |\Delta_i|$, gilt $\kappa_{ij} = \kappa_{ik}$.
Für alle $\omega \in \Omega$ gilt dann:

$$\kappa_{\oplus 2}(\omega) = \sum_{\omega \not\models \varphi_{ij}} \kappa_{ij}$$

Sei $\Delta = \Delta_1, \ldots, \Delta_n$ eine gegebene Stratifikation.

Satz 5.2.1 (Lexikograph. System/konditional indifferente OCFs)
Die von der Stratifikation Δ induzierte Präferenzrelation $\mid\!\sim_{lex}$ lässt sich durch eine bezüglich Δ konditional indifferente OCF $\kappa_{\oplus 2} \in K_{\oplus 2}$ auf Δ darstellen. Eine Interpretation ω ist genau dann $\kappa_{\oplus 2}$-bevorzugt zu einer Interpretation ω', wenn ω lex-bevorzugt zu ω' ist.

Beweis: $\Longrightarrow$ [3]
Sei also ω lex-bevorzugt zu ω' und sei i so gewählt, dass

$$\begin{aligned} \forall j > i : |\Delta_j^{sat}(\omega)| &= |\Delta_j^{sat}(\omega')| \text{ und} \quad (5.2) \\ |\Delta_i^{sat}(\omega)| &> |\Delta_i^{sat}(\omega')| \quad (5.1) \end{aligned}$$

Sei $\Delta_k^{fals}(\omega)$ die Menge aller Defaults in Δ_k, die von ω falsifiziert werden, und $|\Delta_k^{fals}(\omega)|$ sei die Anzahl der Defaults in dieser Menge. Sei κ eine OCF aus $K_{\oplus 2}$. Dann gilt

$$\begin{aligned} \kappa(\omega) &= \sum_{k=1}^{n} \sum_{\substack{j=1 \\ \omega \not\models \varphi_{kj}}}^{|\Delta_k|} \kappa_{kj} \\ &= \sum_{k=1}^{n} \kappa_k |\Delta_k^{fals}(\omega)| \text{ (wegen } \kappa_{ki} = \kappa_{kj} \text{ f. alle } 1 \leq \{i,j\} \leq |\Delta_k|) \\ &= \sum_{k=1}^{i-1} \kappa_k |\Delta_k^{fals}(\omega)| + \kappa_i |\Delta_i^{fals}(\omega)| + \sum_{k=i+1}^{n} \kappa_k |\Delta_k^{fals}(\omega)| \\ &= \sum_{k=1}^{i-1} \kappa_k |\Delta_k^{fals}(\omega)| + \kappa_i |\Delta_i^{fals}(\omega)| + \sum_{k=i+1}^{n} \kappa_k |\Delta_k^{fals}(\boldsymbol{\omega'})| \end{aligned}$$

[2] genauer: $K_{\oplus 2}(\boldsymbol{\Delta_1, \ldots, \Delta_n})$
[3] vgl. [2] Lemma 11

$$\text{wegen (5.2)}: \forall n \geq j > i : |\Delta_j^{fals}(\omega)| = |\Delta_j^{fals}(\omega')|$$

$$< \quad \kappa_i + \qquad \kappa_i|\Delta_i^{fals}(\omega)| + \sum_{k=i+1}^{n} \kappa_k|\Delta_k^{fals}(\omega')|$$

$$(\text{wegen (5.3) und} \Delta_k^{fals}(\omega) \subseteq \Delta_k \Rightarrow |\Delta_k^{fals}(\omega)| \leq |\Delta_k|)$$

$$\leq \kappa_i \cdot (|\Delta_i^{fals}(\omega')| - |\Delta_i^{fals}(\omega)|) + \kappa_i|\Delta_i^{fals}(\omega)| + \sum_{k=i+1}^{n} \kappa_k|\Delta_k^{fals}(\omega')|$$

$$(\text{da wegen (5.1)}: |\Delta_i^{sat}(\omega)| > |\Delta_i^{sat}(\omega')| \Rightarrow |\Delta_i^{fals}(\omega)| < |\Delta_i^{fals}(\omega')|$$

$$\Rightarrow (|\Delta_i^{fals}(\omega')| - |\Delta_i^{fals}(\omega)|) \geq 1)$$

Damit gilt:

$$\kappa(\omega) < \kappa_i \cdot (|\Delta_i^{fals}(\omega')| - |\Delta_i^{fals}(\omega)|) + \kappa_i \cdot |\Delta_i^{fals}(\omega)|$$

$$+ \sum_{k=i+1}^{n} \kappa_k \cdot |\Delta_k^{fals}(\omega')|$$

$$\Rightarrow \kappa(\omega) < \kappa_i|\Delta_i^{fals}(\omega')| + \sum_{k=i+1}^{n} \kappa_k|\Delta_k^{fals}(\omega')|$$

$$\Rightarrow \kappa(\omega) < \sum_{k=1}^{i-1} \kappa_k \cdot |\Delta_k^{fals}(\omega')| + \kappa_i \cdot |\Delta_i^{fals}(\omega')| + \sum_{k=i+1}^{n} \kappa_k \cdot |\Delta_k^{fals}(\omega')|$$

$$\Rightarrow \kappa(\omega) < \kappa(\omega')$$

$\Longleftarrow$

Sei nun $\kappa \in K_{\oplus 2}$ eine OCF auf der gegebenen Stratifikation $\Delta = \Delta_1, \ldots, \Delta_n$ und $\kappa(\omega) < \kappa(\omega')$. Angenommen, ω sei nicht lex-bevorzugt zu ω'. Dann gibt es zwei mögliche Fälle:

1. ω' ist auch nicht lex-bevorzugt zu ω. In diesem Fall falsifizieren ω und ω' genau dieselbe Anzahl von Defaults in jedem Δ_i. Das bedeutet dann $\kappa(\omega) = \kappa(\omega')$, und dies widerspricht der Hypothese.

2. Der zweite mögliche Fall ist: ω' ist lex-bevorzugt zu ω. Dann wäre aber nach dem $\Rightarrow$-Teil des Beweises $\kappa(\omega') < \kappa(\omega)$. Dies ist ebenfalls ein Widerspruch.

Daher muss ω lex-bevorzugt zu ω' sein. $\square$

Korollar 5.2.1 (Lexikographisches System und $K_{\oplus 2}$). Sei Δ eine Default-Basis mit einer vorgegebenen Stratifikation. Sei $\mid\!\sim_{lex}$ die Folgerungsrelation des lexikographischen Systems auf Δ und $\kappa_{\oplus 2}$ eine entsprechende konditional indifferente OCF aus $K_{\oplus 2}$. Dann gilt $\alpha \mid\!\sim_{lex} \beta$ genau dann, wenn $\alpha \mid\!\sim_{\oplus 2} \beta$.

BEWEIS: Der Beweis ergibt sich aus Satz 2.5.1 und Satz 5.2.1, da die bevorzugten Welten in beiden Logiken gleich sind. $\square$

Aus Korollar 5.2.1 ergibt sich auch, dass es zu jedem lexikographischen Modell für eine Default-Menge, d.h. einem lex-Modell, in dem alle Defaults ableitbar sind, auch eine c-Repräsentation $\kappa_{\oplus 2} \in K_{\oplus 2}$ gibt, die dieselbe Inferenzrelation induziert. Eine bezüglich der lexikographischen Logik konsistente Default-Menge ist daher auch konsistent im Sinne von Definition 2.4.1.

Kapitel 6

Brewkas bevorzugte Subtheorien

6.1 Grundlagen der bevorzugten Subtheorien von Brewka

Der Ansatz der bevorzugten Subtheorien von Brewka hat viel Ähnlichkeit mit dem lexikographischen System. Der Haupt-Unterschied zwischen diesen beiden Logiken ist die Definition der Ordnung auf Ω. Sei Δ wieder eine Menge von Defaults und $\Delta = \Delta_1, \ldots, \Delta_n$ eine vorgegebene Stratifikation auf Δ.

Definition 6.1.1 (B-bevorzugt $<_B$ [6])
Eine Interpretation ω wird in der Theorie von Brewka als „B-bevorzugt“ zu einer Interpretation ω' betrachtet,

$$\omega <_B \omega'$$

wenn es einen Ebenenindex i gibt, so dass [1]

1. $$\text{falls } i < n, \Delta_i^{sat}(\omega') \subset \Delta_i^{sat}(\omega) \tag{6.1}$$

   und

2. $$\forall n \geq j > i : \Delta_j^{sat}(\omega') = \Delta_j^{sat}(\omega) \tag{6.2}$$

Eine Formel β ist wieder eine Brewka-Folgerung aus α,

$$\alpha \mid\!\sim_B \beta$$

wenn jede α_B-bevorzugte Welt auch β erfüllt, wobei die α_B-bevorzugten Welten wieder wie in Definition 2.5.4 definiert werden: Eine Welt ω wird als α_B-bevorzugt bezeichnet, wenn $\omega \models \alpha$ und es gibt keine Welt $\omega' \in \Omega$ mit $\omega' \models \alpha$ und $\omega' <_B \omega$.

In Brewkas bevorzugten Subtheorien werden also Inklusionsbeziehungen zwischen Defaultmengen untersucht, während sich das lexikographische System nur für die Kardinalität der Defaultmengen interessiert. Es ist offensichtlich, dass B-bevorzugte Welten auch lex-bevorzugt sind, da aus $\Delta_i^{sat}(\omega') \subset \Delta_i^{sat}(\omega)$ (6.1) folgt, dass $|\Delta_i^{sat}(\omega')| < |\Delta_i^{sat}(\omega)|$ (5.1). Ausserdem gilt natürlich für $\Delta_j^{sat}(\omega') = \Delta_j^{sat}(\omega)$ (6.2) auch $|\Delta_j^{sat}(\omega')| = |\Delta_j^{sat}(\omega)|$ (5.2). Brewkas Inferenzrelation $\mid\!\sim_B$ hängt wie die Relation $\mid\!\sim_{lex}$ von der vorgegebenen Stratifikation der Defaultmenge ab.

[1] Zur Erklärung von $\Delta_i^{sat}(\omega)$ s. 2.2.2

Beispiel 6.1.1 (Brewkas bevorzugte Subtheorien)
Man kann die Logik von Brewka analog auf Beispiel 5.1.1 anwenden. Für die zwei Welten $\omega_1 := v \wedge \neg p \wedge f$ und $\omega_2 := v \wedge \neg p \wedge \neg f$ gilt $\Delta_1^{sat}(\omega_1) = \{v \rightarrow f\} \supset \emptyset = \Delta_1^{sat}(\omega_2)$ und $\Delta_2^{sat}(\omega_1) = \{p \rightarrow v, p \rightarrow \neg f\} = \Delta_2^{sat}(\omega_2)$. Daher ist auch in Brewkas Logik ω_1 B-bevorzugt zu ω_2, so dass man auch aus dieser Logik ableitet, dass Vögel, die keine Pinguine sind, fliegen können.

Um den Unterschied zwischen Brewkas bevorzugten Subtheorien und dem lexikographischen System zu veranschaulichen, wird jetzt das Beispiel 5.1.2 unter der Anwendung von Brewkas Logik fortgesetzt.

Beispiel 6.1.2 (Unterschied zwischen $\mid\!\sim_B$ und $\mid\!\sim_{lex}$)
Beispiel 5.1.2 in der Logik von Brewka:
In dem folgenden Beispiel soll vfp eine Abkürzung für $v \wedge f \wedge p$ sein. Ausserdem wird die Defaultvoraussetzung $\top \rightarrow$ zur Abkürzung weggelassen, d.h. v ist eine Abkürzung für $\top \rightarrow v$.

ω_j	$\Delta_1^{sat}(\omega_j)$	$\Delta_2^{sat}(\omega_j)$	$\Delta_3^{sat}(\omega_j)$	$\Delta_4^{sat}(\omega_j)$
$\omega_1 = vfp$	$\emptyset$	$\emptyset$	$\{\neg v \vee f, v \vee p\}$	$\{v\}$
$\omega_2 = vf\neg p$	$\{\neg p\}$	$\emptyset$	$\{\neg v \vee f, v \vee p\}$	$\{v\}$
$\omega_3 = v\neg fp$	$\emptyset$	$\{\neg f\}$	$\{v \vee p\}$	$\{v\}$
$\omega_4 = v\neg f\neg p$	$\{\neg p\}$	$\{\neg f\}$	$\{v \vee p\}$	$\{v\}$
$\omega_5 = \neg vfp$	$\emptyset$	$\emptyset$	$\{\neg v \vee f, v \vee p\}$	$\{\neg v\}$
$\omega_6 = \neg vf\neg p$	$\{\neg p\}$	$\emptyset$	$\{\neg v \vee f\}$	$\{\neg v\}$
$\omega_7 = \neg v\neg fp$	$\emptyset$	$\{\neg f\}$	$\{\neg v \vee f, v \vee p\}$	$\{\neg v\}$
$\omega_8 = \neg v\neg f\neg p$	$\{\neg p\}$	$\{\neg f\}$	$\{\neg v \vee f\}$	$\{\neg v\}$

Tabelle 6.1: Logik von Brewka

Da weder $\{\top \to v\} \subset \{\top \to \neg v\}$ noch $\{\top \to \neg v\} \subset \{\top \to v\}$ gilt, gibt es in diesem Beispiel keine Welt, die man gegenüber allen anderen Welten als $\top_B$-bevorzugt bezeichnen könnte. Deshalb kann man in diesem Beispiel im Gegensatz zum lexikographischen System auch nicht $\top |\!\sim_B \neg v$ ableiten.

Das folgende Beispiel aus [2] Beispiel 12 wird im nächsten Abschnitt aufgegriffen, um den Unterschied zwischen Brewkas bevorzugten Subtheorien und der im nächsten Abschnitt vorgestellten Theorie von Geffner zu veranschaulichen:

Beispiel 6.1.3 (Unterschied zwischen $|\!\sim_B$ und Geffners Logik)
Die Sprache $\mathcal{L}$ bestehe in diesem Beispiel aus vier Symbolen, $\mathcal{L} := \{p(\text{inguin}), mv(\text{männlicher Vogel}), wv(\text{weiblicher Vogel}), f(\text{liegt})\}$. Die Default-Basis Δ sei $\Delta := \{\Delta_1, \Delta_2\}$ mit $\Delta_1 = \{mv \to f, wv \to f\}$ und $\Delta_2 := \{p \to mv \vee wv, p \to \neg f\}$. Die Default-Basis ist konsistent im Sinne von Definition 2.4.1, da die Welt $\omega_{kons1} := mv \wedge wv \wedge \neg p \wedge f$ die Defaults aus Δ_1 verifiziert und den Defaults aus Δ_2 genügt. Ausserdem verifiziert die Welt $\omega_{kons2} := p \wedge mv \wedge wv \wedge \neg f$ alle Defaults aus Δ_2. Das bedeutet, dass alle Defaults aus Δ_2 von Δ_2 toleriert werden.

Wenn man nun wissen möchte, ob ein Pinguin, der ein weiblicher und kein männlicher Vogel ist, fliegen kann, dann muss man überprüfen, welche der beiden Welten, die die Bedingungen $p \wedge wv \wedge \neg mv$ erfüllen, B-bevorzugt ist. Für $\omega_1 := p \wedge wv \wedge \neg mv \wedge f$ gilt $\Delta_2^{sat}(\omega_1) = \{p \to mv \vee wv\}$ und $\Delta_2^{fals}(\omega_1) = \{p \to \neg f\}$. Für $\omega_2 := p \wedge wv \wedge \neg mv \wedge \neg f$ gilt $\Delta_2^{sat}(\omega_2) = \{p \to mv \vee wv, p \to \neg f\}$ und $\Delta_2^{fals}(\omega_2) = \emptyset$. Wegen $\Delta_2^{sat}(\omega_1) \subset \Delta_2^{sat}(\omega_2)$ ist ω_2 B-bevorzugt zu ω_1, und da in ω_2 $\neg f$ gilt, folgert man in Brewkas Logik, dass Pinguine, die weibliche und keine männlichen Vögel sind, nicht fliegen können.

6.2 Konditionale Indifferenz und Brewkas bevorzugte Subtheorien

Man erhält wieder eine äquivalente Logik aus konditional indifferenten OCFs auf der gegebenen Stratifikation Δ, indem man Zusatzbedingungen an die κ stellt. Sei $\Delta = \Delta_1, \ldots, \Delta_n$ wieder eine vorgegebene Stratifikation. Sei d_{ij} der j-te Default aus der i-ten Ebene.

Definition 6.2.1 (Bewka-OCFs $\kappa_{\oplus 3}$)
Man erhält Rangfunktionen vom Typ $\kappa_{\oplus 3}$, indem man zunächst den Defaults d_{1s} aus Δ_1 beliebige $\kappa_{1s} \in \mathbb{N}$ zuordnet und danach, beginnend bei i:=2, jedem d_{ij}, $j = 1, \ldots, |\Delta_i|$, ein κ_{ij} zuordnet, so dass

$$\text{für alle } i = 1 \ldots n : \sum_{k=1}^{i-1} \sum_{j=1}^{|\Delta_k|} \kappa_{kj} < \kappa_{is}. \tag{6.3}$$

für alle s=1,$\cdots$ $|\Delta_i|$.
$K_{\oplus 3}$[2] sei dann die Menge aller OCFs $\kappa_{\oplus 3}$, die die Bedingung (6.3) erfüllen. Für alle $\omega \in \Omega$ wird $\kappa(\omega)$ dann berechnet durch

$$\kappa(\omega) = \sum_{k=1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj}$$

Man sieht sofort, dass es immer natürliche Zahlen gibt, die die konstruktive Bildungsvorschrift der Bedingung (6.3) erfüllen. Diese Bedingung ist der Bedingung (5.3) des lexikographischen Systems ähnlich. Während in dem lexikographischen System jedoch alle Defaults einer Ebene als gleich wichtig angesehen werden, werden in Brewka's Logik keine Bedingungen an das Verhältnis der Defaults einer Ebene gestellt.

[2]genauer: $K_{\oplus 3}(\boldsymbol{\Delta_1}, \ldots, \boldsymbol{\Delta_n})$,vgl.Def.5.2.1

Satz 6.2.1 (Brewkas bevorzugte Subtheorien/kond.indiff.OCFs) *Jede Präferenzrelation aus Brewkas bevorzugten Subtheorien lässt sich durch eine Menge von konditional indifferenten OCFs $\kappa_{\oplus 3}$ auf Δ darstellen. Genauer gilt: Sei $\Delta = \Delta_1, \ldots, \Delta_n$ eine Stratifikation. Seien ω und ω' Welten aus Ω. Dann ist ω B-bevorzugt zu ω', $\omega <_B \omega'$, genau dann, wenn für alle $\kappa_{\oplus 3} \in K_{\oplus 3}$ auf $\Delta = \{\Delta_1, \ldots, \Delta_n\}$ $\kappa_{\oplus 3}(\omega) < \kappa_{\oplus 3}(\omega')$, d.h. genau dann, wenn ω für alle $\kappa_{\oplus 3}$ $\kappa_{\oplus 3}$-bevorzugt zu ω' ist.*

Die Präferenzrelationen der bevorzugten Subtheorien von Brewka lassen sich also nicht mehr wie die Präferenzrelationen des lexikographischen Systems durch jeweils eine konditional indifferente OCF darstellen, sondern man muss jeder Präferenzrelation eine Menge von OCFs zuordnen.

BEWEIS: (von Satz 6.2.1)

$\Longrightarrow$[3] Es gilt

$$\kappa(\omega) = \sum_{k=1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj}$$

Angenommen, ω ist B-bevorzugt zu ω', und sei i so gewählt, dass

$$\begin{aligned} \forall j > i, \Delta_j^{sat}(\omega) &= \Delta_j^{sat}(\omega') \text{ (6.2) und} \\ \Delta_i^{sat}(\omega) &\supset \Delta_i^{sat}(\omega') \text{ (6.1) d.h.} \\ \forall j > i, \Delta_j^{fals}(\omega) &= \Delta_j^{fals}(\omega') \text{ und} \qquad (6.4) \\ \Delta_i^{fals}(\omega) &\subset \Delta_i^{fals}(\omega') \qquad (6.5) \end{aligned}$$

Dann gilt für alle $\kappa \in K_{\oplus 3}(\Delta)$ auf der vorgegebenen Stratifikation Δ

$$\begin{aligned} \kappa(\omega) &= \sum_{k=1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \qquad (6.6) \\ &= \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \end{aligned}$$

[3] vgl.[2] Lemma 12

$$= \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj}$$

$$\text{(wegen (6.4): } \Delta_k^{fals}(\omega) = \Delta_k^{fals}(\omega') \text{ für } k = i+1 \ldots n)$$

$$< \sum_{d_{ij} \in (\Delta_i^{fals}(\omega') - \Delta_i^{fals}(\omega))} \kappa_{ij} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj}$$

$$\text{(wegen (6.3) } \forall d_{is} \in \Delta_i \text{ und } \Delta_k^{fals}(\omega) \subseteq \Delta_k)$$

$$((\Delta_i^{fals}(\omega') - \Delta_i^{fals}(\omega)) \text{ ist wegen (6.5) nicht leer})$$

Also erhält man

$$\text{(wegen } \sum_{d_{ij} \in \Delta_i^{fals}(\omega') - \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} = \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij})$$

$$\kappa(\omega) < \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj}$$

$$\leq \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj}$$

$\Rightarrow \kappa(\omega) < \kappa(\omega')$

$\Longleftarrow$

Angenommen $\kappa(\omega) < \kappa(\omega')$ für alle $\kappa \in K_{\oplus 3}(\Delta)$ und angenommen, dass ω nicht B-bevorzugt zu ω' ist. Man unterscheidet zwei Fälle:

1. ω' ist B-bevorzugt zu ω. Dann gilt laut dem $\Rightarrow$-Teil des Beweises $\kappa(\omega') < \kappa(\omega)$ für alle κ aus $\kappa_{\oplus 3}$, was der Hypothese widerspricht.

2. ω' ist auch nicht B-bevorzugt zu ω. Dann gilt entweder

(a) für beliebige j $\Delta_j^{sat}(\omega) = \Delta_j^{sat}(\omega')$, aber dies bedeutet, dass $\kappa(\omega') = \kappa(\omega)$, da ω und ω' genau dieselbe Menge von Defaults falsifizieren, was der Hypothese widerspricht, oder

(b) es gibt einen Index i, so dass

$$\forall j > i : \Delta_j^{sat}(\omega) = \Delta_j^{sat}(\omega') \tag{6.7}$$

und

$$\text{weder } \Delta_i^{sat}(\omega) \supset \Delta_i^{sat}(\omega') \text{ noch } \Delta_i^{sat}(\omega') \supset \Delta_i^{sat}(\omega) \tag{6.8}$$

In Fall 2b kann man ein κ konstruieren, das zu $K_{\oplus 3}$ gehört mit $\kappa(\omega') < \kappa(\omega)$ (was der Hypothese widerspricht). κ sollte die zwei folgenden Bedingungen erfüllen: für i=2,...n:

$$(i) \quad \sum_{k=1}^{i-1} \sum_{h=1}^{|\Delta_k|} \kappa_{kh} < \kappa_{is} \text{ für alle } s = 1, \ldots, |\Delta_i| \tag{6.9}$$

$$(ii) \quad \sum_{k=1}^{i-1} \sum_{h=1}^{|\Delta_k|} \kappa_{kh} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij} < \kappa_{is} \text{ für alle } d_{is} \notin \Delta_i^{fals}(\omega') \tag{6.10}$$

Diese Bedingungen teilen jedes Δ_i in zwei Teile, wobei die Defaults d_{ij} in Δ_i, die von ω' falsifiziert werden, eine geringere Priorität, also einen kleineren Wert κ_{ij}, zugewiesen bekommen als diejenigen, denen ω' genügt. Die Bedingung (i) garantiert, dass κ zu $K_{\oplus 3}$ gehört (Gleichung(6.3)), und Bedingung (ii) splittet jedes Δ_i in die zwei Teile der falsifizierten und erfüllten Defaults auf, indem die Defaults nach ihrer Priorität sortiert werden.

Daher erhält man:

$$
\begin{aligned}
\kappa(\omega') &= \sum_{k=1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} \\
&= \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} \\
&= \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega')} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\boldsymbol{\omega})} \kappa_{kj} \\
&\qquad \text{(da für k=i+1,...n } \Delta_k^{fals}(\omega') = \Delta_k^{fals}(\omega) \text{ (6.7))} \\
&= \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega')} \kappa_{kj} + \sum_{\boldsymbol{d_{ij} \in \Delta_i^{fals}(\omega') - \Delta_i^{fals}(\omega)}} \kappa_{ij} \\
&\qquad + \sum_{\boldsymbol{d_{ij} \in \Delta_i^{fals}(\omega') \cap \Delta_i^{fals}(\omega)}} \kappa_{kj} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \\
&< \sum_{\boldsymbol{k=1}}^{\boldsymbol{i-1}} \sum_{\boldsymbol{h=1}}^{\boldsymbol{|\Delta_k|}} \boldsymbol{\kappa_{kh}} + \sum_{\boldsymbol{d_{ij} \in \Delta_i^{fals}(\omega')}} \kappa_{ij} \\
&\qquad + \sum_{d_{ij} \in \Delta_i^{fals}(\omega') \cap \Delta_i^{fals}(\omega)} \kappa_{kj} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \\
&\qquad \text{(wegen } \Delta_k^{fals}(\omega') \subset \Delta_k \text{)} \\
&\qquad \text{und } \Delta_i^{fals}(\omega') - \Delta_i^{fals}(\omega) \subset \Delta_i^{fals}(\omega') \\
&< \sum_{\boldsymbol{d_{ij} \in \Delta_i^{fals}(\omega) - \Delta_i^{fals}(\omega')}} \kappa_{ij} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega') \cap \Delta_i^{fals}(\omega)} \kappa_{ij} \\
&\qquad + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \\
&\qquad \text{(wegen (6.10)) für alle } d_{is} \notin \Delta_i^{fals}(\omega')
\end{aligned}
$$

$$
\begin{aligned}
&(\Delta_i^{fals}(\omega) - \Delta_i^{fals}(\omega') \text{ ist wegen (6.8) nicht leer)} \\
\leq & \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} \\
\leq & \sum_{k=1}^{i-1} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} + \sum_{d_{ij} \in \Delta_i^{fals}(\omega)} \kappa_{ij} + \sum_{k=i+1}^{n} \sum_{d_{kj} \in \Delta_k^{fals}(\omega)} \kappa_{kj} = \kappa(\omega)
\end{aligned}
$$

$\Rightarrow \kappa(\omega') < \kappa(\omega)$ und daher die These

□

Korollar 6.2.1 (Folgerungsrelation von $\mid\!\sim_B$ und $\kappa_{\oplus 3}$). Sei $\mid\!\sim_B$ die Inferenzrelation aus Brewkas System der bevorzugten Subtheorien, und sei $\mid\!\sim_{\oplus 3}$ die Folgerungsrelation der entprechenden OCFs. D.h. $\alpha \mid\!\sim_{\oplus 3} \beta$ genau dann, wenn für alle OCFs aus $K_{\oplus 3}$, die auf derselben Stratifikation wie $\mid\!\sim_B$ basieren, β in allen $\alpha_{\oplus 3}$-bevorzugten Welten gilt. Dann gilt für alle Δ $\alpha \mid\!\sim_{\oplus 3} \beta$ genau dann, wenn $\alpha \mid\!\sim_B \beta$.

Beweis: Da die bevorzugten Welten für beide Präferenzrelationen $<_B$ und $<_{\oplus 3}$ gleich sind, sind auch die Inferenzrelationen $\mid\!\sim_B$ und $\mid\!\sim_{\oplus 3}$ gleich. Die Inferenzrelation $\mid\!\sim_B$ basiert auf Definition 2.5.2, also $\alpha \mid\!\sim_B \beta$ genau dann, wenn jede α_B-bevorzugte Welt auch β erfüllt, und für alle $\kappa_{\oplus 3} \in K_{\oplus 3}$ gilt $\alpha \mid\!\sim_{\oplus 3} \beta$ genau dann, wenn nach Satz 2.5.1 jede $\alpha_{\oplus 3}$-bevorzugte Welt auch β erfüllt. □

Es folgt, dass es für jedes Modell aus Brewkas Logik für eine Default-Menge Δ auch eine Menge von c-Repräsentationen $\kappa_{\oplus 3}$ für Δ gibt, so dass eine bezüglich Brewkas Logik konsistente Default-Menge auch konsistent im Sinne von Definition 2.4.1 ist.

An Beispiel 6.1.2 sieht man, dass man in Brewkas Logik nicht allen Welten eine bestimmte Position in einem Ordnungssystem auf Ω zuordnen kann. Wegen $\Delta_4^{sat}(\omega_1) = \{\top \to v\}$ und $\Delta_4^{sat}(\omega_5) = \{\top \to \neg v\}$, aber weder $\{\top \to v\} \subseteq \{\top \to \neg v\}$ noch $\{\top \to \neg v\} \subseteq \{\top \to v\}$ gilt laut Definition 6.1.1 weder $\omega_1 <_B \omega_5$ noch $\omega_5 <_B \omega_1$. Man kann aber auch nicht $\omega_1 =_B \omega_5$ ableiten.

Da man den Welten ω_1 und ω_5 in diesem Beispiel keinen eindeutigen Rang P in dem Relationensystem $<_B$ zuordnen kann, ist es auch schwierig, dem Produkt $\omega_1 \cdot \omega_5$ einen Rang bezüglich $<_B$ zuzuordnen. Die Tatsache, dass die Ordnung $<_B$ auf Ω nicht total, sondern nur partiell ist, ist der Grund dafür, dass eine einzelne OCF zur Darstellung der Inferenzrelation $\mid\!\sim_B$ nicht genügt.

Das Prinzip der konditionalen Indifferenz ist aber in Brewkas Logik insoweit gültig, als für alle Welten ω_i, ω_j, denen eine Position P in dem Relationensystem $<_B$ zugewiesen werden kann und die dieselbe konditionale Struktur $\sigma_\Delta(\omega_i) = \sigma_\Delta(\omega_j)$ haben, diese Position in dem Ordnungssystem gleich ist, also $P(\omega_i) = P(\omega_j)$, falls $\sigma_\Delta(\omega_i) = \sigma_\Delta(\omega_j)$ für $\omega_i \equiv_\top \omega_j$. Und falls $P(\omega_i)$ undefiniert ist mit $\sigma_\Delta(\omega_i) = \sigma_\Delta(\omega_j)$, dann ist auch $P(\omega_j)$ undefiniert.

Die Begründung für diese Eigenschaft der „partiellen konditionalen Indifferenz“ ergibt sich daraus, dass die Inferenzrelation $\mid\!\sim_B$ durch eine Menge von OCFs dargestellt werden kann und dass alle diese OCFs konditional indifferent bezüglich Δ sind.

Kapitel 7

Geffners conditional entailment

7.1 Grundlagen der konditionalen Folgerungsrelation nach Geffner

Die vierte Logik, die untersucht werden soll, ist Geffners konditionale Folgerungsrelation (conditional entailment) [7]. Geffners Logik basiert auf einer partiellen Ordnung zwischen den Defaults einer Default-Basis Δ. Eine Ordnung zwischen den Defaults einer Default-Basis Δ, bezeichnet durch $>_\Delta$, wird als *erreichbar* auf Δ bezeichnet, wenn eine Subbasis $D \subseteq \Delta$ immer, wenn D einen Default $d \in \Delta$ nicht toleriert, einen Default d' enthält, so dass $d >_\Delta d'$. Jede erreichbare Ordnung $>_\Delta$ auf Δ induziert eine Ordnung auf Ω:

Definition 7.1.1 (Geffner-Präferenzrelation $<_\Omega$)
Sei $>_\Delta$ eine erreichbare Ordnung auf der gegebenen Defaultmenge Δ. Es gilt

$$\omega >_\Omega \omega'$$

genau dann, wenn für jeden Default d, der von ω falsifiziert wird und dem ω' genügt, ein Default d' existiert, dem ω genügt und der von ω' falsifiziert wird, so dass $d' >_\Delta d$.

Jede dieser Ordnungen wird als „erreichbar durch die Default-Basis Δ“ bezeichnet. Das Quadrupel $(\Omega, >_\Omega, \Delta, >_\Delta)$ wird bevorzugte erreichbare Struktur genannt. Geffner definiert dann die conditional-entailment-Relation $\mid\!\sim_G$ wie gewohnt durch

$$\alpha \mid\!\sim_G \beta$$

genau dann, wenn für jede erreichbare bevorzugte Struktur $(\Omega, >_\Omega, \Delta, >_\Delta)$ β in allen $>_\Omega$-bevorzugten Modellen von α gilt.

Beispiel 7.1.1 (Geffner-Logik)
Die Default-Basis sei wieder $\Delta = \{d_1 = v \to f, d_2 = p \to v, d_3 = p \to \neg f\}$. Man legt die Ordnung auf den Defaults so fest, dass $d_1 <_\Delta d_2$ und $d_1 <_\Delta d_3$. Da die Ordnung nur partiell und nicht total sein muss, definiert man für die Defaults d_2 und d_3 keine Priorität, also weder $d_2 <_\Delta d_3$ noch $d_3 <_\Delta d_2$. Sei $\omega_1 := v \wedge p \wedge f$ und $\omega_2 := v \wedge p \wedge \neg f$. Es gilt $\Delta^{sat}(\omega_1) = \{v \to f, p \to v\}$, $\Delta^{fals}(\omega_1) = \{p \to \neg f\}$, $\Delta^{sat}(\omega_2) = \{p \to v, p \to \neg f\}$ und $\Delta^{fals}(\omega_2) = \{v \to f\}$. Es gilt $\omega_2 >_\Omega \omega_1$, da für den Default $d_1 = v \to f$, der von ω_2 falsifiziert wird und dem ω_1 genügt, der Default $d_3 = p \to \neg f$ existiert, der von ω_1 falsifiziert wird und dem ω_2 genügt, mit $d_3 >_\Delta d_1$. Deshalb schließt man auch in Geffners conditional entailment bezüglich $>_\Delta$, dass Pinguine, die Vögel sind, nicht fliegen können, da ω_2 das einzige $(p \wedge v)_\Omega$-bevorzugte Modell bezüglich $>_\Delta$ ist und da dieses Modell $\neg f$ erfüllt.

Um zu zeigen, dass aus Geffners Theorie andere Folgerungen als aus Brewkas Theorie abgeleitet werden können, wird auf das Beispiel 6.1.3 nun Geffners Theorie angewandt:

Beispiel 7.1.2 (Unterschied zwischen Theorie $\mid\!\sim_B$ und $\mid\!\sim_G$)
Sei $\mathcal{L} := \{p, mv, wv, f\}$ wieder die Symbolmenge aus Beispiel 6.1.3 und $\Delta := \{d_1 = p \to mv \vee wv, d_2 = mv \to f, d_3 = wv \to f, d_4 = p \to \neg f\}$

die Default-Basis aus diesem Beispiel. In Geffners Logik kann man folgende erreichbare Ordnung auf der Default-Basis definieren: $d_1 >_\Delta d_3 >_\Delta d_4 >_\Delta d_2$. *Man kann nachprüfen, dass diese Ordnung erreichbar ist. Z.B. toleriert die Defaultmenge* $\Delta_{erreichb} := \{d_1, d_2, d_3\}$ *den Default* d_4 *nicht, aber* $\Delta_{erreichb}$ *enthält den Default* d_2 *mit* $d_2 <_\Delta d_4$. *Um zu zeigen, dass die Ordnung* $>_\Delta$ *erreichbar ist, muss man diesen Test für alle Defaultmengen, die man aus den vier Defaults bilden kann, durchführen.*

Wenn man nun wieder wissen möchte, ob ein Pinguin, der ein weiblicher und kein männlicher Vogel ist, fliegen kann, dann muss man wieder die zwei Welten $\omega_1 := p \wedge wv \wedge \neg mv \wedge f$ *und* $\omega_2 := p \wedge wv \wedge \neg mv \wedge \neg f$ *miteinander vergleichen. Es gilt* $\Delta^{sat}(\omega_1) = \{d_1, d_2, d_3\}$, $\Delta^{fals}(\omega_1) = \{d_4\}$, $\Delta^{sat}(\omega_2) = \{d_1, d_2, d_4\}$ *und* $\Delta_2^{fals}(\omega_2) = \{d_3\}$. *Da die Welt* ω_2 *den Default* $d_3 = wv \rightarrow f$ *falsifiziert, dem* ω_1 *genügt, und da* ω_1 *den Default* $d_4 = p \rightarrow \neg f$ *falsifiziert, dem* ω_2 *genügt, gilt in dieser Logik wegen* $d_3 >_\Delta d_4$: $\omega_1 >_\Omega \omega_2$ *bezüglich* $>_\Delta$. *Da* ω_1 f *erfüllt und da die obige Ordnung mit* $d_3 >_\Delta d_4$ *erreichbar ist, gibt es in dieser bevorzugten erreichbaren Struktur ein* $>_\Omega$*-bevorzugtes Modell, das* f *erfüllt, und deshalb kann man in diesem Modell im Gegensatz zu Brewkas Logik nicht folgern, dass Pinguine, die weibliche und keine männlichen Vögel sind, nicht fliegen können (vgl. Beispiel 6.1.3).*

7.2 Konditionale Indifferenz und Geffners conditional entailment

Man kann auch Geffners conditional entailment durch konditional indifferente OCFs beschreiben.

Definition 7.2.1 (Geffner-OCFs $\kappa_{\oplus 4}$)
Für jede erreichbare Präferenzrelation $>_\Delta$ auf Δ konstruiert man eine Menge von Funktionen $\kappa_{\oplus 4}$, indem man, beginnend mit dem bezüglich $>_\Delta$ kleinsten Default, jedem Default $d \in \Delta$ ein κ_d zuordnet, so dass

1. für alle d, d'$\in \Delta$, falls $d >_\Delta d'$, dann $\kappa_{d'} < \kappa_d$; und
2. für alle $d \in \Delta$ gilt $\sum_{d':\kappa_{d'}<\kappa_d} \kappa_{d'} < \kappa_d$.

Für alle $\omega \in \Omega$ gilt wieder $\kappa(\omega) = \sum_{\omega \not\models \varphi_i} \kappa_i$.
$K_{\oplus 4}(\Delta, > \Delta)$ sei die Menge $\{\kappa_{\oplus 4}(\Delta, >_\Delta)\}$ aller OCFs, die man anhand der Bedingungen 1. und 2. auf einer vorgegebenen Defaultmenge Δ mit einer vorgegebenen Präferenzrelation $>_\Delta$ auf Δ bilden kann.

Durch die zweite Bedingung wird folgende Forderung an die κ_d gestellt: falls ein κ_d, das einem Default d entspricht, größer als die κ'_{d_i} ist, die anderen Defaults $d_{i'}$ entsprechen, dann muss es auch größer als deren Summe sein. Es ist klar, dass es zu jeder Ordnung $>_\Delta$ auf Δ Zuweisungen κ_d für alle $d \in \Delta$ gibt, die beide Bedingungen erfüllen.
Sei $>_\Delta$ eine erreichbare Präferenzrelation auf Δ. Jedes κ aus $K_{\oplus 4}(\Delta, >_\Delta)$ erzeugt die erreichbaren Prioritätsstrukturen, die zu $>_\Delta$ assoziiert sind:

Satz 7.2.1 (Geffner-Logik und konditionale Indifferenz)
Sei $>_\Delta$ eine erreichbare Präferenzrelation auf Δ, und $(\Omega, >_\Omega, \Delta, >_\Delta)$ sei eine erreichbare Prioritätsstruktur. Dann gilt für alle $\omega, \omega' \in \Omega$ $\omega >_\Omega \omega'$ genau dann, wenn für alle $\kappa \in K_{\oplus 4}(\Delta, >_\Delta)$ gilt $\kappa(\omega) < \kappa(\omega')$.

Beweis: [1] Für alle ω aus Ω gilt $\kappa(\omega) = \sum_{\omega \not\models \varphi_i} \kappa_i$. Bezeichne $\Delta^{fals}(\omega)$ die Menge der Defaults in Δ, die ω falsifiziert. Für alle ω und ω' in Ω sei Δ_L^{fals},

[1] vgl.[2] Lemma 13

Δ_C^{fals} und Δ_R^{fals} definiert durch

$$\Delta_L^{fals} := \Delta^{fals}(\omega) - \Delta^{fals}(\omega') \tag{7.1}$$
$$\Delta_C^{fals} := \Delta^{fals}(\omega) \cap \Delta^{fals}(\omega') \text{ und} \tag{7.2}$$
$$\Delta_R^{fals} := \Delta^{fals}(\omega') - \Delta^{fals}(\omega) \tag{7.3}$$

Es gilt:

$$\kappa(\omega) = \sum_{d_i \in \Delta_L^{fals}} \kappa_i + \sum_{d_i \in \Delta_C^{fals}} \kappa_i$$
$$\kappa(\omega') = \sum_{d_j \in \Delta_R^{fals}} \kappa_j + \sum_{d_j \in \Delta_C^{fals}} \kappa_j$$

$\Longrightarrow$

Angenommen, dass $\omega >_\Omega \omega'$. Da $(\Omega, >_\Omega, \Delta, >_\Delta)$ eine erreichbare bevorzugte Struktur ist, gilt $\omega >_\Omega \omega'$ genau dann, wenn für alle d_j in Δ_L^{fals} ein d_i' in Δ_R^{fals} existiert so dass $d_i' >_\Delta d_j$, d.h. $\kappa_j < \kappa_{i'}$. Zuerst ordnet man jedem κ_j aus Δ_L^{fals} ein solches $\kappa_{i'}$ aus Δ_R^{fals} zu. Dann benennt man alle κ_j, die einem $\kappa_{i'}$ zugeordnet wurden, um in κ_{ij}. Dadurch kann man Δ_L^{fals} nach dem Index i partitionieren in $\Delta_L^{fals} := (\Delta_{L1}^{fals}, \ldots, \Delta_{Ln}^{fals})$.
Sei $(\Delta_{L1}^{fals}, ...\Delta_{Lk}^{fals})$ eine solche Partition von Δ_L^{fals}, so dass Δ_{Li}^{fals} alle κ_{ij} enthält, denen dasselbe κ_i' in Δ_R^{fals} entspricht, nämlich $\kappa_i' > \kappa_{ij}$ für j=1,...$|\Delta_{Li}|$. Verwendet man die Bedingung (2) aus Definition 7.2.1an die κ, um Geffners System zu konstruieren, dann gilt auch für alle i für die Summe der $\kappa_{ij} \in \Delta_{Li}^{fals}$:

$$\sum_{\kappa_{ij} \in \Delta_{Li}} \kappa_{ij} < \kappa_i' \tag{7.4}$$

Daher gilt

$$\kappa(\omega) = \sum_{d_i \in \Delta_L^{fals}} \kappa_i + \sum_{d_i \in \Delta_C^{fals}} \kappa_i$$
$$\Leftrightarrow \kappa(\omega) = \sum_{\kappa_{1j} \in \Delta_{L1}^{fals}} \kappa_{1j} + \ldots \sum_{\kappa_{kj} \in \Delta_{Lk}^{fals}} \kappa_{kj} + \sum_{d_i \in \Delta_C^{fals}} \kappa_i$$

$$\begin{aligned}
\Rightarrow \kappa(\omega) &< \kappa'_1 + \ldots + \kappa_{k'} + \sum_{d_i \in \Delta_C^{fals}} \kappa_i \text{ (wegen (7.4))}\\
\Rightarrow \kappa(\omega) &< \sum_{d_{i'} \in \Delta_R^{fals}} \kappa_{i'} + \sum_{d_i \in \Delta_C^{fals}} \kappa_i\\
\Rightarrow \kappa(\omega) &< \kappa(\omega')
\end{aligned}$$

$\Longleftarrow$:

Man nimmt jetzt an, dass $\kappa(\omega) < \kappa(\omega')$ für alle $\kappa \in K_{\oplus 4}(\Delta, >_\Delta)$ und dass ω nicht Ω-bevorzugt zu ω' ist, $\omega \not>_\Omega \omega'$. Das bedeutet, dass ein d_{gr} aus Δ_L existiert, so dass für kein $d_{i'}$ aus Δ_R $d_{i'} >_\Delta d_{gr}$ gilt. Man kann nun wie in Satz 6.2.1 eine Funktion κ konstruieren, die zu $K_{\oplus 4}(\Delta, >_\Delta)$ gehört und die $\kappa(\omega') \leq \kappa(\omega)$ erfüllt (Widerspruch). κ wird so konstruiert, dass es die drei folgenden Bedingungen erfüllt:

1. für alle $d, d' \in \Delta$ gilt: falls $d >_\Delta d'$, dann gilt $\kappa_{d'} < \kappa_d$ und
2. für alle $d \in \Delta : \sum_{d' : \kappa_{d'} < \kappa_d} \kappa_{d'} < \kappa_d$ und
3. für alle $d_{i'} \in \Delta_R : \kappa_{di'} < \kappa_{d_{gr}}$

Die Bedingungen (1) und (2) wurden schon bei der Definition von $\kappa_{\oplus 4}$ gefordert. Da für kein $d_{i'}$ aus Δ_R $d_{i'} >_\Delta d_{gr}$ gilt, widersprechen sich die Bedingungen (1) und (3) nicht. Aus diesen Bedingungen ergibt sich $\kappa(\omega') < \kappa(\omega)$ (Widerspruch):

$$\begin{aligned}
\kappa(\omega') &= \sum_{d_i \in \Delta_R} \kappa_i + \sum_{d_i \in \Delta_C} \kappa_i\\
&< \kappa_{d_{gr}} + \sum_{d_i \in \Delta_C} \kappa_i\\
&< \sum_{d_i \in \Delta_L} \kappa_i + \sum_{d_i \in \Delta_C} \kappa_i\\
(&= \kappa(\omega))
\end{aligned}$$

$\square$

Durch $K_{\oplus 4}(\Delta) = \cup\{K_{\oplus 4}(\Delta, >_\Delta)\}$,so dass $>_\Delta$ erreichbar durch Δ ist, wird die Familie der OCFs aus $\kappa_{\oplus 4}(\Delta)$ bezeichnet, die man dadurch erhält, dass man alle Ordnungen $>_\Delta$ in Betracht zieht, die erreichbar durch Δ sind. $K_{\oplus 4}(\Delta)$ hängt also nicht mehr einer bestimmten vorgegebenen Ordnung $>_\Delta$ auf Δ ab. Sei dann $\mid\!\sim_{\oplus 4}$ die Konsequenzrelation, die von $K_{\oplus 4}$ induziert wird, also $\alpha \mid\!\sim_{\oplus 4} \beta$ genau dann, wenn für alle $\kappa \in K_{\oplus 4}$ gilt $\kappa \models \alpha \rightarrow \beta$. Dann gilt

Korollar 7.2.1 (conditional entailment und $\kappa_{\oplus 4}$). Für ein gegebenes Δ gilt $\alpha \mid\!\sim_G \beta$ genau dann, wenn $\alpha \mid\!\sim_{\oplus 4} \beta$.

BEWEIS: Der Satz ist eine direkte Konsequenz aus Satz 7.2.1. □

Man hat in Geffners Logik also dasselbe Problem der „partiellen konditionalen Indifferenz“ wie bei Brewkas bevorzugten Subtheorien. Jede Präferenzrelation $<_\Omega (\Delta, >_\Delta)$ und damit auch die Inferenzrelation $\mid\!\sim_G$ kann nicht durch eine einzelne OCF dargestellt werden. Da jedoch jede Präferenzrelation $>_\Omega (\Delta, >_\Delta)$ durch eine Menge von bezüglich Δ konditional indifferenten OCFs beschrieben werden kann, könnte man Geffners Folgerungsrelation als „partiell konditional indifferent“ bezeichnen.

Kapitel 8

Dempster-Shafer-Theorie

8.1 Basismaße, Glaubens- und Plausibilitätsfunktionen

Die Dempster-Shafer-Theorie (vgl. [10]) ermöglicht die Modellierung und Kombination unsicheren Glaubens. In der Wahrscheinlichkeitstheorie schließt man wegen Definition 2.3.1 (3) (endliche Additivität) aus $V_W(F) = x$ immer $V_W(\neg F) = 1 - x$, was manchmal problematisch sein kann. Wenn man eine Eigenschaft wie „Vögel können fliegen“ zu 80 Prozent glaubt, was durch $\text{bel}(\neg v \vee f) = 0,8$ dargestellt werden kann, dann muss man nicht unbedingt zu 20 Prozent glauben, dass Vögel nicht fliegen können. Es kann auch sein, dass man zu 20 Prozent alle denkbaren Welten für glaubwürdig hält, dass man also zu 20 Prozent völlig unwissend ist. In der Dempster-Shafer-Theorie ist es möglich, solche Evidenzen zu formulieren. Man kann in dieser Theorie sogar sowohl der Regel „Vögel können fliegen“ als auch der Regel „Vögel können nicht fliegen“ das Glaubensmaß 0 zuordnen. Dies drückt die völlige Unkenntnis über die Eigenschaften von Vögeln aus, und das Äquivalent $V_W(v \rightarrow f) = V_W(v \rightarrow \neg f) = 0$ ist in der Wahrscheinlichkeitstheorie wegen $V_W(f) + V_W(\neg f) = 1$ nicht möglich. In der Wahrscheinlichkeitstheorie könnte man diese Unwissenheit nur durch $V_W(v \rightarrow f) = 0,5 = V_W(v \rightarrow \neg f)$ formulieren.

Ausserdem kann man in der Wahrscheinlichkeitstheorie normalerweise aus der Kenntnis von $V_W(A)$ und $V_W(B)$ nicht die Wahrscheinlichkeit $V_W(A \wedge B)$ von A und B berechnen. Genau diese Kombination von Evidenzen ermöglicht in der Dempster-Shafer-Theorie die Dempster'sche Kombinationsregel. Grundlage der Dempster-Shafer-Theorie ist ein Basismaß:

Definition 8.1.1 (Basismaß)
Ein Basismaß über Ω ist eine Funktion $\mathrm{m} : 2^\Omega \rightarrow [0,1]$, die den folgenden beiden Bedingungen genügt:

$$\mathrm{m}(\emptyset) = 0 \text{ und}$$

$$\sum_{A \subseteq \Omega} \mathrm{m}(A) = 1$$

$\mathrm{m}(A)$ beinhaltet das Maß an Glauben daran, dass die aktuelle Welt exakt zu A und nicht etwa zu einer Teilmenge von A gehört. Diese beiden Bedingungen stellen sicher, dass das unmögliche Ereignis bzw. die leere Menge mit einem Glaubensmaß von 0, also gar nicht, geglaubt wird, und dass das Gesamtmaß des Glaubens 1 beträgt.

Definition 8.1.2 (fokales Element)
Ist $\mathrm{m} : 2^\Omega \rightarrow [0,1]$ ein Basismaß, dann heisst jede Teilmenge $A \subseteq \Omega$ mit $\mathrm{m}(A) > 0$ ein fokales Element von m.

Aus vorhandenen Basismaßen kann man durch Anwendung der folgenden Definitionen Glaubens- und Plausibilitätsfunktionen berechnen.

Definition 8.1.3 (Belief- bzw. Glaubensfunktion)
Sei m ein Basismaß über Ω. Die durch m induzierte Glaubensfunktion, bel, wird definiert durch

$$\mathrm{bel} : 2^{\Omega} \longrightarrow [0,1],\ \ \mathrm{bel}(A) := \sum_{B \subseteq A} \mathrm{m}(B)$$

Der Term bel(A) repräsentiert den Grad an Glauben bzw. notwendiger Unterstützung dafür, dass die aktuelle Welt zu A gehört.

Definition 8.1.4 (Plausibilitätsfunktion)
Sei m ein Basismaß über Ω. Die durch m induzierte Plausibilitätsfunktion, pl, wird definiert durch

$$\mathrm{pl} : 2^{\Omega} \longrightarrow [0,1],\ \ \mathrm{pl}(A) := \sum_{B \cap A \neq \emptyset} \mathrm{m}(B)$$

Der Term pl(A) quantifiziert den Grad an Plausibilität, d.h. möglicher Unterstützung dafür, dass die aktuelle Welt zu A gehört. Er enthält diejenigen Basismaße, die man Eigenschaften oder Aussagen zuschreibt, die A nicht widersprechen, denn für alle B, die A widersprechen, gilt $A \cap B = \emptyset$. Es lässt sich zeigen:

Satz 8.1.1 (Dualität von Belief- und Plausibilitätsfunktionen)
Es gilt $\mathrm{bel}(A) = 1 - \mathrm{pl}(\overline{A})$ *und* $\mathrm{pl}(A) = 1 - \mathrm{bel}(\overline{A})$.

Wenn zwei Funktionen diese Eigenschaften haben, dann nennt man diese Funktionen *dual* zueinander.

BEWEIS:

$$1 - \mathrm{bel}(\overline{A}) = 1 - \sum_{B \subseteq \overline{A}} \mathrm{m}(B) = \sum_{B \subseteq \Omega} \mathrm{m}(B) - \sum_{B \subseteq \overline{A}} \mathrm{m}(B) = \sum_{B \cap A \neq \emptyset} \mathrm{m}(B) = \mathrm{pl}(A)$$

$$\begin{aligned}1-\mathrm{pl}(\overline{A}) &= 1-\sum_{B\cap\overline{A}\neq\emptyset}\mathrm{m}(B)=\sum_{B\subseteq\Omega}\mathrm{m}(B)-(\sum_{B\subseteq\Omega}\mathrm{m}(B)-\sum_{B\subseteq A}\mathrm{m}(B))\\ &= \sum_{B\subseteq A}\mathrm{m}(B)=\mathrm{bel}(A)\end{aligned}$$

$\square$

8.2 Kombinationsregel von Dempster

Die Dempster'sche Kombinationsregel ermöglicht es, verschiedene Anhaltspunkte, die auch *Evidenzen* genannt werden und durch Basismaße ausgedrückt werden können, zu kombinieren.

Definition 8.2.1 ($\mathrm{m}_\oplus$: Kombinationsregel von Dempster)
Seien $\mathrm{m}_1,\ldots,\mathrm{m}_n$ verschiedene Basismaße. Die Kombinationsfunktion $\mathrm{m}_\oplus$ dieser einzelnen Basismaße erhält man durch die Regel

$$\mathrm{m}_\oplus(F)=\begin{cases}\frac{\sum_{X_1\cap\ldots\cap X_n=F}\prod_{i=1}^n \mathrm{m}_i(X_i)}{\sum_{X_1\cap\ldots\cap X_n\neq\emptyset}\prod_{i=1}^n \mathrm{m}_i(X_i)}, & \text{wenn } \emptyset\neq F\subseteq\Omega\\ 0, & \text{wenn } F=\emptyset\end{cases}\tag{8.1}$$

Wegen

$$\sum_{X_1\cap\ldots\cap X_n\neq\emptyset}\prod_{i=1}^n \mathrm{m}_i(X_i)=1-\sum_{X_1\cap\ldots\cap X_n=\emptyset}\prod_{i=1}^n \mathrm{m}_i(X_i)\tag{8.2}$$

wird der Nenner der Dempster'schen Kombinationsregel in Gleichung (8.1) manchmal auch durch die rechte Seite der Gleichung (8.2) beschrieben. Diesen Nenner nennt man *Normalisierungskonstante* c. Gleichung (8.2) ergibt sich aus der Tatsache, dass $\mathrm{m}_\oplus$ ein Basismaß ist, was in Satz 8.2.1 bewiesen wird:

Satz 8.2.1 (Basismaß $\mathrm{m}_\oplus$)
Alle $\mathrm{m}_\oplus$, die aus der Dempsterschen Kombinationsregel gewonnen wurden, sind Basismaße.

BEWEIS: Es ist nach Definition 8.1.1 zu zeigen, dass

$$\sum_{Y_i \in \Omega} \mathrm{m}_{\oplus}(Y_i) = 1$$

und

$$\mathrm{m}_{\oplus}(\emptyset) = 0$$

$\mathrm{m}_{\oplus}(\emptyset) = 0$ gilt laut Definition 8.2.1. Alle Mengen F, für die $\mathrm{m}_{\oplus}(F) > 0$ gilt, werden laut Definition 8.1.2 fokale Elemente von $\mathrm{m}_{\oplus}$ genannt. Seien F_{ik} jeweils die fokalen Elemente X_{ik} der m_i, $k = 1, \ldots, |F_i|$. An Gleichung (8.1) sieht man, dass auch nur diese fokalen Elemente $X_{ik} := F_{ik}$ und $X_{i\Omega} := \Omega$ zu den fokalen Elementen von $\mathrm{m}_{\oplus}$ beitragen, denn sobald für ein X_{jk} gilt: $X_{jk} \neq \Omega$ und $X_{jk} \neq F_{jk}$, ergibt sich $\mathrm{m}_j(X_{jk}) = 0$ und damit auch $\prod_{i=1}^{n} \mathrm{m}_i(X_i) = \prod_{i \neq j} \mathrm{m}_i(X_i) \cdot \mathrm{m}_j(X_{jk}) = 0$. Deshalb folgt (es wird hier zunächst vorausgesetzt, dass die Normalisierungskonstante c=1 ist, dass also $F_{1k} \cap \ldots \cap F_{nl} \neq \emptyset$ für alle fokalen Elemente F_{ik} aller m_i):

$$\begin{aligned}
\sum_{Y_i \in \Omega} \mathrm{m}_{\oplus}(Y_i) &= \sum_{\substack{X_1 \cap \ldots \cap X_n = Y_i \\ X_i \in \{F_{ik}, k=1,\ldots,|F_i|\}}} \prod_{i=1}^{n} \mathrm{m}_i(X_i) \\
&= \mathbf{m_1(\boldsymbol{F_{11}})} \sum \prod_{\substack{\mathbf{i=2} \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i) + \ldots + \mathbf{m_1(\boldsymbol{F_{1l}})} \sum \prod_{\substack{\mathbf{i=2} \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i) \\
& \qquad l = |F_1| \\
&= \sum_{\mathbf{k=1}}^{\mathbf{|F_1|}} \mathbf{m_1(\boldsymbol{F_{1k}})} \sum \prod_{\substack{\mathbf{i=2} \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i) \\
&\overset{Def 8.1.1}{=} \mathbf{1\cdot} \sum \prod_{\substack{i=2 \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i) \\
&= \mathbf{m_2(\boldsymbol{F_{21}})} \sum \prod_{\substack{\mathbf{i=3} \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i) + \ldots + \mathbf{m_2(\boldsymbol{F_{2m}})} \sum \prod_{\substack{\mathbf{i=3} \\ X_i \in \{F_i\}}}^{n} \mathrm{m}_i(X_i)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{|F_2|} m_1(F_{2k}) \sum_{X_i \in \{F_i\}} \prod_{i=3}^{n} \mathrm{m}_i(X_i) \\
&\overset{Def 8.1.1}{=} 1 \cdot \sum_{X_i \in \{F_i\}} \prod_{i=3}^{n} \mathrm{m}_i(X_i) \\
&\quad \vdots \\
&= \\
&= \sum_{k=1}^{|F_n|} m_n(F_{nk}) \\
&\overset{Def 8.1.1}{=} 1
\end{aligned}
$$

Falls die Normalisierungskonstante c doch nicht gleich 1 ist, falls es also Evidenzen $F_{1k}, \ldots, F_{nl}$ gibt mit $F_{1k} \cap \ldots \cap F_{nl} = \emptyset$, dann muss man $\sum_{F_{1k} \cap \ldots \cap F_{nl} = \emptyset} \mathrm{m}_1(F_{1k}) \cdot \ldots \cdot \mathrm{m}_n(F_{nl})$ wieder von dem obigen Ergebnis 1 abziehen und erhält damit im Zähler von Gleichung (8.1) die Normalisierungskonstante $c = 1 - \sum_{X_1 \cap \ldots \cap X_n = \emptyset} \prod_{i=1}^{n} \mathrm{m}_i(X_i)$, also $\sum\limits_{Y_i \in \Omega} \mathrm{m}_{\oplus}(Y_i) = \frac{c}{c} = 1$. Deshalb ist auch $\mathrm{m}_{\oplus}$ ein Basismaß. □

Belieffunktionen, die aus Basismaßen induziert wurden, die höchstens ein fokales Element $F \neq \Omega$, $F \neq \emptyset$, haben, nennt man *simple-support-functions*.

Definition 8.2.2 (Simple-support-Funktionen)
Simple-support-Funktionen sind Belief-Funktionen, die aus Basisfunktionen m mit

$$
\mathrm{m}(\Omega) = s, \;\; \mathrm{m}(F) = 1 - s \text{ und } \mathrm{m}(X) = 0 \text{ sonst}, \; s \in [0, 1] \tag{8.3}
$$

induziert wurden

Seien nun $F_1, \ldots, F_n$ verschiedene Evidenzen, denen jeweils eine simple-support-function m_i zugeordnet wurde, also $\mathrm{m}_i(\Omega) = s_i$, $\mathrm{m}_i(F_i) = 1 - s_i$, $\mathrm{m}_i(X_i) = 0$ sonst. Diese Evidenzen werden durch die Dempster'sche Kombinationsregel zu einem Basismaß $\mathrm{m}_\oplus$ kombiniert. Es soll nun gezeigt werden, dass für Plausibilitätsfunktionen, die man aus solchen simple-support-functions mit der Dempster´schen Kombinationsregel erhält, gilt

Satz 8.2.2 (Berechnung der Plausibilität einer Welt ω)
Für alle Plausibilitätsfunktionen, die man aus simple-support-functions mit der Dempster'schen Kombinationsregel erhält, gilt[1]

$$\mathrm{pl}_\oplus(\omega) = \prod_{\omega \not\models F_i} s_i \tag{8.4}$$

Beweis:

$$\begin{aligned}\mathrm{pl}_\oplus(\omega) &\overset{Def8.1.4}{=} \sum_{\omega\cap[F]\neq\emptyset} \mathrm{m}_\oplus(F) \overset{Def8.2.1}{=} \sum_{(\omega\cap[F]\neq\emptyset)} \sum_{([X_1]\cap\cdots\cap[X_n]=[F])} \prod_{i=1}^{n} \mathrm{m}_i(X_i) \\ &= \sum_{\omega\in([X_1]\cap\cdots\cap[X_n])} \prod_{i=1}^{n} \mathrm{m}_i(X_i) \end{aligned} \tag{8.5}$$

Die m_i sind simple-support-functions mit $\mathrm{m}_i(\Omega) = s_i$, $\mathrm{m}_i([F_i]) = 1 - s_i$ und $\mathrm{m}_i(X_i) = 0$ sonst. Alle X_i können nur die zwei Werte $[F_i]$ und Ω annehmen, da sonst

$$\mathrm{m}_\oplus(X_1 \cap \cdots \cap X_n) = \mathrm{m}_j(X_j) \cdot \prod_{i\neq j} \mathrm{m}_i(X_i) = 0 \;\forall\; F_j \neq X_j \neq \Omega$$
$$\text{wegen } \mathrm{m}_j(X_j) = 0 \text{ für alle } X_j \neq \Omega \text{ und } X_j \neq [F_j]$$

Falls es nur eine Evidenz F_1 gibt, ist die Behauptung klar:
n=1
Es gilt nach (8.5)

$$\mathrm{pl}_\oplus(\omega) = \sum_{(\omega\in X_1)} \prod_{i=1}^{1} \mathrm{m}_1(X_1) = \sum_{\omega\in X_1} \mathrm{m}_1(X_1) \tag{8.6}$$

[1] vgl.[8] 3.1

Da $\omega \in \Omega$ immer gilt, gibt es die zwei Möglichkeiten

$$\omega \in [F_1] \text{ oder } \omega \notin [F_1]$$

- **1.Fall:** Es gelte $\omega \in [F_1]$. Da $\omega \in \Omega$ immer gilt, folgt nach (8.6):

$$\mathrm{pl}_{\oplus}(\omega) = \mathrm{m}_1(\Omega) + \mathrm{m}_1(F_1) \stackrel{(Def.8.2.2)}{=} s_1 + (1 - s_1) = 1$$

- **2.Fall:** Es gelte $\omega \notin [F_1]$. Dann folgt aus $\omega \in \Omega$ mit (8.6):

$$\mathrm{pl}_{\oplus}(\omega) = \mathrm{m}_1(\Omega) = s_1$$

Daraus folgt die Behauptung $\mathrm{pl}_{\oplus}(\omega) = 1$, falls $\omega \models F_1$, und $\mathrm{pl}_{\oplus}(\omega) = s_1$, falls $\omega \not\models F_1$.
Dieser Beweis muss nun auf eine beliebige Menge von Evidenzen $F_1, \ldots, F_n$ erweitert werden:
$\boldsymbol{n > 1}$:
Nach (8.5) gilt:

$$\mathrm{pl}_{\oplus}(\omega) = \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{i=1}^{n} m_i(X_i),\ \ X_i \in \{F_i, \Omega\}$$

Für alle i, $1 \leq i \leq n$, gibt es wieder die zwei Möglichkeiten $\omega \models F_i$ und $\omega \not\models F_i$. Man kann daher zwei Mengen

$$Y^+(\omega) := \{k |\ \omega \models F_k\}$$

und

$$Y^-(\omega) := \{k' |\ \omega \not\models F_k\}$$

bilden. Da klar ist, dass diese Mengen von ω abhängen, wird $Y^+(\omega)$ im folgenden Text durch Y^+ und $Y^-(\omega)$ durch Y^- abgekürzt. Da es auf die Reihenfolge der X_i in der Schnittmenge $X_1 \cap \ldots \cap X_n$ nicht ankommt, kann man die X_i so umsortieren, dass alle X_i, für die $\omega \models F_i$ gilt, vorne stehen:

$$\mathrm{pl}_{\oplus}(\omega) = \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+}}^{|Y^+|} m_i(X_i) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} m_i(X_i), \; X_i \in \{F_i, \Omega\}$$

Für alle $i \in Y^+$ gilt: $\omega \in [F_i]$ und $\omega \in \Omega$, und für alle $i \in Y^-$ gilt nur $\omega \in \Omega$. Für alle beliebigen $j \in Y^+$ gilt:

$$\begin{aligned}
\mathrm{pl}_{\oplus}(\omega) &= \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+, i \neq j}}^{|Y^+|} \mathrm{m}_i(X_i) \cdot \mathbf{m}_{\boldsymbol{j}}(\boldsymbol{F_j}) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} \mathrm{m}_i(X_i) \\
&+ \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+, i \neq j}}^{|Y^+|} \mathrm{m}_i(X_i) \cdot \mathbf{m}_{\boldsymbol{j}}(\boldsymbol{\Omega}) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} \mathrm{m}_i(X_i), \; X_i \in \{F_i, \Omega\} \\
&= \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+, i \neq j}}^{|Y^+|} \mathrm{m}_i(X_i) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} \mathrm{m}_i(X_i) \cdot (\mathbf{m}_{\boldsymbol{j}}(\boldsymbol{\Omega}) + \mathbf{m}_{\boldsymbol{j}}(\boldsymbol{F_j})) \\
&= \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+, i \neq j}}^{|Y^+|} \mathrm{m}_i(X_i) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} \mathrm{m}_i(X_i) \cdot (\boldsymbol{s_j} + (\mathbf{1} - \boldsymbol{s_j})) \\
&= \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^+, i \neq j}}^{|Y^+|} \mathrm{m}_i(X_i) \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} \mathrm{m}_i(X_i)
\end{aligned}$$

Da dieser Nachweis für beliebige $j \in Y^+$ fortgesetzt werden kann, bleiben nur noch alle $i \in Y^-$ übrig:

$$\mathrm{pl}_{\oplus}(\omega) = \sum_{\omega \in ([X_1] \cap \cdots \cap [X_n])} \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} m_i(X_i)$$

Da für alle $i \in Y^-$ gilt $\omega \not\models F_i$, kann man für diese X_i nur Ω einsetzen:

$$\mathrm{pl}_{\oplus}(\omega) = \sum_{\omega \in ([X_1] \cap \cdots \cap [X_k] \cap \Omega \cap \ldots \cap \Omega)} \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} m_i(\Omega)$$

$$= \sum_{i \in Y^-, X_i = \Omega} \prod_{\substack{i=1 \\ i \in Y^-}}^{|Y^-|} s_i$$

$$= \prod_{\omega \not\models F_i} s_i$$

□

In [2] werden in Kapitel 2.3 folgende Eigenschaften der Dempster-Shafer-Theorie erwähnt, die das Verhalten der Belief- und Plausibilitätsfunktionen nach Erweiterung des Wissens beschreiben. Enthält das vorhanden Wissen z.B. eine Formel X und wird das Wissen durch eine neue Evidenz erweitert, die darauf hinweist, dass die aktuelle Welt zu A gehört, dann unterstützt das Basis-Belief-Maß, das vorher ein Maß für den Glauben an X war, nun $X \cap A$. Das Belief-Maß $\text{bel}(A \to X)$ ist der Wert für den Glauben an X, nachdem das bereits bekannte Basismaß m an die neue Evidenz A angepasst wurde. $\text{bel}(A \to X)$ ist ein Maß für den Glauben an X, wenn man A sicher weiss. $\text{bel}(A \to X)$ beschreibt also den Wert, den man dem Glauben an den Default $A \to X$ zuordnet, und $\text{bel}(A \to X)$ muss nicht gleich $\text{bel}(\neg A \vee X)$ sein.

Für Belief- und Plausibilitätsfunktion nach der Erweiterung des Wissens um die neue Evidenz A gilt:

Satz 8.2.3 (Konditionierungsregel von Dempster)

$$\text{bel}(A \to X) = \frac{\text{bel}(X \cup \overline{A}) - \text{bel}(\overline{A})}{1 - \text{bel}(\overline{A})} \tag{8.7}$$

$$\text{pl}(A \to X) = \frac{\text{pl}(X \cap A)}{\text{pl}(A)} \tag{8.8}$$

Beweis: Das Beliefmaß $\text{bel}(A \to X)$ ist ein Maß für den Glauben an X unter der Voraussetzung, dass man A schon sicher weiss. Diese sichere Evidenz für

A kann man durch ein Maß m_A ausdrücken, für das gilt[2]:

$$\mathrm{m}_A(Y) = \begin{cases} 1, & \text{falls } Y = A \\ 0, & \text{falls } Y \neq A \end{cases}$$

Das einzige fokale Element von m_A ist A. Wenn man das vorhandene Basismaß m_{alt} nun durch die Dempster'sche Kombinationsregel mit diesem neuen Maß m_A kombiniert, dann erhält man:

$$\mathrm{m}_{\oplus}(Y) = \frac{\sum\limits_{D,\ A\cap D=Y} \mathrm{m}_{alt}(D)\cdot \mathrm{m}_A(A)}{1-\sum\limits_{A\cap D=\emptyset} \mathrm{m}_{alt}(D) m_A(A)} = \frac{\sum\limits_{D,\ A\cap D=Y} \mathrm{m}_{alt}(D)}{1-\mathrm{bel}(\overline{A})} \tag{8.9}$$

Das Maß bel für den Glauben an X unter diesem Basismaß ist dann:

$$\begin{aligned}
\mathrm{bel}(A\to X) &= \sum_{\emptyset\neq C\subset X} \mathrm{m}_{\oplus}(C) \overset{(8.9)}{=} \frac{\sum\limits_{C,\ C\subset X}\ \sum\limits_{D, A\cap D=C} \mathrm{m}_{alt}(D)}{1-\mathrm{bel}(\overline{A})} \\
&= \frac{\sum\limits_{\emptyset\neq A\cap D\subset X} \mathrm{m}_{alt}(D)}{1-\mathrm{bel}(\overline{A})} = \frac{\sum\limits_{D\not\subset\overline{A},\ D\subset X\cup\overline{A}} \mathrm{m}_{alt}(D)}{1-\mathrm{bel}(\overline{A})} \\
&= \frac{\mathrm{bel}(X\cup\overline{A})-\mathrm{bel}(\overline{A})}{1-\mathrm{bel}(\overline{A})}
\end{aligned}$$

Daraus ergibt sich für die Plausibilität:

$$\begin{aligned}
\mathrm{pl}(A\to X) &= 1-\mathrm{bel}(A\to\overline{X}) \\
&\overset{(8.7)}{=} \frac{1-\mathrm{bel}(\overline{A})}{1-\mathrm{bel}(\overline{A})} - \frac{\mathrm{bel}(\overline{A}\cup\overline{X})-\mathrm{bel}(\overline{A})}{1-\mathrm{bel}(\overline{A})} \\
&= \frac{1-\mathrm{bel}(\overline{A})-\mathrm{bel}(\overline{A}\cup\overline{X})+\mathrm{bel}(\overline{A})}{1-\mathrm{bel}(\overline{A})} \\
&= \frac{1-\mathrm{bel}(\overline{A\cap X})}{1-\mathrm{bel}(\overline{A})} \\
&= \frac{\mathrm{pl}(A\cap X)}{\mathrm{pl}(A)}
\end{aligned}$$

□

[2](vgl.[9] Theorem 3.6)

8.3 Dempster'sche Kombinationsregel und konditionale Indifferenz

Es wird nun gezeigt, dass alle Plausibilitäts- und Belieffunktionen, die man aus simple-support-functions mit der Dempster'schen Kombinationsregel erhält, konditional indifferent sind. Die Definition der konditionalen Indifferenz für Wahrscheinlichkeitsfunktionen und OCFs ist in Definition 3.0.10 auf S.27 angegeben.

Man muss diese Definition zunächst auf die zu untersuchenden Plausibilitäts- und Belieffunktionen übertragen.

8.3.1 Plausibilitätsfunktionen und konditionale Indifferenz

Man kann die konditionale Indifferenz für Plausibilitätsfunktionen analog zu den Bedingungen der konditionalen Indifferenz für Wahrscheinlichkeitsfunktionen und OCFs definieren, indem man fordert, dass aus $\sigma(\hat{\omega}) = 1$ folgt, dass auch $\mathrm{pl}(\hat{\omega}) = 1$ gelten muss. Das bedeutet

Definition 8.3.1 (Konditionale Indifferenz von Plausifunktionen)
Eine Plausibilitätsfunktion $\mathrm{pl}_{\oplus}$ ist konditional indifferent in bezug auf eine Defaultmenge Δ genau dann, wenn $\mathrm{pl}_{\oplus}(\hat{\omega}_1) = \mathrm{pl}_{\oplus}(\hat{\omega}_2)$ immer dann, wenn $\sigma_{\Delta}(\hat{\omega}_1) = \sigma_{\Delta}(\hat{\omega}_2)$ für $\hat{\omega}_1 \equiv_{\top} \hat{\omega}_1 \in \hat{\Omega}$.

Um die Kriterien dieser Definition anwenden zu können, muss man noch festlegen, was $\mathrm{pl}_{\oplus}(\hat{\omega})$ bedeutet. Man kann analog zu Definition 3.0.8 mit $\hat{\Omega}^{\times} := \{\hat{\omega} \in \hat{\Omega} \mid \mathrm{pl}_{\oplus}(\hat{\omega}) \neq 0\}$ festlegen:

Definition 8.3.2 (Fortsetzung von pl auf $\hat{\Omega}$)
Eine Plausibilitätsfunktion $\mathrm{pl}_\oplus : \Omega \to [0,1]$ wird zu einem Homomorphismus $\mathrm{pl}_\oplus : \hat{\Omega}^\times \to \mathbb{R}^+$ fortgesetzt durch

$$\mathrm{pl}_\oplus(\hat{\omega}) = \mathrm{pl}_\oplus(\omega_1^{r_1} \cdots \omega_r^{r_n}) := \mathrm{pl}_\oplus(\omega_1)^{r_1} \ldots \mathrm{pl}_\oplus(\omega_k)^{r_k}$$

$((\mathbb{R}^+\backslash 0), \cdot)$ ist eine Gruppe. Man muss die Division durch 0 zunächst verbieten.

Satz 8.3.1 (Plausibilitätsfunktionen und konditionale Indifferenz)
Alle Plausibilitätsfunktionen $\mathrm{pl}_\oplus$, die aus simple-support-Funktionen mit der Dempster'schen Kombinationsregel gebildet wurden, erfüllen die Kriterien von Definition 8.3.1 und sind damit konditional indifferent bezüglich der zugrundeliegenden Defaultmenge Δ.

Beweis: Die kombinierte Plausibilitätsfunktion erhält man nach Lemma 9.2.1 aus simple-support-functions, die in Definition 8.2.2 auf Seite 72 definiert wurden.

Sei $\mathrm{pl}_\oplus$ eine Plausibilitätsfunktion. Dann gilt für alle Welten $\omega \in \Omega$

$$\mathrm{pl}_\oplus(\omega) = \prod_{\omega \not\models F_i} s_i \ (8.4)$$

Die Evidenz F_i entspricht hier dem i-ten Default $F_i = d_i = A_i \to B_i$. Wegen

$$\sigma_\Delta(\omega) \overset{Def.3.0.5}{=} \prod_{1 \leqslant i \leqslant n} \sigma_i(\omega) = \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} a_i^+ \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} a_i^- \ (3.1)$$

$$\text{und } \mathrm{pl}_\oplus(\omega) = \prod_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} s_i \ (8.4)$$

erhält man die Plausibilität einer Interpretation ω aus der konditionalen Struktur, indem man z.B. über eine Funktion $V_\oplus$ alle $a_i^+ := 1$ setzt und allen a_i^- die entsprechenden s_i zuweist:

$$V_\oplus(a_i^+) := 1 \text{ und } V_\oplus(a_i^-) := s_i \text{ und } V_\oplus(1) := 1 \tag{8.10}$$

Falls ω mindestens einen Default d_i falsifiziert, dann gilt:

$$\begin{aligned} V_\oplus(\sigma_\Delta(\omega)) = V_\oplus(\prod_{\omega \models A_iB_i} a_i^+ \prod_{\omega \models A_i\overline{B}_i} a_i^-) &\overset{(8.10)}{=} \prod_{\omega \models A_iB_i} 1 \cdot \prod_{\omega \models A_i\overline{B}_i} s_i \\ = \prod_{\omega \models A_i\overline{B}_i} s_i \overset{(8.4)}{=} \mathrm{pl}_\oplus(\omega) \end{aligned}$$

Falls ω keinen Default falsifiziert, dann gilt entweder

$$\begin{aligned} \sigma_\Delta(\omega) = \prod_{\omega \models A_iB_i} a_i^+ \text{ und } \mathrm{pl}_\oplus(\omega) &\overset{(8.4)}{=} 1 \overset{(8.10)}{=} V_\oplus(\sigma_\Delta(\omega)) \\ \text{oder } \sigma_\Delta(\omega) \overset{(3.0.5)}{=} 1, \text{ falls } \omega \models \neg A_i &\text{ für alle } i = 1, \ldots n \\ \Rightarrow V_\oplus(\sigma_\Delta(\omega)) &\overset{(8.10)}{=} 1 \overset{(8.4)}{=} \mathrm{pl}_\oplus(\omega) \end{aligned}$$

$$\text{Also gilt für alle } \omega \in \Omega : \mathrm{pl}_\oplus(\omega) = V_\oplus(\sigma_\Delta(\omega)) \tag{8.11}$$

Daraus folgt

$$\forall \hat{\omega} \in \hat{\Omega} : V_\oplus(\sigma_\Delta(\hat{\omega})) = \mathrm{pl}_\oplus(\hat{\omega}), \text{ denn:}$$

$$\begin{aligned} &\forall \hat{\omega} \in \hat{\Omega}, \hat{\omega} = \omega_1^{r_1} \cdot \ldots \cdot \omega_n^{r_n} : \\ &\qquad V_\oplus(\sigma_\Delta(\hat{\omega})) \overset{(3.2)}{=} V_\oplus(\sigma_\Delta(\omega_1)^{r_1}) \cdot \ldots \cdot \sigma_\Delta(\omega_n)^{r_n}) \\ &\qquad \overset{(8.10)}{=} V_\oplus(\sigma_\Delta(\omega_1))^{r_1} \cdot \ldots \cdot V_\oplus(\sigma_\Delta(\omega_n))^{r_n} \\ &\qquad \overset{(8.11)}{=} \mathrm{pl}_\oplus(\omega_1)^{r_1} \cdot \ldots \cdot \mathrm{pl}_\oplus(\omega_n)^{r_n} \overset{(Def.3.0.8)}{=} \mathrm{pl}_\oplus(\hat{\omega}) \end{aligned}$$

Es gelte also $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$, also $\frac{\sigma_\Delta(\hat{\omega}_1)}{\sigma_\Delta(\hat{\omega}_2)} = 1$. Dann folgt aus

$$1 = \frac{\sigma_\Delta(\hat{\omega}_1)}{\sigma_\Delta(\hat{\omega}_2)} = \frac{(a_1^+)^{r_1}(a_1^-)^{r_2} \cdots (a_n^+)^{r_{2n-1}}(a_n^-)^{r_{2n}}}{(a_1^+)^{t_1}(a_1^-)^{t_2} \cdots (a_n^+)^{t_{2n-1}}(a_n^-)^{t_{2n}}}$$

$$\Rightarrow r_i = t_i \text{ für alle } i = 1, \ldots 2n \tag{8.12}$$

und damit, falls $s_i \neq 0$ für alle $t_{2i} > 0$

$$\begin{aligned} \frac{\mathrm{pl}_\oplus(\hat{\omega}_1)}{\mathrm{pl}_\oplus(\hat{\omega}_2)} &= \frac{V_\oplus(\sigma_\Delta(\hat{\omega}_1))}{V_\oplus(\sigma_\Delta(\hat{\omega}_2))} = \frac{s_1^{r_2} \cdots s_n^{r_{2n}}}{s_1^{t_2} \cdots s_n^{t_{2n}}} = 1 \text{ wegen } (8.12) \\ \Rightarrow \mathrm{pl}_\oplus(\hat{\omega}_1) &= \mathrm{pl}_\oplus(\hat{\omega}_2) \end{aligned}$$

und falls für ein $t_{2i} > 0$ ein $s_i = 0$ ist, $1 \leq i \leq n$, dann gilt

$$\begin{aligned} \mathrm{pl}_\oplus(\hat{\omega}_1) &= s_1^{r_2} \cdot \ldots \cdot s_n^{r_{2n}} = 0 = s_1^{t_2} \cdot \ldots \cdot s_n^{t_{2n}} = \mathrm{pl}_\oplus(\hat{\omega}_2) \\ &\Rightarrow \mathrm{pl}_\oplus(\hat{\omega}_1) = \mathrm{pl}_\oplus(\hat{\omega}_2) \end{aligned}$$

Insgesamt folgt: falls

$$\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2) \text{ für } \omega_1 \equiv_\top \omega_2$$

dann gilt auch

$$\mathrm{pl}_\oplus(\hat{\omega}_1) = pl_\oplus(\hat{\omega}_2)$$

und damit erfüllen diese Plausibilitätsfunktionen die Kriterien aus Definition 8.3.1 und sind daher konditional indifferent. □

8.3.2 Belief-Funktionen und konditionale Indifferenz

Es soll nun gezeigt werden, dass auch die entsprechenden Belieffunktionen, die man aus der Dempster'schen Kombinationsregel erhält, die Eigenschaft der konditionalen Indifferenz haben. Dazu muss man die Definition der konditionalen Indifferenz zunächst auf Belief-Funktionen übertragen.

Definition 8.3.3 (Konditionale Indifferenz von Belieffunktionen)
Sei $\Delta = \{(A_i \rightarrow B_i)\} \subseteq (\mathcal{L} \rightarrow \mathcal{L})$ eine Menge von Defaults. Eine Belieffunktion $\mathrm{bel}_\oplus$ ist indifferent in bezug auf Δ genau dann, wenn $\mathrm{bel}_\oplus(\hat{\omega}_1) = \mathrm{bel}_\oplus(\hat{\omega}_2)$ immer dann, wenn $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$ für $\hat{\omega}_1 \equiv_\top \hat{\omega}_2$

Man muss auch hier wieder definieren, was $\mathrm{bel}_\oplus(\hat{\omega})$ bedeutet. Man kann wieder analog zu Definition 3.0.8 mit $\hat{\Omega}^\times := \{\hat{\omega} \in \hat{\Omega} \mid \mathrm{bel}_\oplus(\hat{\omega}) \neq 0\}$ festlegen:

Definition 8.3.4 (Fortsetzung von bel auf $\hat{\Omega}$)
Eine Belieffunktion $\mathrm{bel}_{\oplus} : \Omega \to [0,1]$ wird zu einem Homomorphismus $\mathrm{bel}_{\oplus} : \hat{\Omega}^{\times} \to \mathbb{R}^{+}$ fortgesetzt durch

$$\mathrm{bel}_{\oplus}(\hat{\omega}) = \mathrm{bel}_{\oplus}(\omega_1^{r_1} \cdots \omega_r^{r_n}) := \mathrm{bel}_{\oplus}(\omega_1)^{r_1} \ldots \mathrm{bel}_{\oplus}(\omega_k)^{r_k}$$

Wenn man die Belieffunktionen wie in dem BSS-Ansatz aus simple-support-functions mit der Dempster´schen Kombinationsregel gewinnt, gilt

$$\mathrm{bel}_{\oplus}(\omega) \overset{(Def.8.1.3)}{=} \sum_{B:B\subseteq\omega} \mathrm{m}_{\oplus}(B) = \sum_{F_{\oplus}:F_{\oplus}\subseteq\omega} \prod_{X_1\cap\ldots\cap X_n=F_{\oplus}} \mathrm{m}_i(X_i)$$

In der Regel besteht die Schnittmenge $F_{\oplus} = X_1 \cap \ldots \cap X_n$ der fokalen Elemente der m_i nicht aus einer einzigen Welt ω, sondern aus einer Vereinigung von Welten, so dass man einzelnen Welten oft nur den Belief-Wert 0 zuordnen kann.

Beispiel 8.3.1 (Belief-Wert 0)
In dem obigen Beispiel 9.2.3 auf Seite 101 muss man allen einzelnen Welten den Belief-Wert 0 zuordnen. Zum Beispiel genügt die Welt $\omega_1 := v \wedge \neg p \wedge f$ *allen drei Defaults, aber da es kein fokales Element* $F_{\oplus}$ *von* $\mathrm{m}_{\oplus}$ *mit* $F_{\oplus} \subseteq \{\omega_1\}$ *gibt, gilt* $\mathrm{bel}_{\oplus}(\omega_1) = 0$*. Das „kleinste" fokale Element* F_1 *von* $\mathrm{m}_{\oplus}$ *ist* $F_1 := [\neg v \vee f] \cap [\neg p \vee v] \cap [\neg p \vee \neg f] = [v \wedge \neg p \wedge f] \cup [\neg v \wedge \neg p \wedge f] \cup [\neg v \wedge \neg p \wedge \neg f] \not\subseteq \{\omega_1\}$

In diesem Fall kann man die Belieffunktion als konditional indifferent bezeichnen, da der Beliefwert aller Welten unabhängig von der konditionalen Struktur immer gleich 0 ist: $\mathrm{bel}_{\oplus}(\hat{\omega}) = \mathrm{bel}_{\oplus}(\hat{\omega}') = 0$ immer dann, wenn $\sigma_{\Delta}(\hat{\omega}) = \sigma_{\Delta}(\hat{\omega'})$

Beispiel 8.3.2 (Eindeutigkeit der Belieffunktion)
Man kann in dem Beispiel nur der Vereinigung aller drei Welten, die allen Defaults genügen und zu denen ω_1 gehört, den Belief-Wert $(1-\varepsilon_1)\cdot(1-\varepsilon_2)\cdot(1-\varepsilon_3)$ zuordnen. Dieser Belief-Wert entspricht dem "Glaubensmaß", dass alle drei Defaults gleichzeitig gelten.

Um in dem obigen Beispiel einer einzelnen Welt einen Belief-Wert $\neq 0$ zuzuordnen, müsste man z.B. den Fakt „v", den man ja auch als Regel $d_4 := \top \rightarrow v$ ausdrücken kann, zu den Defaults hinzufügen. Mit dieser neuen Regel hat die Welt $\omega_1 := v \wedge \neg p \wedge f = [\varphi_1] \wedge [\varphi_2] \wedge [\varphi_3] \wedge [\varphi_4]$ den Beliefwert $\mathrm{bel}(\omega_1) = (1-s_1)\cdot(1-s_2)\cdot(1-s_3)\cdot(1-s_4)$, und alle übrigen Welten erhalten den Beliefwert 0.

Wenn man den Fakt „p" hinzufügen würde, dann müsste man z.B. die Regel $v \rightarrow f$ entfernen, um einer einzelnen Welt $\omega_2 := p \wedge v \wedge \neg f$ einen positiven Beliefwert zuordnen zu können.

Satz 8.3.2 (Eindeutigkeit der Belieffunktion)
Es gibt für alle Belieffunktionen $\mathrm{bel}_\oplus$ höchstens eine Welt $\omega_{\mathrm{bel}} \in \Omega$ mit $\mathrm{bel}_\oplus(\omega_{\mathrm{bel}}) \neq 0$.

BEWEIS: Die fokalen Elemente F_i von $\mathrm{m}_\oplus$ sind partiell geschachtelt. Sei F_{bel} dasjenige fokale Element von $\mathrm{m}_\oplus$, das keinen Default falsifiziert. Wegen $[\varphi_i] \subseteq \Omega$ gilt

$$F_{\mathrm{bel}} := \bigcap_{i=1}^{n} [\varphi_i] \;\subseteq\; [\varphi_1] \cap \ldots \cap \Omega \cap \ldots \cap [\varphi_n] \text{ für alle i=1...n, also}$$

$$F_{\text{bel}} \subseteq F_q \text{ für alle fokalen Elemente } F_q \text{ von } \mathrm{m}_\oplus$$
$$\text{mit } F_q := \bigcap_{1\leq i\leq n} [\varphi_i] \bigcap_{\substack{1\leq j\leq n\\ j\neq i}} \Omega_j,\ \Omega_j = \Omega,\ q = 1, \ldots 2^n$$

Daher gilt

$$\text{falls } F_{\text{bel}} \not\subseteq \{\omega_i\}, \text{ wegen } F_q \supseteq F_{\text{bel}} \text{ auch } F_q \not\subseteq \{\omega_i\}$$

für alle fokalen Elemente F_q und alle i, $1 \leq i \leq n$.

Wenn F_{bel} aus einer Vereinigung von Welten besteht, also

$$F_{\text{bel}} = \{\omega_{F1} \cup \ldots \cup \omega_{Fn}\}$$

dann gilt

$$\text{wegen } \{\omega_{F1} \cup \ldots \cup \omega_{Fn}\} \not\subseteq \{\omega_i\} :$$
$$\mathrm{bel}_\oplus(\omega_i) = 0 \text{ für alle } \omega_i \in \Omega$$

Wenn F_{bel} nur aus einer einzigen Welt besteht, also

$$F_{\text{bel}} = \{\omega_{\text{bel}}\}$$

dann genügt $\{\omega_{\text{bel}}\} = [\varphi_1] \cap \ldots \cap [\varphi_n]$ allen Defaults $\varphi_i \in \Delta$, d.h.

$$\sigma_\Delta(\omega_{\text{bel}}) = \prod_{\omega_{\text{bel}} \models A_i \wedge B_i} a_i^+ \prod_{\omega_{\text{bel}} \models \neg A_i} 1$$

für alle Defaults $d_i = (A_i \to B_i) \in \Delta$, und es gibt keinen Default, den ω_{bel} falsifiziert. Für alle anderen Welten

$$\omega_i \neq \omega_{\text{bel}} \text{ gilt } \{\omega_i\} \cap \{\omega_{\text{bel}}\} = \emptyset$$

und damit

$$F_{\text{bel}} = \{\omega_{\text{bel}}\} \not\subseteq \{\omega_i\}$$

also auch

$$F_q \not\subseteq \{\omega_i\} \text{ wegen } F_q \supseteq F_{\text{bel}}$$

und damit

$$\mathrm{bel}_{\oplus}(\omega_i) = 0$$

für alle $\omega_i \neq \omega_{\mathrm{bel}}$. □

Damit kann man nun zeigen:

Satz 8.3.3 (Belieffunktionen und konditionale Indifferenz)
Alle Belieffunktionen $\mathrm{bel}_{\oplus}$*, die man aus simple-support-Funktionen mit der Dempster'schen Kombinationsregel erhält, sind konditional indifferent.*

BEWEIS: Falls es keine einzelne Welt ω_{bel} mit $\mathrm{bel}_{\oplus}(\omega_{\mathrm{bel}}) \neq 0$ gibt, dann haben alle Welten $\omega_i \in \Omega$ den Beliefwert 0. Damit haben auch alle Welten und auch alle Juxtapositionen von Welten , die dieselbe konditinale Struktur haben, denselben Beliefwert, nämlich 0.

Falls es eine Welt ω_{bel} mit $\mathrm{bel}_{\oplus}(\omega_{\mathrm{bel}}) \neq 0$ gibt, dann ist diese Welt nach Satz 8.3.2 eindeutig bestimmt, und für alle übrigen Welten $\omega_i \neq \omega_{\mathrm{bel}}$ gilt $\mathrm{bel}_{\oplus}(\omega_i) = 0$. Da alle anderen Welten $\omega_i \neq \omega_{\mathrm{bel}}$ mindestens einen Default falsifizieren, erhält man den Beliefwert dieser Welten aus der konditionalen Struktur, indem man den a_i^- den Wert 0 zuweist. Damit ist dann ω_{bel} die einzige Welt, die einen positiven Beliefwert

$$\mathrm{bel}_{\oplus}(\omega_{\mathrm{bel}}) = \prod_{i=1}^{n}(1 - s_i)$$

hat. Deshalb gilt auch für alle

$$\hat{\omega} := \omega_1 \cdot \ldots \cdot \omega_k : \mathrm{bel}_{\oplus}(\hat{\omega}) = \left(\prod_{i=1}^{n}(1 - s_i)\right)^k \neq 0$$

falls $\omega_i = \omega_{\mathrm{bel}}\ \forall i = 1, \ldots, k$, also falls $\hat{\omega} = \omega_{\mathrm{bel}}^k$, und $\mathrm{bel}_{\oplus}(\hat{\omega}) = 0$ sonst. Damit erfüllen die Belief-Funktionen das Kriterium aus Definition 8.3.3 auf

Seite 81:

$$\sigma_\Delta(\hat{\omega}) = \prod_{1 \leq k \leq n} a_k^{+r_k} = \sigma_\Delta(\hat{\omega}')$$

$$\Rightarrow \mathrm{bel}_\oplus(\hat{\omega}) = \prod_{i=1}^{n}(1 - s_i)^{r_i} = \mathrm{bel}_\oplus(\hat{\omega}')$$

und

$$\mathrm{bel}_\oplus(\hat{\omega}) = 0 = \mathrm{bel}_\oplus(\hat{\omega}')$$

sonst.

Es gilt aber im Allgemeinen nicht $\mathrm{bel}_\oplus(\omega_1 \cup \omega_2) = \mathrm{bel}_{\oplus\varepsilon}(\omega_1 \cdot \omega_2)$, wie man an dem oben angegebenen Beispiel sieht. □

Kapitel 9

BSS-Ansatz

Der BSS-Ansatz [2] (benannt nach **B**enferhat, **S**afiotti und **S**mets) basiert auf infinitesimalen Belief- und Plausibilitätsfunktionen, die durch die Dempster'sche Kombinationsregel induziert werden.

9.1 Infinitesimale Funktionen

Grundlegend für den Begriff der infinitesimalen Funktionen ist der Begriff der Ordnung:

Definition 9.1.1 (Ordnung κ einer Funktion)
Sei f eine stetige Funktion von $\mathbb{R}$ nach $\mathbb{R}$. Eine positive ganze Zahl **k** wird Ordnung von f genannt, bezeichnet durch

$$\kappa(f) := \mathbf{k} \text{ genau dann, wenn } \lim_{\eta\to 0} \frac{f(\eta)}{\eta^{\boldsymbol{k}}} \in \mathbb{R} - \{0\}$$

wenn ein solcher Limes existiert.[1]Für reelle Konstanten $c \in \mathbb{R}$ gilt:

$$\kappa(c) = \begin{cases} 0 & \text{falls } c \neq 0 \\ = \infty & \text{falls } c = 0 \end{cases}$$

Beispiel 9.1.1 (Ordnung einer Funktion $f(\eta)$)
Für die Ordnung $\kappa(f)$ der Funktion $f(\eta) := 3\eta^9 + \eta^5 + 5\eta^2$ gilt wegen

$$\lim_{\eta\to 0} \frac{3\eta^9 + \eta^5 + 5\eta^2}{\eta^{\mathbf{2}}} = \lim_{\eta\to 0} 3\eta^7 + \eta^3 + 5 = 5 \in \mathbb{R} - \{0\}$$

$$\kappa(f) = \mathbf{2}$$

Durch die Ordnung $\kappa(f)$ wird also allen infinitesimalen Funktionen f ein Rang k in einem Ordnungssystem zugeordnet.

[1]Für die Funktionen $f(\eta) = \sqrt{\eta}$ oder $f(\eta) = \frac{1}{\eta}$ existiert z.B. keine Ordnung

Definition 9.1.2 (Infinitesimale ε)
Ein Infinitesimal ε ist eine reelle stetige Funktion von (dem reellen Intervall) (0,1) nach (0,1) mit

1. $\lim_{\eta \to 0} \varepsilon(\eta) = 0$ und
2. Die Ordnung $\kappa(\varepsilon)$ ist definiert

Es ist klar, dass ε umso dichter an 0 ist, je größer die Ordnung $\kappa(\varepsilon)$ ist. Zum Beispiel ist die Ordnung 9 des Infinitesimals $\varepsilon(\eta) = \eta^9$ größer als die Ordnung 2 des Infinitesimals $\varepsilon(\eta) = \eta^2$, und η^9 geht für $\eta \to 0$ schneller gegen 0 als η^2.

Definition 9.1.3 ($\mathbb{E}^0$, $\mathbb{E}^1$ und $\mathbb{E}$)
$\mathbb{E}^0$ ist die Menge aller Infinitesimale gemäß Definition 9.1.2.
Die Menge $\mathbb{E}^1$ wird daraufhin festgelegt durch

$$\mathbb{E}^1 = \{1 - \varepsilon \mid \varepsilon \in \mathbb{E}^0\}$$

Schließlich definiert man eine Menge $\mathbb{E}$ durch

$$\mathbb{E} := \mathbb{E}^0 \cup \mathbb{E}^1 \cup \{0\} \cup \{1\}$$

Die Infinitesimale aus $\mathbb{E}^0$ haben also wegen Definition 9.1.2 (1) den Grenzwert 0, und die Elemente $1 - \varepsilon(\eta)$ aus $\mathbb{E}^1$ nähern sich für $\eta \to 0$ dem Grenzwert 1. Die Ordnung k eines beliebigen Elements $(1-\varepsilon)$ aus $\mathbb{E}^1 \cup \{1\}$ ist $\kappa(1-\varepsilon) = 0$ wegen $\lim_{\eta \to 0} \frac{1-\varepsilon(\eta)}{\eta^0} = 1$ für $\varepsilon \in \mathbb{E}^0 \cup \{0\}$.

Um Infinitesimale miteinander zu vergleichen, verwendet man ihre Ordnung:

Definition 9.1.4 (Vergleich von Infinitesimalen)
Eine infinitesimale Funktion ε_1 wird als „infinitesimal kleiner als" eine infinitesimale Funktion ε_2 bezeichnet, notiert als

$$\varepsilon_1 <_\infty \varepsilon_2$$

genau dann, wenn $\kappa(\varepsilon_1) > \kappa(\varepsilon_2)$.
Entsprechend ist ε_1 „infinitesimal größer als" ε_2,

$$\varepsilon_1 >_\infty \varepsilon_2$$

genau dann, wenn $\kappa(\varepsilon_1) < \kappa(\varepsilon_2)$.
ε_1 und ε_2 werden als „infinitesimal gleich groß" bezeichnet, geschrieben als

$$\varepsilon_1 \approx_\infty \varepsilon_2$$

genau dann, wenn $\kappa(\varepsilon_1) = \kappa(\varepsilon_2)$.

Folgende nützliche Eigenschaften erleichtern den Vergleich verschiedener Infinitesimale (vgl. [2] (Anhang A))

Lemma 9.1.1 (ε-Arithmetik)
Seien t, t' zwei beliebige Elemente aus $\mathbb{E}$. Es gilt

1. $\kappa(t+t') = \min(\kappa(t), \kappa(t'))$ *und für* $t + t' <_\infty 1$ *oder* $t(\eta) + t'(\eta) = 1$ *gilt* $t + t' \in \mathbb{E}$ *([2]) Lemma A2.a)*
   Eine Summe von Infinitesimalen oder beliebigen Elementen aus $\mathbb{E}$ wird von dem Summanden mit dem kleinsten Rang dominiert.
   Diese Eigenschaft lässt sich dadurch begründen, dass eine Summe sich nur so schnell wie der infinitesimal größte Summand dem Wert 0 nähern kann, und dieser „größte" Summand ist derjenige mit der kleinsten Ordnung k.

2. $\kappa(\max(t, t')) = \min(\kappa(t), \kappa(t'))$ *und* $\max(t, t') \in \mathbb{E}$ *([2] Lemma A2.b)*
   Das bezüglich der Ordnung κ infinitesimal größte Element aus einer

Menge von Funktionen aus $\mathbb{E}$ ist dasjenige Element mit dem kleinsten Rang.
Diese Eigenschaft folgt direkt aus Definition 9.1.4: $\varepsilon_1 >_\infty \varepsilon_2$ genau dann, wenn $\kappa(\varepsilon_1) < \kappa(\varepsilon_2)$

3. $\kappa(t \cdot t') = \kappa(t) + \kappa(t')$ und $t \cdot t' \in \mathbb{E}$ ([2] Lemma A2.c)
Der Rang eines Produktes von Elementen aus $\mathbb{E}$ ist gleich der Summe der Ränge aller Faktoren.
Diese Eigenschaft wird dadurch begründet, dass die Ordnung k der Infinitesimalrechnung von den Exponenten der Elemente aus $\mathbb{E}$ abhängt und dass bei der Produktbildung die Exponenten der Faktoren addiert werden.

4. $t >_\infty t'$ genau dann, wenn $\kappa(t) < \kappa(t')$ genau dann, wenn $\lim_{\eta\to\infty} \frac{t'(\eta)}{t(\eta)} = 0$ ([2] Lemma A.4b)
t ist infinitesimal größer als t' genau dann, wenn $\kappa(t) < \kappa(t')$ genau dann, wenn sich t für $\eta \to 0$ langsamer als t' dem Grenzwert 0 nähert.
Diese Eigenschaft folgt aus den Definitionen 9.1.1 und 9.1.4

5. $t >_\infty t'$ genau dann, wenn $t' + t >_\infty t'$ ([2] Lemma A8.b)
t ist infinitesimal größer als t' genau dann, wenn $t' + t$ infinitesimal größer als t' ist. Die Richtung $\Rightarrow$ gilt auch in der Standard-Analysis, also z.B. für $t, t' \in \mathbb{N}$ oder $t, t' \in \mathbb{R}$. Deshalb wird diese Eigenschaft meistens in der Richtung $\Leftarrow$ verwendet, die in der Standard-Analysis nicht gilt, wie das Beispiel 2+1>2, aber 1<2 zeigt.
Die Begründung für die $\Leftarrow$-Richtung ergibt sich daraus, dass die Ordnung einer Summe t+t' nach Punkt 1 dieses Lemma von dem Summanden mit dem kleinsten Rang, also von dem infinitesimal größten Summanden, bestimmt wird. Wegen $t' \approx_\infty t'$ kann dieser infinitesimal größte Summand nur der Summand t sein.

6. Für alle $r \in \mathbb{R}^+$ gilt $t \approx_\infty r \cdot t$, vorausgesetzt, dass $r \cdot t \in \mathbb{E}$. ([2] Lemma A8.e)

Die Multiplikation mit beliebigen Konstanten aus $\mathbb{R}^+$ *hat keinen Einfluss auf den Rang einer infinitesimalen Funktion* t *aus* $\mathbb{E}^0$ *oder auf den Rang eines beliebigen Elements* t *aus* $\mathbb{E}^1 \cup \{0\} \cup \{1\}$.
Wenn $\lim_{\eta \to 0} \frac{t(\eta)}{\eta^k} \in \mathbb{R}\backslash 0$ *ist, dann ist auch* $\lim_{\eta \to 0} \frac{r \cdot t(\eta)}{\eta^k} \in \mathbb{R}\backslash 0$

7. *Seien* $s, s' \in \mathbb{E}$. *Wenn* $s \approx_\infty s'$ *und* $t \approx_\infty t'$, *dann* $s >_\infty t$ *genau dann, wenn* $s' >_\infty t'$ *([2] Lemma A9.b)*
Wenn zwei Infinitesimale oder sonstige Elemente aus $\mathbb{E}$ *infinitesimal gleich sind, dann haben sie auch dieselbe Position in einem Ordnungssystem bezüglich der Ordnung „*$>_\infty$*“.*
$s \approx_\infty s'$ *und* $t \approx_\infty t'$ *genau dann, wenn* $\kappa(s) = \kappa(s')$ *und* $\kappa(t) = \kappa(t')$.
Daraus folgt mit den Eigenschaften der natürlichen Zahlen $\kappa(s) < \kappa(t)$ *genau dann, wenn* $\kappa(s') < \kappa(t')$, *mit* $\kappa(x) \in \mathbb{N}\ \forall\ x \in \{s, s', t, t'\}$

8. *Seien* $t_1, \ldots, t_n \in \mathbb{E}$. *Dann gilt* $\max_{i=1,\ldots,n} t_i \approx_\infty \sum_{i=1}^{n} t_i$ *([2] Lemma A11)*
Eine Summe von beliebigen Elementen aus $\mathbb{E}$ *ist infinitesimal gleich dem infinitesimal größten Summanden, also dem Summanden mit dem kleinsten Rang bezüglich* κ.
Die Begründung dieser Eigenschaften ergibt sich aus Punkt 1 und 2 dieses Lemmas, da sowohl der Rang einer Summe als auch der Rang des Maximums einer Menge von dem Summanden bzw. Element mit dem kleinsten Rang bestimmt wird.

9.2 BSS-Ansatz: Infinitesimale Belief- und Plausibilitätsfunktionen

Sei Δ wieder eine Menge von Defaults d_i mit $n := |\Delta|$, und für alle $i \in \mathbb{N}, 1 \leq i \leq n$, sei φ_i die materiale Entsprechung zu d_i. Zunächst wird jedem Default

$d_i \in \Delta$ eine simple-support-function zugeordnet:

$$\mathrm{m}_i(\Omega) = \varepsilon_i,\ \mathrm{m}_i(\varphi_i) = 1 - \varepsilon_i \text{ und } \mathrm{m}_i(X) = 0 \text{ sonst} \tag{9.1}$$

Den ε_i werden in dem BSS-Ansatz entweder konstante Funktionen $\varepsilon_i := 0$ $\varepsilon_i := 1$ oder infinitesimale Funktionen zugeordnet, d.h. $\varepsilon_i \in \mathbb{E}\ \forall\ i = 1, \ldots, n$.

Beispiel 9.2.1 (BSS-Ansatz)

Sei $\Delta = \{d_1 = v \rightarrow f, d_2 = p \rightarrow v, d_3 = p \rightarrow \neg f\}$ wieder die bekannte Default-Basis aus Beispiel 2.4.1. Es sei $[\varphi_i]$ die Menge aller Welten, die φ_i nicht falsifizieren, z.B.

$$[\varphi_1] \cap [\varphi_2] \cap [\varphi_3] = [\neg v \vee f] \cap [\neg p \vee v] \cap [\neg p \vee \neg f] = \\ \{[v \wedge \neg p \wedge f], [\neg v \wedge \neg p \wedge f], [\neg v \wedge \neg p \wedge \neg f]\}$$

Man ordnet dann allen Defaults eine simple-support-function m_i z.B. folgendermaßen zu:
$\mathrm{m}_1(\neg v \vee f) = 1 - \varepsilon_1$, $\mathrm{m}_2(\neg p \vee v) = 1 - \varepsilon_2$ *und* $\mathrm{m}_3(\neg p \vee \neg f) = 1 - \varepsilon_3$. *Den einzelnen ε_i können beliebige konkrete Infinitesimale zugeordnet sein wie z.B. $\varepsilon_1(\eta) = \eta^2$, $\varepsilon_2(\eta) = \eta^3$ und ebenfalls $\varepsilon_3(\eta) = \eta^3$ mit $\kappa(\varepsilon_1) = 2$, $\kappa(\varepsilon_2) = \kappa(\varepsilon_3) = 3$. Es folgt nach Gleichung (9.1): $\mathrm{m}_1(\Omega) = \varepsilon_1$, $\mathrm{m}_2(\Omega) = \varepsilon_2$, $\mathrm{m}_3(\Omega) = \varepsilon_3$ und $\mathrm{m}_i(X_i)$=0 für alle $X_i \neq \Omega$ und $X_i \neq \varphi_i$, $i = 1, 2, 3$.*

Die einzelnen simple-support-functions m_i werden durch die Dempster'sche Kombinationsregel aus Definition 8.2.1 zu einer Funktion $\mathrm{m}_\oplus$ verknüpft.

$$\mathrm{m}_\oplus(F) = \sum_{X_1 \cap \ldots \cap X_n = F} \prod \mathrm{m}_i(X_i)$$

wobei nach Gleichung (9.1) nur alle $X_i \in \{[\varphi_i], \Omega\}$ zu $\mathrm{m}_\oplus(F) > 0$ beitragen, da für alle $X_i \notin \{[\varphi_i], \Omega\}$ gilt $\mathrm{m}_i(X_i) = 0$. Die Normalisierungskonstante c, die man im Allgemeinen zur Normalisierung von $m_\oplus$ benötigt, ist in dem BSS-Ansatz für konsistente Default-Basen immer gleich 1 und spielt somit keine Rolle.[2]. Die X_i können nämlich in diesem Ansatz nur die Werte $X_i = \Omega$ oder $X_i = [\varphi_i]$ annehmen, wobei $[\varphi_i]$ die materiale Entsprechung zu dem i-ten Default ist, und für konsistente Default-Basen gibt es immer eine Welt ω, die allen Defaults genügt, und damit ist $\omega \subseteq [\varphi_1 \cap \ldots \cap \varphi_n] \neq \emptyset$.

Beispiel 9.2.2 (BSS-Ansatz)

Die fokalen Elemente von $m_\oplus$ sind dann für das obige Beispiel 9.2.1

F_i	*Fokale Elemente F_i von $m_\oplus$*	$m_\oplus(F_i)$
F_1	$([\neg v \vee f] \cap [\neg p \vee v] \cap [\neg p \vee \neg f])$	$(1-\varepsilon_1)(1-\varepsilon_2)(1-\varepsilon_3)$
F_2	$([\neg v \vee f] \cap [\neg p \vee v] \cap \Omega)$	$(1-\varepsilon_1)(1-\varepsilon_2)\varepsilon_3$
F_3	$([\neg v \vee f] \cap \Omega \cap [\neg p \vee \neg f])$	$(1-\varepsilon_1)\varepsilon_2(1-\varepsilon_3)$
F_4	$([\neg v \vee f] \cap \Omega \cap \Omega)$	$(1-\varepsilon_1)\varepsilon_2\varepsilon_3$
F_5	$(\Omega \cap [\neg p \vee v] \cap [\neg p \vee \neg f])$	$\varepsilon_1(1-\varepsilon_2)(1-\varepsilon_3)$
F_6	$(\Omega \cap [\neg p \vee v] \cap \Omega)$	$\varepsilon_1(1-\varepsilon_2)\varepsilon_3$
F_7	$\Omega \cap \Omega \cap [\neg p \vee \neg f])$	$\varepsilon_1\varepsilon_2(1-\varepsilon_3)$
F_8	$(\Omega \cap \Omega \cap \Omega)$	$\varepsilon_1\varepsilon_2\varepsilon_3$

Tabelle 9.1: $m_\oplus$ für Beispiel 9.2.1

[2]([2] S. 11 oben)

Man kann leicht nachrechnen, dass $\sum_{i=1}^{8} \mathrm{m}_{\oplus}(F_i) = 1$ *und* $\mathrm{m}_{\oplus}(\emptyset) = 0$*, so dass auch* $\mathrm{m}_{\oplus}$ *ein Basismaß ist.*

Wegen Lemma 9.1.1 (3) gilt für alle F_i*:* $\mathrm{m}_{\oplus}(F_i) \in \mathbb{E}$*, da alle Produkte* $t \cdot t'$ *von Infinitesimalen* $t, t' \in \mathbb{E}$ *wieder Elemente aus* $\mathbb{E}$ *sind. Man sieht, dass nur für ein einziges fokales Element, nämlich* F_1*, gilt* $\mathrm{m}_{\oplus}(F_1) \in \mathbb{E}^1$ *wegen* $\mathrm{m}_{\oplus}(F_1) = \mathrm{m}_{\oplus}([\neg v \vee f] \cap [\neg p \vee v] \cap [\neg p \vee \neg f]) = (1-\varepsilon_1)(1-\varepsilon_2)(1-\varepsilon_3) \in \mathbb{E}^1$ *und* $\mathrm{m}_{\oplus}(X) \in \mathbb{E}^0$ *für alle* $X \neq ([\neg v \vee f] \cap [\neg p \vee v] \cap [\neg p \vee \neg f])$*. Dies entspricht der Tatsache, dass man nur alle Welten, die allen vorhandenen Defaults genügen, für glaubwürdig hält, und alle Welten, die einen Default falsifizieren, als unglaubwürdig einschätzt.*

Der BSS-Ansatz gewinnt aus den $\mathrm{m}_{\oplus}$ Belief- und Plausibilitätsfunktionen durch die bekannten Definitionen: $\mathrm{bel}_{\oplus} : 2^{\Omega} \to [0,1]$ ist für alle $A \subseteq \Omega$ gegeben durch

$$\mathrm{bel}_{\oplus}(A) \overset{Def. 8.1.3}{=} \sum_{B: B \subseteq A} \mathrm{m}_{\oplus}(B)$$

und die Plausibilität $\mathrm{pl}_{\oplus} : 2^{\Omega} \to [0,1]$ erhält man aus $\mathrm{m}_{\oplus}$ durch

$$\mathrm{pl}_{\oplus}(A) \overset{Def. 8.1.4}{=} \sum_{B: B \cap A \neq \emptyset} \mathrm{m}_{\oplus}(B)$$

Da der BSS-Ansatz nur die Grenzwerte der Belief- und Plausibilitätsfunktionen für $\varepsilon(\eta) \to 0$ untersucht, genügt es, die Belief- und Plausibilitätsfunktionen auf die entsprechenden Infinitesimale abzubilden, also ist $\mathrm{bel}_{\oplus\varepsilon}$ für $A \subseteq \Omega$ gegeben durch

Definition 9.2.1 (Infinitesimale Belieffunktion)

$$\mathrm{bel}_{\oplus\varepsilon} : 2^{\Omega} \to \mathbb{E}, \ \mathrm{bel}_{\oplus\varepsilon}(A) \approx_{\infty} \sum_{B: B \subseteq A} \mathrm{m}_{\oplus}(B)$$

und für die Plausibilität $\mathrm{pl}_{\oplus\varepsilon} : 2^{\Omega} \to \mathbb{E}$ gilt

Definition 9.2.2 (Infinitesimale Plausibilitätsfunktion)

$$\mathrm{pl}_{\oplus\varepsilon} : 2^{\Omega} \to \mathbb{E},\ \mathrm{pl}_{\oplus\varepsilon}(A) \approx_{\infty} \sum_{B:B\cap A\neq\emptyset} \mathrm{m}_{\oplus}(B)$$

wobei „$\approx_{\infty}$“ das Verhalten der Belief- und Plausibilitätsfunktionen für $\varepsilon_i(\eta)_{\eta\to 0}$ beschreibt.

Im BSS-Ansatz werden den einzelnen ε_i und damit auch den einzelnen simple-support-Funktionen $\mathrm{m}_i(d_i) = 1 - \varepsilon_i$ keine konkreten Werte aus dem Intervall $[0,1]$, sondern infinitesimale Funktionen $\varepsilon_i(\eta) \in \mathbb{E}$ zugeordnet, wobei diese einzelnen infinitesimalen Funktionen und damit auch alle m_i wiederum von einem η abhängen. Damit hängen natürlich auch die aus den m_i gebildeten Funktionen $\mathrm{m}_{\oplus}$ und die aus diesen $\mathrm{m}_{\oplus}$ gebildeten Belief- und Plausibilitätsfunktionen von η ab. Streng genommen müsste man deshalb das η immer zusätzlich angeben und z.B. in Definition 9.2.2 $\mathrm{pl}_{\oplus\varepsilon}(A)(\eta) \approx_{\infty} \sum_{B:B\cap A\neq\emptyset} \mathrm{m}_{\oplus}(B)(\eta)$ schreiben. Der Einfachheit halber wird die Angabe von η im folgenden Text weggelassen, wenn klar ist, dass es sich um infinitesimale Funktionen handelt. Im folgenden Text sollen nicht-infinitesimale Belief-bzw. Plausibilitätsfunktionen durch $\mathrm{bel}_{\oplus}$ bzw. $\mathrm{pl}_{\oplus}$ und ininitesimale Belief-bzw. Plausibilitätsfunktionen durch $\mathrm{bel}_{\oplus\varepsilon}$ bzw. $\mathrm{pl}_{\oplus\varepsilon}$ bezeichnet werden.

Die Folgerungsrelation des BSS-Ansatzes ist schließlich gegeben durch die folgende Definition:

Definition 9.2.3 (Folgerungsrelation des BSS-Ansatzes)
Sei $\mathrm{bel}_{\oplus\varepsilon}$ eine infinitesimale Belieffunktion, und sei $A \to B$ eine Default-Regel. Es gilt

$$\mathrm{bel}_{\oplus\varepsilon} \models A \to B$$

auch geschrieben als

$$A \mathrel{|\!\sim}_{\oplus} B$$

genau dann, wenn

$$\mathrm{pl}_{\oplus\varepsilon}(A \wedge B) >_{\infty} \mathrm{pl}_{\oplus\varepsilon}(A \wedge \neg B)$$

(vgl.[2] Definition 4)

Die Plausibilität einer Formel A kann man laut [2] Lemma 3 aus der Summe oder dem Maximum der Plausibilitäten aller Welten ω berechnen, die die Formel verifizieren:

Satz 9.2.1 ($\mathrm{pl}_{\oplus\varepsilon}$, Summe und Maximum)
Für alle $A \subseteq \Omega$ gilt

1. $\mathrm{pl}_{\oplus\varepsilon}(A) \approx_{\infty} \sum_{\omega\in[A]} \mathrm{pl}_{\oplus\varepsilon}(\omega)$

2. $\mathrm{pl}_{\oplus\varepsilon}(A) \approx_{\infty} \max_{\omega\in[A]}\{\mathrm{pl}_{\oplus\varepsilon}(\omega)\}$

BEWEIS: [3] Laut Definition der Plausibilität gilt

$$\sum_{\omega\in[A]} \mathrm{pl}_{\oplus\varepsilon}(\omega) = \sum_{\omega\in[A]} \sum_{Y_i:\omega\in Y_i} \mathrm{m}_{\oplus}(Y_i) = \sum_{Y_i\cap A\neq\emptyset} k_{Y_i} \cdot \mathrm{m}_{\oplus}(Y_i)$$

k_{Y_i} bezeichnet die Anzahl der Welten, die Y_i und A verifizieren, also

$$k_{Y_i} = |Y_i \cap A|$$

[3] vgl.[2] Lemma 3

Wegen $\mathrm{m}_\oplus(Y_i) \in \mathbb{E}$ (Bemerkung nach Satz 8.2.1) für alle $A_i \in \Omega$ gilt

$$k_{Y_i} \cdot \mathrm{m}_\oplus(Y_i) \approx_\infty \mathrm{m}_\oplus(Y_i)$$

nach Lemma 9.1.1 (6), und deshalb gilt

$$\sum_{\omega \in [A]} \mathrm{pl}_{\oplus\varepsilon}(\omega) \approx_\infty \sum_{Y_i \cap A \neq \emptyset} \mathrm{m}_\oplus(Y_i)$$

und die rechte Seite dieser Formel ist genau die Definition von $\mathrm{pl}_{\oplus\varepsilon}(A)$ und beweist damit Teil 1 des Satzes.
Teil 2 des Satzes erhält man aus Teil 1 direkt mit Lemma 9.1.1 (8). □

In [2] wird in Lemma 4 gezeigt, dass die Definition 9.2.3 äquivalent zu der Forderung $\lim_{\eta\to 0}(\mathrm{bel}_{\oplus\varepsilon}(A_i \to B_i)) = 1$ ist:

Satz 9.2.2 (Zusammenhang zwischen Belief und Plausibilität)

$$\mathrm{bel}_{\oplus\varepsilon} \models \alpha \to \beta, \text{ also } \mathrm{pl}_{\oplus\varepsilon}(\alpha \wedge \beta) >_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)$$

genau dann, wenn

$$\lim_{\eta\to 0} \mathrm{bel}_{\oplus\varepsilon}(\alpha \to \beta) = 1$$

Beweis: [4] $\Longrightarrow$:
Es gilt

$$\mathrm{bel}_{\oplus\varepsilon}(\alpha \to \beta) \overset{(Satz\ 8.1.1)}{=} 1 - \mathrm{pl}_{\oplus\varepsilon}(\overline{\alpha \to \beta})$$

$\mathrm{bel}_{\oplus\varepsilon}(\alpha \to \beta)$ ist das Beliefmaß, das man β zuordnet unter der Bedingung, dass man α sicher weiss. $\mathrm{pl}_{\oplus\varepsilon}(\alpha \to \neg\beta)$ ist der Plausibilitätswert, den man $\neg\beta$ zuordnet unter der Voraussetzung, dass man α sicher weiss. Wegen

[4] vgl. [2] Lemma 4

$\text{bel}_{\oplus\varepsilon}(\beta) \overset{(Satz\ 8.1.1)}{=} 1 - \text{pl}_{\oplus\varepsilon}(\beta)$ gilt $\text{bel}_{\oplus\varepsilon}(\alpha \to \beta) = 1 - \text{pl}_{\oplus\varepsilon}(\alpha \to \neg\beta)$, also

$$\text{bel}_{\oplus\varepsilon}(\alpha \to \beta) = 1 - \text{pl}_{\oplus\varepsilon}(\alpha \to \neg\beta) \overset{(8.8)}{=} 1 - \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)}$$

Wegen

$$\text{pl}_{\oplus\varepsilon}(\alpha) \geq \text{pl}_{\oplus\varepsilon}(\alpha \wedge \beta)$$

gilt

$$\frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)} \leq \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \beta)}$$

Wegen

$$\text{pl}_{\oplus\varepsilon}(\alpha \wedge \beta) >_\infty \text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta) \text{ (Def.9.2.3)}$$

gilt

$$\lim_{\eta \to 0} \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \beta)} = 0 \text{ (Lemma 9.1.1(4))}$$

und damit auch

$$\lim_{\eta \to 0} \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)} \leq \lim_{\eta \to 0} \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \beta)} = 0$$

Also folgt

$$\lim_{\eta \to 0} \text{bel}_{\oplus\varepsilon}(\alpha \to \beta) = \lim_{\eta \to 0} \left(1 - \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)}\right) = 1$$

$\Longleftarrow$: Es gelte

$$\lim_{\eta \to 0} \text{bel}_{\oplus\varepsilon}(\alpha \to \beta) = 1$$

Dann folgt

$$\lim_{\eta \to 0} (1 - \text{pl}(\alpha \to \neg\beta)) \overset{(8.8)}{=} \lim_{\eta \to 0} \left(1 - \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)}\right) = 1$$

und deshalb

$$\lim_{\eta \to 0} \frac{\text{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta)}{\text{pl}_{\oplus\varepsilon}(\alpha)} = 0$$

Das bedeutet

$$\mathrm{pl}_{\oplus\varepsilon}(\alpha) >_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha \wedge \neg\beta) \text{ (Lemma 9.1.1(4))} \tag{9.2}$$

Wegen

$$\begin{aligned}
\mathrm{pl}_{\oplus\varepsilon}(\alpha) &\overset{Def.9.2.2}{\approx_\infty} \sum_{C:C\cap[\alpha]\neq\emptyset} \mathrm{m}_\oplus(C) \\
&\approx_\infty \sum_{C:C\cap[\neg\beta\wedge\alpha]\neq\emptyset} \mathrm{m}_\oplus(C) + \sum_{C:C\cap[\beta\wedge\alpha]\neq\emptyset} \mathrm{m}_\oplus(C) \\
&\qquad \text{(wenn ein fokales Element C sowohl } [\alpha\wedge\beta] \\
&\qquad \text{ als auch } [\alpha\wedge\neg\beta] \text{ schneidet, dann hat dies} \\
&\qquad \text{wegen } pl_{\oplus\varepsilon}(C) \approx_\infty \mathrm{pl}_{\oplus\varepsilon}(C) + \mathrm{pl}_{\oplus\varepsilon}(C) \text{ (Lemma9.1.1)} \\
&\qquad \text{keinen Einfluss auf die infinitesimale Plausibilität)} \\
&\approx_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\neg\beta) + \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\beta)
\end{aligned}$$

gilt

$$\mathrm{pl}_{\oplus\varepsilon}(\alpha) \approx_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\neg\beta) + \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\beta)$$

und daraus und aus (9.2) folgt mit Lemma 9.1.1(7)

$$\mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\neg\beta) + \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\beta) >_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\neg\beta)$$

und mit Lemma 9.1.1 (5) ergibt sich

$$\mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\beta) >_\infty \mathrm{pl}_{\oplus\varepsilon}(\alpha\wedge\neg\beta)$$

□

Beispiel 9.2.3 (Folgerungsrelation des BSS-Ansatzes)
Aus der obigen Tabelle auf Seite 95 entnimmt man, dass $[p\wedge f]$ *die fokalen Elemente von* $\mathrm{m}_\oplus$ F_2, F_4, F_6 *und* F_8 *aus der Tabelle schneidet, also* $\mathrm{pl}_{\oplus\varepsilon}(p\wedge f) = \sum_{F\cap[p\wedge f]\neq\emptyset} \mathrm{m}_\oplus(F) = (1-\varepsilon_1)(1-\varepsilon_2)\varepsilon_3 + (1-\varepsilon_1)\varepsilon_2\varepsilon_3 + \varepsilon_1(1-\varepsilon_2)\varepsilon_3 + \varepsilon_1\varepsilon_2\varepsilon_3 = \varepsilon_3$. *Für* $[p\wedge\neg f]$ *erhält man, da* $[p\wedge\neg f]$ *alle fokalen Elemente aus der Tabelle*

F_3 bis F_8 einschließlich schneidet, $\mathrm{pl}_{\oplus\varepsilon}(p \wedge \neg f) = (1-\varepsilon_1)\varepsilon_2(1-\varepsilon_3) + (1-\varepsilon_1)\varepsilon_2\varepsilon_3 + \varepsilon_1(1-\varepsilon_2)(1-\varepsilon_3) + \varepsilon_1(1-\varepsilon_2)\varepsilon_3 + \varepsilon_1\varepsilon_2(1-\varepsilon_3) + \varepsilon_1(1-\varepsilon_2)\varepsilon_3 = (1-\varepsilon_1)\varepsilon_2 + \varepsilon_1(1-\varepsilon_2) + \varepsilon_1\varepsilon_2 = \varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2$. *Wenn man wie im vorherigen Beispiel* $\varepsilon_1(\eta) = \eta^2$, $\varepsilon_2(\eta) = \eta^3 = \varepsilon_3(\eta)$ *gewählt hat, dann gilt* $\kappa(\varepsilon_3) = 3$ *und* $\kappa(\varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2) = 2$. *Daraus folgt wegen* $\kappa(\mathrm{pl}_{\oplus\varepsilon}(p \wedge f)) > \kappa(\mathrm{pl}_{\oplus\varepsilon}(p \wedge \neg f))$: $\mathrm{pl}_{\oplus\varepsilon}(p \wedge f) <_\infty \mathrm{pl}_{\oplus\varepsilon}(p \wedge \neg f)$, *und deshalb gilt in diesem Modell* $\mathrm{bel}_{\oplus\varepsilon} \models p \to \neg f$.

In [2] wird ausserdem gezeigt:

Satz 9.2.3 ($\mid\!\sim_{\oplus\varepsilon}$ und Maximum von $\mathrm{pl}_{\oplus\varepsilon}$)
Es gilt $\mathrm{bel}_{\oplus\varepsilon} \models A_i \to B_i$ *genau dann, wenn*

$$\max_{\omega \models A_i \wedge B_i} \mathrm{pl}_{\oplus\varepsilon}(\omega) >_\infty \max_{\omega \models A_i \wedge \neg B_i} \mathrm{pl}_{\oplus\varepsilon}(\omega)$$

([2] Lemma 4 und (10))

Beweis: Diese Eigenschaft folgt direkt aus Satz 9.2.1 und Definition 9.2.3. □

In dem BSS-Ansatz wird in [2] Lemma 9 die Berechnungsgrundlage für die Plausibilität von Welten gezeigt:

Lemma 9.2.1 (Berechnung von $\mathrm{pl}_{\oplus\varepsilon}$)
Sei $\mathrm{bel}_{\oplus\varepsilon}$ *eine Belieffunktion. Dann gilt für alle Welten* $\omega \in \Omega$

$$\mathrm{pl}_{\oplus\varepsilon}(\omega) \quad \approx_\infty \quad \prod_{\omega \not\models \varphi_d} \varepsilon_d \tag{9.3}$$

Lemma 9.2.1 wird in [2] direkt bewiesen, man kann das Lemma aber auch aus der nicht-infinitesimalen Plausibilität mit Satz 8.2.2 herleiten, denn aus

$$\mathrm{pl}_{\oplus\varepsilon}(\omega) = \prod_{\omega \not\models \varphi_d} s_d$$

folgt, wenn man für die s_d die infinitesimalen Funktionen ε_d einsetzt:

$$\mathrm{pl}_{\oplus\varepsilon}(\omega) \approx_\infty \prod_{\omega \not\models \varphi_d} \varepsilon_d$$

9.3 BSS-Ansatz und konditional indifferente OCFs

Es soll in diesem Abschnitt zunächst gezeigt werden, dass man alle Logiken, die man aus dem BSS-Ansatz erhält, auch durch konditional indifferente OCFs ausdrücken kann.

Satz 9.3.1 (BSS-Ansatz und OCFs)
Zu jeder infinitesimalen Plausibilitätsfunktion $\mathrm{pl}_{\oplus\varepsilon}$ des BSS-Ansatzes, die man aus simple-support-functions mit der Dempster´schen Kombinationsregel nach Definition 9.2.2 erhält, gibt es eine konditional indifferente OCF $\kappa_\oplus$ mit $\kappa_{\oplus i}^- = \kappa(\varepsilon_i)$.

Die konditionale Indifferenz von OCFs kann man laut Definition 3.0.11 überprüfen durch folgende Bedingung: Eine (endliche) OCF ist indifferent in bezug auf eine Menge von Defaults $\Delta = \{(A_1 \to B_1), ..., (A_n \to B_n)\}$ genau dann, wenn es rationale Zahlen $\kappa_0, \kappa_i^+, \kappa_i^- \in \mathbb{Q}, 1 \leq i \leq n$ gibt, so dass für alle $\omega \in \Omega$ gilt:

$$\kappa(\omega) = \kappa_0 + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i B_i}} \kappa_i^+ + \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_i^-$$

Laut Lemma 9.2.1, S.101, gilt für alle $\omega \in \Omega$ und für alle Plausibilitätsfunktionen, die man mit der Dempster´schen Kombinationsregel erhält:

$$\mathrm{pl}_{\oplus\varepsilon}(\omega) \quad \approx_\infty \prod_{\omega \not\models \varphi_i} \varepsilon_i$$

Nach Lemma 9.1.1 folgt

$$\kappa(\mathrm{pl}_{\oplus\varepsilon}(\omega)) = \kappa(\prod_{\omega \not\models \varphi_i} \varepsilon_i) = \sum_{\omega \not\models \varphi_i} \kappa(\varepsilon_i) \tag{9.4}$$

Ausserdem gilt $\kappa(\mathrm{pl}_{\oplus\varepsilon}(\omega)) = \kappa(1) = 0$, falls $\omega \models \varphi_i \ \forall i, 1 \leq i \leq n$. Setzt man also

$$\kappa_{\oplus i} := \kappa(\varepsilon_i)$$

und

$$\kappa_{\oplus}(\omega) := \kappa(\mathrm{pl}_{\oplus\varepsilon}(\omega)),$$

dann erhält man

$$\kappa_{\oplus}(\omega) = \sum_{\omega \not\models \varphi_i} \kappa_{\oplus i} = \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_{\oplus i}.$$

Damit ist $\kappa_{\oplus}$ konditional indifferent in bezug auf Δ, wenn man $\kappa_0 := 0$, $\kappa_i^+ := 0$ und $\kappa_i^- := \kappa_{\oplus i}$ setzt.

Satz 9.3.2 (Folgerungsrelation von $\kappa_{\oplus}$)
Für alle OCFs $\kappa_{\oplus}$, die man aus den $\mathrm{pl}_{\oplus\varepsilon}$ bzw. den dazu passenden $\mathrm{bel}_{\oplus\varepsilon}$ nach der oben genannten Methode erhält, gilt $\kappa_{\oplus} \models A \rightarrow B$ genau dann, wenn $\mathrm{bel}_{\oplus\varepsilon} \models A \rightarrow B$.

Beweis: Laut Satz 9.2.3 gilt:

$$\mathrm{bel}_{\oplus\varepsilon} \models A \rightarrow B \text{ genau dann, wenn } \max_{\omega \models A \wedge B} \mathrm{pl}_{\oplus\varepsilon}(\omega) >_{\infty} \max_{\omega \models A \wedge \neg B} \mathrm{pl}_{\oplus\varepsilon}(\omega)$$

für alle $\mathrm{bel}_{\oplus\varepsilon} \in \mathrm{bel}_{\oplus}(\Delta)$.
Es gilt

$$\max_{\omega \models A \wedge B} \mathrm{pl}_{\oplus\varepsilon}(\omega) >_{\infty} \max_{\omega \models A \wedge \neg B} \mathrm{pl}_{\oplus\varepsilon}(\omega)$$

genau dann, wenn

$$\kappa(\max_{\omega \models A \wedge B} \mathrm{pl}_{\oplus\varepsilon}(\omega)) < \kappa(\max_{\omega \models A \wedge \neg B} \mathrm{pl}_{\oplus\varepsilon}(\omega)) \text{ (mit Lemma 9.1.1(4))}$$

genau dann, wenn

$$\min_{\omega \models A \wedge B} \kappa(\mathrm{pl}_{\oplus\varepsilon}(\omega)) < \min_{\omega \models A \wedge \neg B} \kappa(\mathrm{pl}_{\oplus\varepsilon}(\omega)) \text{ (mit Lemma 9.1.1(2))}$$

genau dann, wenn

$$\min_{\omega \models AB} \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_{\oplus i} < \min_{\omega \models A\overline{B}} \sum_{\substack{1 \leqslant i \leqslant n \\ \omega \models A_i \overline{B}_i}} \kappa_{\oplus i} \text{ (wegen (9.4))}$$

genau dann, wenn

$$\kappa_{\oplus}(AB) < \kappa_{\oplus}(A\overline{B})$$

also $\mathrm{bel}_{\oplus\varepsilon} \models A \to B$ genau dann, wenn $\kappa_{\oplus} \models A \to B$ □

In [1] Definition 2 werden alle OCFs, die eine Menge Δ von Defaults repräsentieren und die konditional indifferent in bezug auf diese Menge Δ sind, als c-Repräsentationen bezeichnet.

Satz 9.3.3 (BSS-Ansatz und c-Repräsentationen)

Zu jeder Belieffunktion $\mathrm{bel}_{\oplus\varepsilon}$ *des BSS-Ansatzes gibt es eine c-Repräsentation* $\kappa_{\oplus}$ *mit*

$$\mathrm{bel}_{\oplus\varepsilon} \models A \to B \text{ genau dann, wenn } \kappa_{\oplus} \models A \to B$$

Diese c-Repräsentationen erfüllen die Bedingung (12) aus [1] und sind daher eine Variante des generalisierten System-Z, für das gilt*[5]*:*

$$\kappa_j^- > \min_{\omega \models A_j B_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \kappa_i^- - \min_{\omega \models A_j \overline{B}_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \kappa_i^- \tag{9.5}$$

BEWEIS: Aus

$$\min_{\omega\models A_jB_j} \sum_{\substack{1\leqslant i\leqslant n\\ \omega\models A_i\overline{B}_i}} \kappa_{\oplus i} < \min_{\omega\models A_j\overline{B}_j} \sum_{\substack{1\leqslant i\leqslant n\\ \omega\models A_i\overline{B}_i}} \kappa_{\oplus i}$$

folgt

$$\kappa_{\oplus j}^{(-)} > \min_{\omega\models A_j\wedge B_j} \sum_{\substack{i\neq j\\ \omega\models A_i\overline{B}_i}} \kappa_{\oplus i}^{(-)} - \min_{\omega\models A_j\wedge\neg B_j} \sum_{\substack{i\neq j\\ \omega\models A_i\overline{B}_i}} \kappa_{\oplus i}^{(-)}$$

und damit die Behauptung. $\square$

9.4 BSS-Ansatz und konditionale Indifferenz

Um nachzuweisen, dass die Belief- und Plausibilitätsfunktionen des BSS-Ansatzes konditional indifferent bezüglich der zugrundeliegenden Defaultmenge Δ sind, kann man eine abgeschwächte Form der konditionalen Indifferenz definieren. In dieser Definition dürfen in den verallgemeinerten Welten von Definition 3.0.6 nur nicht-negative Exponenten vorkommen, so dass die $\hat{\omega}$ eine Halbgruppe bilden. Damit wird der wesentliche Aspekt der konditionalen Indifferenz eingefangen, ohne dass eine Gruppenstruktur mit der Existenz von Inversen im Wertebereich notwendig ist. Die Fortsetzung von $\mathrm{pl}_{\oplus\varepsilon}$ auf $\hat{\Omega}$ wird folgendermaßen definiert:

Definition 9.4.1 (Fortsetzung von $\mathrm{pl}_{\oplus\varepsilon}$ auf $\hat{\Omega}$)
Eine infinitesimale Plausibilitätsfunktion $\mathrm{pl}_{\oplus\varepsilon} : \Omega \to \mathbb{E}$ wird zu einem Homomorphismus $\mathrm{pl}_{\oplus} : \hat{\Omega}^{\times} \to \hat{\mathbb{E}}$ fortgesetzt durch

$$\mathrm{pl}_{\oplus\varepsilon}(\hat{\omega}) \approx_{\infty} \mathrm{pl}_{\oplus\varepsilon}(\omega_1^{r_1} \cdots \omega_r^{r_n}) \approx_{\infty} \mathrm{pl}_{\oplus\varepsilon}(\omega_1)^{r_1} \dots \mathrm{pl}_{\oplus\varepsilon}(\omega_k)^{r_k}$$

[5] [1] Bedingung 12

Die Definition der konditionalen Indifferenz für infinitesimale Plausibilitätsfunktionen lautet:

Definition 9.4.2 (Konditionale Indifferenz v.BSS-Plausifunktionen)
Eine Plausibilitätsfunktion $\mathrm{pl}_{\oplus\varepsilon}$ ist konditional indifferent in bezug auf eine Defaultmenge Δ genau dann, wenn $\mathrm{pl}_{\oplus\varepsilon}(\hat{\omega}_1) \approx_\infty \mathrm{pl}_{\oplus\varepsilon}(\hat{\omega}_2)$ immer dann, wenn $\sigma_\Delta(\hat{\omega}_1) = \sigma_\Delta(\hat{\omega}_2)$ für $\hat{\omega}_1 \equiv_\top \hat{\omega}_1 \in \hat{\Omega}$.

Da laut Lemma 9.3 gilt

$$\mathrm{pl}_{\oplus\varepsilon}(\omega) \approx_\infty \prod_{\omega \not\models \varphi_i} \varepsilon_i$$

kann man den Beweis für die konditionale Indifferenz der finitesimalen Funktionen von Satz 8.3.1 in diesem Rahmenwerk in Analogie zum endlichen Bereich direkt auf die infinitesimalen Funktionen des BSS-Ansatzes übertragen, indem man für alle s_i die infinitesimalen Funktionen ε_i einsetzt und alle Gleichheitszeichen durch $\approx_\infty$ ersetzt. Analog zu Satz 8.3.3 kann man die konditionale Indifferenz der infinitesimalen Belief-Funktionen des BSS-Ansatzes zeigen.

Wenn man die komplette konditionale Indifferenz der Funktionen des BSS-Ansatzes zeigen möchte, d.h. wenn in den verallgemeinerten Welten $\hat{\omega}$ auch negative Exponenten vorkommen dürfen, dann muss man zunächst eine Gruppe $\hat{\mathbb{E}}$ bilden, zu der man die Inversen aller Elemente aus $\mathbb{E}$ hinzunimmt:

Definition 9.4.3 ($\hat{\mathbb{E}}$)
$\hat{\mathbb{E}} := \mathbb{E}\backslash 0 \cup \{\frac{1}{\varepsilon} \mid \varepsilon \in \mathbb{E}^0\} \cup \{\frac{1}{1-\varepsilon} \mid 1-\varepsilon \in \mathbb{E}^1\}$ ist die Menge $\mathbb{E}\backslash 0$, ergänzt um die Inversen aller Elemente aus $\mathbb{E}\backslash 0$.

Wegen $\lim_{\eta\to 0} \varepsilon(\eta) = 0$ für alle $\varepsilon \in \mathbb{E}^0$ gilt $\lim_{\eta\to 0} \frac{1}{\varepsilon(\eta)} = \infty \;\forall \varepsilon \in \mathbb{E}^0$. Die Ordnung für die Inversen der Elemente aus $\mathbb{E}$ wird in Analogie zu Definition 9.1.1 auf $\mathbb{Z}$ fortgesetzt:

Definition 9.4.4 (Negative Ordnung der Elemente aus $\hat{\mathbb{E}}$)
Eine negative ganze Zahl **k** wird Ordnung von f genannt, bezeichnet durch

$$\kappa(f) = \mathbf{k} \text{ genau dann, wenn } \lim_{\eta \to 0} \frac{f(\eta)}{\eta^{\mathbf{k}}} \in \mathbb{R} - \{0\}$$

wenn ein solcher Limes existiert.

Aus $\lim_{\eta \to 0} \frac{f(\eta)}{\eta^k} \in \mathbb{R} - \{0\}$ folgt $\lim_{\eta \to 0} \frac{\frac{1}{f(\eta)}}{\eta^{-k}} = \lim_{\eta \to 0} \frac{\eta^k}{f(\eta)} \in \mathbb{R} - \{0\}$, d.h. aus $\kappa(f) = k$ folgt $\kappa(\frac{1}{f}) = -k$, so dass die Ordnung aller Inversen der Elemente f aus $\mathbb{E}$ das Negative der Ordnung von f ist.

Die Verknüpfung in dieser Gruppe $\hat{\mathbb{E}}$ wird als Multiplikation der Funktionen aus $\hat{\mathbb{E}}$ definiert.

Falls $\kappa(f_1) = a$ und $\kappa(f_2) = b$ gilt, dann folgt $\kappa(\frac{f_1}{f_2}) = a - b$ wegen $\lim_{\eta \to 0} \frac{f_1 \cdot \eta^b}{f_2 \cdot \eta^a} \in \mathbb{R} - \{0\}$.

Die Elemente aus $\hat{\mathbb{E}}$ bilden eine multiplikative Gruppe mit neutralem Element 1, und die Ordnungen $\kappa(f)$ der Elemente f aus $\hat{\mathbb{E}}$ bilden als Fortsetzung der $\kappa_\oplus$ aus Abschnitt 9.3 die additive Gruppe $(\mathbb{Z}, +)$ mit neutralem Element 0.

Die Definition 9.1.4 für den Vergleich von Infinitesimalen wird für die Gruppe $\hat{\mathbb{E}}$ identisch übernommen, und die Definitionen 9.4.2 und 9.4.1 gelten hier genauso wie bei der obigen abgeschwächten Form der konditionalen Indifferenz. Der einzige Unterschied zu dieser abgeschwächten Form ist, dass nun für die $\hat{\omega}$ auch negative Exponenten zugelassen werden, so dass $\hat{\Omega}$ nun eine Gruppe bildet. Der Beweis für die Belief- und Plausibilitätsfunktionen, die man aus diesem Ansatz erhält, erfolgt wieder in Analogie zum endlichen Bereich, indem man alle s_i durch ε_i und alle Gleichheitszeichen durch $\approx_\infty$ ersetzt.

Kapitel 10

LCD

10.1 Einführung in LCD

LCD ist eine Logik, die auf der Dempster'schen Kombinationsregel und den infinitesimalen Belief- bzw. den entsprechenden Plausibilitätsfunktionen des BSS-Ansatzes basiert und folgende Zusatzbedingungen an die Plausibilitätsfunktionen stellt:

1. Das Autodeduktions-Prinzip muss erfüllt sein. Das bedeutet, dass alle Defaults ableitbar sein müssen.

2. Das least-commitment-Prinzip, das in Definition 10.3.3 angegeben wird, muss erfüllt sein.

LCD steht für „**l**east **c**ommitment" und „**D**empster". Least-commitment bedeutet „geringste Festlegung". Eine Logik, die dem least-commitment-Prinzip entspricht, ist durch die vorhandenen Defaults nur insoweit festgelegt, dass alle Defaults ableitbar sind, und leitet nur solche Schlussfolgerungen ab, die durch die Defaultbasis definitiv festgelegt sind, d.h. es wird nur die Information verwendet, die sicher gewusst wird.

Laut Gleichung (9.1) wird in dem BSS-Ansatz jedem Default $d_i \in \Delta$ eine simple-support-Funktion m_i mit $\mathrm{m}_i(\Omega) = \varepsilon_i$, $\mathrm{m}_i(\varphi_i) = 1 - \varepsilon_i$ und $\mathrm{m}_i(X) = 0$ für alle $X \neq \Omega$ und $X \neq \varphi_i$ zugeordnet, wobei $\varepsilon_i \in \mathbb{E}$ gilt. Diese simple-support-Funktionen m_i werden durch die Dempster'sche Kombinationsregel zu einem Basismaß $\mathrm{m}_\oplus$ kombiniert. Wenn die Menge $\mathcal{E} := \{\varepsilon_1, \ldots, \varepsilon_n\}$, die der Menge $\Delta := \{d_1, \ldots, d_n\}$ zugeordnet wurde, dem least-commitment-Prinzip entspricht, dann sollte diese Menge $\mathcal{E}$ „so frei wie möglich" bezüglich der Ordnung $<_\infty$ gewählt werden.[1] Das bedeutet, dass die Elemente aus $\mathcal{E}$ unter Erhaltung des Autodeduktions-Prinzips „so groß wie möglich" gewählt werden müssen.[2]

Da in dieser Arbeit gezeigt werden soll, dass c-Repräsentationen auch eine Basis-Semantik für LCD liefern, werden die Bedingungen des BSS-Ansatzes für LCD immer gleich auf die entsprechenden konditional indifferenten OCFs übertragen.

10.2 Autodeduktions-Prinzip

Das Autodeduktions-Prinzip[3] bedeutet für die BSS-Plausibilitätsfunktionen, dass für alle Defaults $d := \alpha \rightarrow \beta$ die folgende Bedingung erfüllt ist:

$$C_d : \max_{\omega \models \alpha \wedge \beta} \prod_{\omega \not\models \varphi_i} \varepsilon_i >_\infty \max_{\omega \models \alpha \wedge \neg \beta} \prod_{\omega \not\models \varphi_i} \varepsilon_i$$

Diese Bedingungen des Autodeduktions-Prinzips für alle Defaults $d \in \Delta$ werden im folgenden in der Menge $C_\Delta := \{C_d \mid d \in \Delta\}$ zusammengefasst.

Laut Satz 9.3.1 gewinnt man aus den Belief-Funktionen des BSS-Ansatzes entsprechende c-Repräsentationen durch $\kappa_{\oplus i}^- := \kappa(\varepsilon_i)$. Wenn man das Autodeduktions-Prinzip auf eine konditional indifferente OCF mit $\kappa_i^- := \kappa(\varepsilon_i)$

[1] vgl.[2] S.26
[2] vgl.[2]S.27
[3] vgl.[2] Abschnitt 6.1

überträgt, dann erhält man für alle Defaults $d_j = A_j \rightarrow B_j \in \Delta$ die Ungleichung (9.5):

$$C_j : \kappa_j^- > \min_{\omega \models A_j B_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \kappa_i^- - \min_{\omega \models A_j \overline{B}_j} \sum_{\substack{i \neq j \\ \omega \models A_i \overline{B}_i}} \kappa_i^-$$

Das Autodeduktions-Prinzip wird auch von allen oben vorgestellten Logiken erfüllt, wenn es in diesen Logiken Modelle bzw. entsprechende c-Repräsentationen für Δ gibt.

Beispiel 10.2.1 (LCD, Autodeduktion)

Wenn man wieder die Regeln $d_1 = v \rightarrow f$, $d_2 = p \rightarrow v$ *und* $d_3 = p \rightarrow \neg f$ *hat, dann müssen wegen des Autodeduktions-Prinzips (Gleichung 9.5) für die* κ_i^- *folgende Bedingungen gelten:*

Bedingung 1:

$$\kappa_1^- > \min_{\omega \models v \wedge f} \sum_{\substack{j=2,3 \\ \omega \models A_j \overline{B}_j}} \kappa_j^- - \min_{\omega \models v \wedge \neg f} \sum_{\substack{j=2,3 \\ \omega \models A_j \overline{B}_j}} \kappa_j^-$$

Da die Welt $v \wedge f \wedge \neg p$ *allen drei Defaults genügt und die Welt* $v \wedge \neg f \wedge \neg p$ *die Bedingung* $v \wedge \neg f$ *verifiziert und keinen anderen Default* d_2 *oder* d_3 *falsifiziert, folgt* $\kappa_1^- > 0 - 0$.

Bedingung 2:

$$\kappa_2^- > \min_{\omega \models p \wedge v} \sum_{\substack{j=1,3 \\ \omega \models A_j \overline{B}_j}} \kappa_j^- - \min_{\omega \models p \wedge \neg v} \sum_{\substack{j=1,3 \\ \omega \models A_j \overline{B}_j}} \kappa_j^-$$

Eine Welt, die $p \wedge v$ *verifiziert, muss die Variable* f *entweder mit* $\top$ *oder mit* $\perp$ *belegen, und falsifiziert damit entweder* d_3 *oder* d_1*. Die Welt* $\neg v \wedge p \wedge \neg f$ *verifiziert die Bedingung* $p \wedge \neg v$ *und falsifiziert keinen der Defaults* d_1 *oder* d_3*. Daraus folgt* $\kappa_2^- > \min\{\kappa_1^-, \kappa_3^-\} - 0$

Bedingung 3:

$$\kappa_3^- > \min_{\omega \models p \wedge \neg f} \sum_{\substack{j=1,2 \\ \omega \models A_j \overline{B}_j}} \kappa_j^- - \min_{\omega \models p \wedge f} \sum_{\substack{j=1,2 \\ \omega \models A_j \overline{B}_j}} \kappa_j^-$$

Für κ_3^- *erhält man analog:* $\kappa_3^- > \min\{\kappa_1^-, \kappa_2^-\}$.
Diese Bedingungen erfüllen z.B. die Wertzuweisungen $\kappa_1^- := 3$, $\kappa_2^- := 5$ *und* $\kappa_3^- := 7$ *oder auch* $\kappa_1^- := 1$, $\kappa_2^- := 2$ *und* $\kappa_3^- := 2$. *Die Werte der einzelnen* κ_i^- *sind also durch das Autodeduktionsprinzip noch nicht genau festgelegt, aber es müssen gewisse Bedingungen an die Ordnung der Defaults erfüllt sein. Z.B. kann* κ_1^- *nicht größer als* κ_2^- *oder* κ_3^- *sein, da sonst die Ungleichungen* $\kappa_2^- > \min\{\kappa_1^-, \kappa_3^-\}$ *und* $\kappa_3^- > \min\{\kappa_1^-, \kappa_2^-\}$ *nicht gleichzeitig erfüllbar wären.*

Um die Parallelen von LCD-Belief-Funktionen und konditionalen OCFs zu veranschaulichen, wird jetzt ein Beispiel aus [2] zitiert und auf OCFs übertragen.[4]

Beispiel 10.2.2 (LCD)
Angenommen, die Sprache $\mathcal{L}$ *besteht aus den Symbolen p (Pinguin), ft (gefiedertes Tier), v (Vogel) und f (Fliegt), also* $\mathcal{L} = \{p, ft, v, f\}$. *Die Default-Regeln seien*

$$\begin{aligned} d_1 &= p \rightarrow ft \vee v \\ d_2 &= ft \rightarrow v \\ d_3 &= v \rightarrow ft \\ d_4 &= p \rightarrow \neg f \text{ und} \\ d_5 &= v \rightarrow f \end{aligned}$$

Aus Gleichung 9.5 ergibt sich:

$$\begin{aligned} \kappa_1^- &> \min\{\kappa_3^- + \kappa_5^-, \kappa_3^- + \kappa_4^-, \kappa_2^-, \kappa_2^- + \kappa_4^-, \kappa_5^-, \kappa_4^-\} \\ &\quad - \min\{0, \kappa_4^-\} \\ \kappa_2^- &> \min\{\kappa_4^-, \kappa_5^-, 0\} - \min\{0, \kappa_4^-\} \end{aligned}$$

[4] vgl.[2] Beispiel 3

$$\begin{aligned}
\kappa_3^- &> \min\{\kappa_4^-, \kappa_5^-, 0\} - \min\{\kappa_5^-, 0, \kappa_4^-\} \\
\kappa_4^- &> \min\{\kappa_1^-, \kappa_3^- + \kappa_5^-, \kappa_2^-, \kappa_5^-\} \\
&\qquad - \min\{\kappa_1^-, \kappa_3^-, \kappa_2^-, 0\} \\
\kappa_5^- &> \min\{\kappa_3^-, 0, \kappa_3^- + \kappa_4^-, \kappa_4^-\} - \min\{\kappa_3^-, 0\}
\end{aligned}$$

Diese Ungleichungen kann man vereinfachen zu

$$\kappa_1^- > \min\{\kappa_2^-, \kappa_4^-, \kappa_5^-\} \tag{10.1}$$
$$\kappa_2^- > 0$$
$$\kappa_3^- > 0$$
$$\kappa_4^- > \min\{\kappa_1^-, \kappa_2^-, \kappa_5^-\} \tag{10.2}$$
$$\kappa_5^- > 0$$

Es gibt viele verschiedene Lösungen dieser Ungleichungen, z.B.

$$K_1 := \{\kappa_1^- := 4, \kappa_2^- := 1, \kappa_3^- := 1, \kappa_4^- := 2, \kappa_5^- := 3\}$$

oder auch

$$K_2 := \{\kappa_1^- := 2, \kappa_2^- := 1, \kappa_3^- := 1, \kappa_4^- := 2, \kappa_5^- := 1\}$$

Wenn man jetzt die Frage stellt, ob Pinguine, die Federtiere und Vögel sind, fliegen können, ergeben sich zunächst z.B. aus Gleichung 3.3 die Gleichungen $\kappa(p \wedge ft \wedge v \wedge f) = \kappa_4^-$ *und* $\kappa(p \wedge ft \wedge v \wedge \neg f) = \kappa_5^-$. *Aus den Wertzuweisungen von* K_1 *ergibt sich* $\kappa_4^- < \kappa_5^-$, *und in* K_2 *gilt* $\kappa_4^- > \kappa_5^-$. *Aus* K_1 *würde man daher folgern, dass Pinguine, die gefiederte Tiere und Vögel sind, fliegen können, und aus* K_2 *folgert man das Gegenteil, nämlich dass diese Pinguine nicht fliegen können.*

Da die Schlussfolgerung „Pinguine können fliegen“ aus K_1 nicht wünschenswert ist, sieht man anhand dieses Beispiels, dass das Autodeduktions-Prinzip allein noch nicht ausreicht, um zufriedenstellende Schlussfolgerungen zu ziehen.

10.3 Least-commitment-Prinzip

In [2] wird deshalb das „least commitment-Prinzip“ eingeführt, das eine gegebene Default-Basis in verschiedene Ordnungs-Klassen aufteilt. Wenn die infinitesimalen $\varepsilon_i \in \mathcal{E}$ unter Einhaltung des Autodeduktions-Prinzips so groß wie möglich gewählt werden, dann ergibt sich aus Lemma 9.1.1(2), dass die entsprechenden $\kappa_i^- = \kappa(\varepsilon_i)$ unter den entsprechenden Bedingungen so klein wie möglich sein müssen.

Als Grundlage des Least-commitment-Prinzips wird zunächst jedem Default $d_i \in \Delta$ eine Variable ζ_i zugeordnet. Diese ζ_i werden in [2] „Proto-Infinitesimale“ genannt, da ihnen später infinitesimale Funktionen zugeordnet werden. Sei $Z = \{\zeta_1, \ldots, \zeta_n\}$ eine solche Menge von Proto-Infinitesimalen, die der gegebenen Default-Menge $\Delta = \{d_1, \ldots, d_n\}$ entspricht. Eine Interpretation $\mathcal{I}$ ist eine Abbildung der Proto-Infinitesimale auf konkrete Infinitesimale. Sei also $\varepsilon_i := \mathcal{I}(\zeta_i)$ die infinitesimale Funktion, die $\mathcal{I}$ dem Proto-Infinitesimal ζ_i zuordnet, und $\mathcal{I}(Z)$ sei die Menge $\mathcal{E} = \{\varepsilon_1, \ldots, \varepsilon_n\}$ mit $\varepsilon_i = \mathcal{I}(\zeta_i) \in \mathbb{E}^0$ für alle $\zeta_i \in Z$.

Definition 10.3.1 (ε-Stratifikation[5])
Sei $Z = \{\zeta_1, \ldots, \zeta_n\}$ eine Menge von Proto-Infinitesimalen. Eine ε-Stratifikation ist definiert als Partition $\xi_0, \ldots, \xi_m$ von Z (Partition bedeutet $\xi_0 \cup \ldots \cup \xi_m = Z$, $\forall i \neq j, \xi_i \cap \xi_j = \emptyset$ und $\forall i, \xi_i \neq \emptyset$). Eine Interpretation $\mathcal{I}(Z)$ erfüllt eine ε-Stratifikation $\xi = \{\xi_0, \ldots, \xi_m\}$ von Z, geschrieben als

$$\mathcal{I} \models \xi$$

wenn für alle $\zeta_i, \zeta_j \in Z$ mit $\zeta_i \in \xi_i$ und $\zeta_j \in \xi_j$, aus i<j folgt: $\mathcal{I}(\zeta_i) >_\infty \mathcal{I}(\zeta_j)$.

Wenn man die obige Definition auf konditional indifferente OCFs überträgt, dann folgt aus $\mathcal{I}(\zeta_i) >_\infty \mathcal{I}(\zeta_j)$: $\kappa(\mathcal{I}(\zeta_i)) < \kappa(\mathcal{I}(\zeta_j))$. Das bedeutet, dass die Infinitesimale in einer ε-Stratifikation absteigend und die entsprechenden Ordnungen der Infinitesimale aufsteigend angeordnet sind.

Da LCD-Funktionen das Autodeduktions-Prinzip erfüllen sollen, wird jetzt die Verbindung von ε-Stratifikationen und diesem Prinzip festgelegt:

Definition 10.3.2 (ε-Stratifikation und Autodeduktion[6])
Sei $Z = \zeta_1, \ldots, \zeta_n$ eine Menge von Proto-Infinitesimalen. Eine Interpretation $\mathcal{I}(Z)$ erfüllt eine Menge von Bedingungen C_Δ, geschrieben

$$\mathcal{I} \models C_\Delta$$

genau dann, wenn alle Bedingungen $C_d \in C_\Delta$ in $\mathcal{I}$ wahr sind.
Eine ε-Stratifikation ξ von Z erfüllt C_Δ, wenn jede Interpretation $\mathcal{I}$, die ξ erfüllt, auch C_Δ erfüllt. Eine ε-Stratifikation ξ von Z ist kompatibel mit C_Δ, wenn es eine Interpretation gibt, die ξ und C_Δ erfüllt.

Eine least-commited-ε-Stratifikation wird in in [2] folgendermaßen definiert:

[5] vgl.[2] Definition 10
[6] vgl.[2] Def.11

Definition 10.3.3 (Least-commitment-Prinzip)
([2] Definition 12) Sei $Z = \{\zeta_1, \ldots, \zeta_n\}$ eine Menge von Proto-Infinitesimalen, und seien $\xi = \{\xi_0, \ldots, \xi_m\}$ und $\xi' = \{\xi'_0, \ldots, \xi'_m\}$ zwei ε-Stratifikationen von Z. ξ ist

„weniger bestimmt“ (less commited)

als ξ' genau dann, wenn für alle $\zeta_i \in Z$ aus $\zeta_i \in \xi_j$ und $\zeta_i \in \xi'_k$ folgt: $j \leq k$. ξ ist

„strikt weniger bestimmt“

als ξ', wenn ξ weniger bestimmt als ξ' ist und mindestens eine der obigen Ungleichungen $j \leq k$ echt ist, also $j < k$. Wenn $\mathcal{P}$ eine Familie von ε-Stratifikationen von Z ist, dann wird ein Element ξ_{lc} aus $\mathcal{P}$ als

„am wenigsten bestimmt“ (least commited)

in $\mathcal{P}$ bezeichnet, wenn es in $\mathcal{P}$ kein $\xi' \neq \xi_{lc}$ gibt, das „strikt weniger bestimmt“ als ξ_{lc} ist.

Das Least-commitment-Prinzip minimiert also die Anzahl der Teil-Klassen ξ_i einer ε-Stratifikation ξ und ordnet so viele Proto-Infinitesimale wie möglich den Teil-Klassen ξ_i mit niedrigem Index i zu. Da die Interpretationen der Proto-Infinitesimale mit niedrigem Index infinitesimal größer als die Interpretationen der Proto-Infinitesimale mit höherem Index und die entsprechenden Ordnungen der Interpretationen mit niedrigem Index kleiner als die Ordnungen der Interpretationen von Proto-Infinitesimalen mit höherem Index sind, entspricht das Least-commitment-Prinzip einer Auswahl von Interpretationen, in denen die Infinitesimale „so groß wie möglich“ und ihre Ordnungen damit „so klein wie möglich“ sind.

Die Menge aller LCD-Belieffunktion $Bel_{lcd}(\Delta)$ ist definiert als die Menge aller infinitesimalen Belieffunktionen des BSS-Ansatzes, die aus einer Menge von Infinitesimalen aufgebaut werden können, die kompatibel mit allen Bedingungen C_Δ des Autodeduktions-Prinzips sind und die „least-commited“ sind, die also eine „least-commited“-Stratifikation bilden.[7]

Die LCD-Folgerungsrelation, die durch $\mid\!\sim_{lcd}$ notiert wird, ist definiert durch $\alpha \mid\!\sim_{lcd} \beta$ genau dann, wenn für alle $\mathrm{bel}_\varepsilon \in Bel_{lcd}(\Delta)$ gilt $\mathrm{bel}_\varepsilon \models \alpha \to \beta$.[8]

Nach Satz 9.3.3 gibt es zu jeder derartigen LCD-Belieffunktion eine c-Repräsentation κ_{lcd} mit $\mathrm{bel}_{lcd} \models A_i \to B_i$ genau dann, wenn $\kappa_{lcd} \models A_i \to B_i$. Als Folge von Satz 9.3.3 bedeutet die Definition 10.3.3 für OCFs: Das least-commitment-Prinzip teilt alle κ_i^- unter Erhaltung des Autodeduktions-Prinzips der kleinstmöglichen Teilklasse einer Stratifikation zu. Die κ_i^- selbst müssen aber nicht minimal sein, wie man an Beispiel 10.4.1 sehen wird. Ausserdem müssen zwei Werte $\kappa_i, \kappa_j \in \xi_k$, die zu derselben Teilklasse ξ_k der Stratifikation gehören, nicht gleichgesetzt werden, so dass man aus $\kappa_i \in \xi_k$ und $\kappa_j \in \xi_k$ nicht $\kappa_i = \kappa_j$ ableiten kann.

Beispiel 10.3.1 (LCD)

In dem obigen Beispiel 10.2.2 gilt

$$\begin{aligned} \xi &= \xi_1 \cup \xi_2 \cup \xi_3 \cup \xi_4 = \{\kappa_2^-, \kappa_3^-\}, \{\kappa_4^-\}, \{\kappa_5^-\}, \{\kappa_1^-\} \\ \text{und } \xi' &= \xi_1' \cup \xi_2' = \{\kappa_2^-, \kappa_3^-, \kappa_5^-\}, \{\kappa_1^-, \kappa_4^-\} \end{aligned}$$

Alle κ_i^- werden in ξ' in niedrigere Klassen eingeteilt als in ξ, z.B. gilt für $\kappa_1^- \in \xi_4$ und $\kappa_1^- \in \xi_2'$: 4>2, und daher ist ξ' weniger bestimmt als ξ. Das least-commitment-Prinzip und das Autodeduktions-Prinzip machen aber keine Aussagen über das Verhältnis der Defaults in einer Klasse, z.B. kann für $\{\kappa_2^-, \kappa_3^-\} \in \xi_1'$ gelten:

$$\kappa_2^- > \kappa_3^-,\ \kappa_2^- = \kappa_3^- \text{ oder } \kappa_2^- < \kappa_3^-$$

Aus den gebenen Defaults lassen sich keine Aussagen darüber ableiten, ob der Default $d_2 = ft \to v$ eine größere, kleinere oder gleich große Priorität wie der

[7] vgl.[2] S.32

[8] vgl.[2] Abschnitt 6.3

Default $d_3 = v \rightarrow ft$ *hat. Auch über Summen der* κ_i *einer Klasse kann man keine Aussagen machen. Es kann z.B. für* $\{\kappa_2^-, \kappa_3^-\} \in \xi_1'$ *und* $\kappa_4^- \in \xi_2'$ *gelten:*

$$\kappa_2^- + \kappa_3^- < \kappa_4^-,\ \kappa_2^- + \kappa_3^- > \kappa_4^- \text{ oder } \kappa_2^- + \kappa_3^- = \kappa_4^-$$

Man kann nur Aussagen über Elemente verschiedener Teilklassen ableiten. Z.B. sind alle Elemente aus ξ_1' *kleiner als alle Elemente aus* ξ_2'*. Damit lässt sich auch für* $\kappa_2^- \in \xi_1'$ *und* $\{\kappa_1^-, \kappa_4^-\} \in \xi_2'$ *ableiten, dass auch* $\kappa_2^- < \kappa_4^- + \kappa_1^-$ *gelten muss. Es lässt sich beweisen, dass* $\xi' = \xi_1' \cup \xi_2' = \{\kappa_2^-, \kappa_3^-, \kappa_5^-\}, \{\kappa_1^-, \kappa_4^-\}$ *die Stratifikation ist, die von allen Partitionen, in die sich* ξ *aufteilen lässt und die das Autodeduktions-Prinzip bezüglich der gegebenen Defaults erfüllen, die am wenigsten bestimmte ist. Wenn es eine Stratifikation gäbe, die noch weniger festgelegt bzw. bestimmt als* ξ' *ist, dann müsste in dieser Stratifikation* κ_1^- *oder* κ_4^- *zu einer kleineren als der zweiten Teilklasse* ξ_2'*, d.h. zu der ersten Teilklasse* ξ_1' *gehören. Mit dieser Stratifikation wäre aber das Autodeduktions-Prinzip verletzt, da eine der beiden Ungleichungen (10.1) oder (10.2) aus dieser Klassenaufteilung nicht mehr abgeleitet werden kann.*

Aus dieser least-commited-Stratifikation schließt man wegen $\kappa_5^- < \kappa_4^-$ *wie gewünscht, dass Pinguine, die gefiederte Tiere und Vögel sind, nicht fliegen können.*

10.4 LCD und konditionale Indifferenz

An dem vorhergehenden Beispiel sieht man, dass LCD-Belief-Funktionen durch konditional indifferente OCFs beschrieben werden können.

Satz 10.4.1 (LCD und konditionale Indifferenz)
Zu allen infinitesimalen LCD-Belief-Funktionen des BSS-Ansatzes gibt es eine c-Repräsentation κ_{lcd} mit

$$\mathrm{bel}_{lcd} \models \alpha \to \beta \text{ genau dann, wenn } \kappa_{lcd} \models \alpha \to \beta$$

Beweis: Da LCD-Funktionen aus Belief-Funktionen und der Dempster'schen Kombinationsregel abgeleitet wurden, ergibt sich die Darstellbarkeit der LCD-Logik durch c-Repräsentationen aus Satz 9.3.3. Die c-Repräsentation κ_{lcd} zu einer LCD-Belief-Funktion bel_{lcd} erhält man durch $\kappa_i^- := \kappa(\varepsilon_i)$. Es folgt analog zu Satz 9.3.2

$$\mathrm{bel}_{lcd} \models \alpha \to \beta$$

genau dann, wenn

$$\max_{\omega \models \alpha\wedge\beta} \mathrm{pl}_{lcd}(\omega) >_\infty \max_{\omega \models \alpha\wedge\neg\beta} \mathrm{pl}_{lcd}(\omega)$$

genau dann, wenn

$$\max_{\omega \models \alpha\wedge\beta} \prod_{\omega \not\models \varphi_i} \varepsilon_i >_\infty \max_{\omega \models \alpha\wedge\neg\beta} \prod_{\omega \not\models \varphi_i} \varepsilon_i$$

genau dann, wenn

$$\min_{\omega \models \alpha\wedge\beta} \sum_{\omega \not\models \varphi_i} \kappa_i^- < \min_{\omega \models \alpha\wedge\neg\beta} \sum_{\omega \not\models \varphi_i} \kappa_i^-$$

genau dann, wenn

$$\kappa_{lcd} \models \alpha \to \beta$$

□

Es ist jedoch zu beachten, dass LCD keine vollständige Ordnung auf Ω induziert, da nicht allen Defaults d_i ein konkreter Wert κ_i zugeordnet werden muss und manche κ_i, κ_j nicht miteinander verglichen werden können.

In [2] wird gezeigt, dass LCD-Belieffunktionen die folgenden wünschenswerten Eigenschaften haben[9], wobei die Default-Basis in den Beispielen die bekannte Default-Menge $\{v \to f, p \to v, p \to \neg f\}$ sei:

- KLM-Rationalität:[10] beinhaltet die Mindestanforderungen an nichtmonotone Logiken, die aus folgenden Bedingungen bestehen:
  - Reflexivität: $\alpha \mathrel{|\!\sim} \alpha$
  - Links-Äquivalenz: aus $\alpha \models \alpha'$, $\alpha' \models \alpha$ und $\alpha \mathrel{|\!\sim} \beta$ folgt $\alpha' \mathrel{|\!\sim} \beta$
  - Rechts-Abschwächung: Aus $\beta \models \beta'$ und $\alpha \mathrel{|\!\sim} \beta$ folgt $\alpha \mathrel{|\!\sim} \beta'$
  - Oder: Aus $\alpha \mathrel{|\!\sim} \gamma$ und $\beta \mathrel{|\!\sim} \gamma$ folgt $\alpha \vee \beta \mathrel{|\!\sim} \gamma$
  - vorsichtige Monotonie: Aus $\alpha \mathrel{|\!\sim} \beta$ und $\alpha \mathrel{|\!\sim} \gamma$ folgt $\alpha \wedge \beta \mathrel{|\!\sim} \gamma$
  - Schnitt: Aus $\alpha \wedge \beta \mathrel{|\!\sim} \gamma$ und $\alpha \mathrel{|\!\sim} \beta$ folgt $\alpha \mathrel{|\!\sim} \gamma$
- Spezifizität: Da Pinguine eine Subklasse aller Vögel sind und die Eigenschaft „Pinguin“ damit spezifischer als die Eigenschaft „Vogel“ ist, sollte man aus „Vögel können fliegen“ und „Pinguine können nicht fliegen“ ableiten können, dass ein Pinguin, der auch ein Vogel ist, nicht fliegen kann. Aus $p \models v$, $p \mathrel{|\!\sim} \neg f$ und $v \mathrel{|\!\sim} f$ folgt $p \wedge v \mathrel{|\!\sim} \neg f$.
- Irrelevanz: Da die Eigenschaft „rot“ keinen Einfluss auf die Eigenschaft des Fliegen-Könnens hat und damit für Schlussfolgerungen, die das Fliegen-Können betreffen, irrelevant ist, sollte man aus „Vögel können fliegen“ auch ableiten können, dass rote Vögel fliegen können. Irrelevant sind alle Formeln, die aus Symbolen des zugrundeliegenden Alphabets bestehen, die in keinem Default vorkommen. Seien α, β Formeln, die aus Symbolen bestehen, die in Δ vorkommen, z.B. $\alpha := v, \beta := f$ mit $v \to f \in \Delta$, und sei γ eine irrelevante Formel, z.B.$\gamma := r$, wobei „r“ für „rot“ steht. Dann muss aus $\alpha \mathrel{|\!\sim}_{lcd} \beta$ auch $\alpha \wedge \gamma \mathrel{|\!\sim}_{lcd} \beta$ folgen.

[9] vgl.[2] Kap.7 und Tabelle 2,S.46

[10] KLM=Kraus,Lehmann,Magidor

- keine Blockierung von Vererbung: Aus der Eigenschaft „Vögel haben Beine“ sollte man ableiten können, dass auch Pinguine, die eine Subklasse der Vögel bilden, Beine haben, und damit diese Eigenschaft „erben“, da Pinguine nur beüglich der Eigenschaft „Fliegen-Können“ eine Ausnahme zu ihrer Superklasse „Vögel“ bilden. Aus der Aussage „Pinguine haben Beine“ kann man keinen Widerspruch zur Default-Basis ableiten, und solche Eigenschaften, zu denen nichts Gegenteiliges bekannt ist, sollten vererbt werden.

  In [2] Abschnitt 7.3 werden alle Defaults, die in Verbindung mit einer Formel α keine Inkonsistenz verursachen, in einer Menge Free($\Delta \cup \alpha$) zusammengefasst, und die Nicht-Blockierung von Vererbung kann dann so formalisiert werden:

  $$\text{wenn } \alpha \wedge Free(\Delta \cup \{\alpha\}) \vdash \beta, \text{ dann } \alpha {\mid\!\sim}_{lcd} \beta$$

  Zum Beispiel gilt, wenn man den Default $v \rightarrow bein$ zu Δ hinzunimmt, $v \rightarrow bein \in Free(\Delta \cup p)$, aber $v \rightarrow f \notin Free(\Delta \cup p)$ (wegen Widerspruch zu $p \rightarrow \neg f$)

- Ambiguitätsbewahrung: Aus den Aussagen „Quäker sind Pazifisten“ „Republikaner sind keine Pazifisten“ und „Nixon ist Quäker und Republikaner“ sollte man weder ableiten können, dass Nixon Pazifist ist, noch, dass Nixon kein Pazifist ist. Die Ambiguitätsbewahrung kann folgendermaßen formalisiert werden: Aus $\alpha {\mid\!\sim} \gamma$ und $\beta {\mid\!\sim} \neg\gamma$, wobei α und β unabhängig voneinander sind, also β z.B. keine Spezifizierung von α ist, sollte weder $\alpha \wedge \beta {\mid\!\sim} \gamma$ noch $\alpha \wedge \beta {\mid\!\sim} \neg\gamma$ ableitbar sein.

- Syntax-Unabhängigkeit: Schlussfolgerungen sollten nicht von der syntaktischen Repräsentation der Default-Basis abhängen. Wenn man z.B. eine Default-Regel einfach verdoppelt, sollte das keinen Einfluss auf die Schlussfolgerungen aus der Default-Basis haben. Wenn beispielsweise in dem vorigen Beispiel q für Quäker und r für Republikaner steht, dann sollten aus $\Delta := \{q \rightarrow p, r \rightarrow \neg p\}$ dieselben Schlussfolgerungen wie aus $\Delta' := \{q \to p, q \to p, r \to \neg p\}$ gezogen werden. Die Regel „Quäker sind Pazifisten“ wird durch Verdopplung nicht wichtiger als die Regel „Republikaner sind keine Pazifisten“.

Diese Eigenschaften werden nach Satz 10.4.1 direkt auf LCD-c-Repräsentationen übertragen, da aus den κ_{lcd}-Funktionen, die man aus den bel_{lcd}-Funktionen gewinnt, genau dieselben Folgerungen wie aus der jeweils zugrundeliegenden bel_{lcd}-Funktion abgeleitet werden.

Alle anderen Logiken aus den Kapiteln 4 bis 7 ausser der von Geffner haben mindestens eine dieser Eigenschaften nicht.

Aber auch die Geffner-Logik unterscheidet sich von LCD. In [2] Beispiel 11 wird hierzu folgendes Beispiel angegeben:

Beispiel 10.4.1 (LCD und andere Logiken)
Die Default-Basis sei $\Delta = \{d_1 = \top \to a, d_2 = y \to \neg a \wedge c, d_3 = y \wedge s \to \neg c\}$. *Es gibt eine Welt* ω_1, *z.B.* $\omega_1 := a \wedge c \wedge s \wedge \neg y$, *die* d_1 *verifiziert und keinen der Defaults* d_2, d_3 *falsifiziert. Jede Welt* ω_2, *die* d_2 *verifiziert, falsifiziert* d_1, *und es gibt eine Welt* $\omega_2' := y \wedge a \wedge \neg c \wedge s$, *die* d_2 *falsifiziert und keinen der Defaults aus* $\{d_1, d_3\}$ *falsifiziert. Jede Welt* ω_3, *die* d_3 *verifziert, falsifiziert* d_2, *und jede Welt* ω_3', *die* d_3 *falsifiziert, falsifiziert* d_1 *oder* d_2. *Deshalb müssen alle c-Repräsentationen aus den Kapiteln 4 bis 7 wegen des Autodeduktions-Prinzips (Gleichung 9.5) dem Default* d_2 *eine höhere Priorität als dem Default* d_1 *zuweisen.*

$$\kappa_1^- > 0 \tag{10.3}$$
$$\kappa_2^- > \kappa_1^- \tag{10.4}$$
$$\kappa_3^- > \kappa_2^- - \min\{\kappa_1^-, \kappa_2^-\} = \kappa_2^- - \kappa_1^-$$
$$\Rightarrow \kappa_3^- + \kappa_1^- > \kappa_2^- \tag{10.5}$$

Diese Bedingungen werden z.B. durch die Wertzuweisungen $\kappa_1^- := 1$, $\kappa_2^- := 2$ *und* $\kappa_3^- := 3$ *erfüllt, so dass es für alle Logiken aus den Kapiteln 4 bis 7 eine c-Repräsentation auf der Stratifikation* $\Delta := \{\Delta_1, \Delta_2, \Delta_3\} = \{\{d_1\}, \{d_2\}, \{d_3\}\}$

gibt. Wenn man nun wissen möchte, ob aus der Formel $y \wedge s \wedge \neg a$ *folgt:* c *oder* $\neg c$*, dann muss man die zwei Welten* $\omega_c := y \wedge s \wedge \neg a \wedge c$ *und* $\omega_{\overline{c}} := y \wedge s \wedge \neg a \wedge \neg c$ *miteinander vergleichen. Beide Welten falsifizieren* d_1*,* ω_c *verifiziert* d_2 *und falsifiziert* d_3*,* $\omega_{\overline{c}}$ *falsifiziert* d_2 *und verifiziert* d_3*. Es folgt* $\kappa(\omega_c) = \kappa_1^- + \kappa_3^-$ *und* $\kappa(\omega_{\overline{c}}) = \kappa_1^- + \kappa_2^-$*. Alle c-Repräsentationen, die den Logiken aus den Kapiteln 4 bis 7 entsprechen, können aus der Formel* $y \wedge s \wedge \neg a$ *mit der obigen Stratifikation* $\Delta = \{\{d_1\}, \{d_2\}, \{d_3\}\}$ *wegen* $\kappa_3^- > \kappa_2^-$ *die Formel* $\neg c$ *aus* Δ *ableiten, während diese Ableitung in LCD nicht möglich ist:*

Die „least-commited"-Stratifikation der Default-Menge Δ *ist die Stratifikation* $\Delta_{lc} := \{\Delta_1, \Delta_2\} = \{\{d_1, d_3\}, \{d_2\}\}$*. Aus Gleichung (10.4) ergibt sich, dass* κ_2^- *einer höheren Teilklasse als* κ_1^- *zugeordnet werden muss. Die Gleichung (10.5) ist z.B. mit den Wertzuweisungen* $\kappa_1^- := \kappa_3^- := 2$ *und* $\kappa_2^- := 3$ *erfüllbar, so dass* κ_3^- *derselben Teilklasse* Δ_1 *wie* κ_1^- *zugeordnet werden kann. Wegen Gleichung (10.4) gibt es für* Δ *auch keine „less commited"-Stratifikation als* Δ_{lc}*, denn eine solche Stratifikation müsste auch* d_2 *der Teilklasse* Δ_1 *zuordnen. Dies ist wegen* $\kappa_2^- > \kappa_1^-$ *nicht möglich.*

Man sieht an diesem Beispiel, dass die einzelnen κ_i *nicht unbedingt minimal sein müssen. Wenn man die* κ_i *bezüglich der Stratifikation* Δ_{lc} *minimal wählen würde, dann müsste für die Stratifikation* $\{\{d_1, d_3\}, \{d_2\}\}$ *gelten* $\kappa_1^- = \kappa_3^- = 1$ *und* $\kappa_2^- = 2$*, aber mit diesen Werten wäre die Bedingung* $\kappa_1^- + \kappa_3^- > \kappa_2^-$ *nicht mehr erfüllt. Wenn man die* κ_i^- *bezüglich des Autodeduktions-Prinzips minimal wählen würde, dann wäre* $\kappa_1^- = 1$ *wegen (10.3),* $\kappa_2^- = 2$ *wegen (10.4) und* $\kappa_3^- = 2$ *wegen (10.5). Diese Stratifikation* $\{\kappa_1\}, \{\kappa_2, \kappa_3\}$ *ist aber nicht „least commited", da die Stratifikation* Δ_{lc} *ebenfalls die Bedingungen des Autodeduktions-Prinzips erfüllt und less commited als* $\{\kappa_1\}, \{\kappa_2, \kappa_3\}$ *ist.*

Wegen $\kappa_3^- < \kappa_2^-$ *gilt in LCD* $\kappa(\omega_c) < \kappa(\omega_{\overline{c}})$*, und deshalb gilt in LCD für die gegebene Default-Menge:* $y \wedge s \wedge \neg a \mid\!\sim_{lcd} c$*. Da alle anderen Logiken aus derselben Formel* $\neg c$ *ableiten können, ist LCD nicht vergleichbar mit der Penalty-Logik, dem lexikographischen System, Brewkas bevorzugten Subtheorien oder Geffners conditional entailment.*

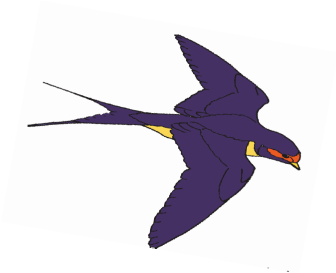
Kapitel 11

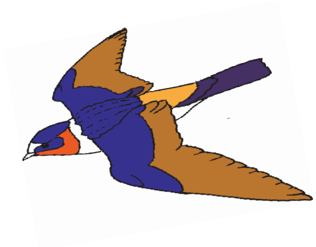
Zusammenfassung

In dieser Arbeit wurde das von Frau Prof. Kern-Isberner entwickelte Prinzip der konditionalen Indifferenz vorgestellt. Danach wurden die verschiedenen Default-Logiken Penalty-Logik, das lexikographische System, Brewkas bevorzugte Subtheorien und Geffners conditional entailment vorgestellt, und es wurde untersucht, inwieweit diese Logiken diesem Prinzip der konditionalen Indifferenz entsprechen bzw. inwieweit c-Repräsentationen eine Basis-Semantik für diese Logiken bereitstellen.

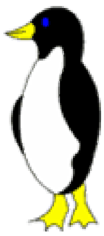
Ausserdem wurde auch der Zusammenhang zwischen den nicht-infinitesimalen Belief- und Plausibilitätsfunktionen der Dempster-Shafer-Theorie und dem Prinzip der konditionalen Indifferenz untersucht, und daraufhin wurden die infinitesimalen Belief- und Plausibilitätsfunktionen des BSS-Ansatzes ebenfalls auf konditionale Indifferenz überprüft. Zum Schluss wurde die Default-Logik LCD vorgestellt und der Zusammenhang zwischen LCD und konditionaler Indifferenz untersucht.

Das Ergebnis des Vergleichs spiegelt sich in der folgenden Tabelle wider, wobei man berücksichtigen muss, dass das Prinzip der konditionalen Indifferenz ursprünglich von einer vollständigen Ordnung auf allen möglichen Welten Ω ausgeht und deshalb auf manche Default-Logiken nur partiell angewandt werden kann:





Logik	konditional indifferent ?
Penalty-Logik	ja
Lexikographische Logik	ja
Brewkas bevorzugte Subtheorien	partiell
Geffners conditional entailment	partiell
DS[1]-Plausibilitätsfunktionen	ja
DS-Belieffunktionen	ja
infinitesimale Plausibilitätsfunktionen(BSS)	ja
infinitesimale Belieffunktionen(BSS)	ja
LCD	ja



In den Teilen, in denen keine volle Entsprechung möglich ist, liefert das Prinzip der konditionalen Indifferenz eine Basis-Semantik für alle betrachteten nicht-monotonen Logiken. Jede Inferenzrelation der betrachteten Logiken kann durch konditional indifferente Funktionen beschrieben werden.

[1] DS=Dempster-Shafer

Verzeichnis der Definitionen, Lemmas, Sätze und Beispiele

Definitionen:

Korollare:

Index

Literaturverzeichnis

[1] G. Kern-Isberner. Handling conditionals adequately in uncertain reasoning. Proc. 6th European Conference on Symbolic and Quantiative Approaches to Reasoning with Uncertainty, ECSQARU-2001

[2] S. Benferhat, A. Saffiotti and P. Smets. Belief functions and default reasoning. Artificial Intelligence, 122: 1-69, 2000

[3] G. Kern-Isberner. Representing and learning conditional information in possibility theory. In Proceedings 7th Fuzzy Days, Dortmund, Germany. Springer, LNCS-Series, 2001

[4] G. Pinkas. Propositional Non-Monotonic Reasoning and Inconsistency in Symmetric Neural Networks. Proc. IJCAI91, Sydney, Australia, Morgan Kaufmann, San Monteo, CA. 1991, S. 525-530

[5] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, H. Prade. Inconsistency management and prioritized syntax-based entailment. Proc. IJCAI 93. Chambéry, France 1993, S.640-645

[6] G. Brewka. Preferred subtheories: An extended logical framework for default reasoning. Proc.IJCAI 89, Detroit, MI, 1989, S.1043-1048

[7] H.Geffner. Default Reasoning: Causal and Conditional Theories. MIT Press, Cambridge, MA, 1992

[8] F. Dupin de Saint Cyr, J. Lang, T.Schiex. Penalty logic and its link with Dempster-Shafer theory. Proc.10th Conference on Uncertainty in Artificial Intelligence(UAI94), Seattle, WA, 1994, S.204-211

[9] G. Shafer. A mathematical theory of evidence. Princeton University Press, 1976

[10] G. Kern-Isberner, C. Beierle. Methoden der Wissensrepräsentation und -verarbeitung. Vieweg-Verlag, 2003

[11] M. Goldszmidt, J.Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. Artificial Intelligence 84 (1996), S. 57-112

[12] R. Reiter. A Logic for Default Reasoning. Artificial Intelligence 13, 1980. S. 81-132

[13] E.W.Adams. Probability and the Logic of Conditionals, in: J.Hintikka, P.Suppes(Eds.), Aspects of inductive Logic, North-Holland, Amsterdam, 1966, S. 253-316.

[14] E.W.Adams. The Logic of Conditionals. Reidel, Dortrecht, 1975.

[15] A.P. Dempster. Upper and lower probabilities induced by a multiple-valued mapping. Ann.Math.Stat.38, 1967. S.325-339.

[16] L. Sombé. Schliessen bei unsicherem Wissen in der Künstlichen Intelligenz. Vieweg Verlag, 1992.

[17] D. McDermott. A critique of pure reason. Computational Intelligence, 3, 1987, S. 151-160.

[18] J. Pearl. System Z: a natural ordering of defaults with tractable applications to default reasoning. Proceedings Third Conference on Theoretical Aspects of Reasoning about Knowledge. Pacific Grove, CA, 1990, S. 121-135.