Fuzzysysteme

Thomas Vetterlein

Institut für medizinische Experten- und wissensbasierte Systeme

Thomas.Vetterlein@meduniwien.ac.at

Sommersemester 2008

Zusammenfassung

In dem vorliegenden Bericht befassen wir uns mit Methoden der computergestützter Informationsverarbeitung, deren Besonderheit darin besteht, daß Begriffe in den Kalkül mit einbezogen werden können, die nicht in jeder Hinsicht präzise umrissene Inhalte vertreten. Insbesondere geht es darum, das Schließen mit natürlichsprachlichen Begriffen zu ergründen und dabei die unvermeidlich auftretende Vagheit mit zu berücksichtigen.

Als vage bezeichnet man eine Eigenschaft, die sich auf einen gewissen Typ von Gegenständen bezieht, ohne daß dabei in jedem Fall eindeutig auszumachen wäre, ob die Eigenschaft gilt oder nicht. Die Aussage „K. ist groß“ beispielsweise ist offenbar zutreffend, falls K. $1,80$ m groß ist, wohingegen man im Fall, daß K. eine Körpergröße von $1,70$ m aufweist, die Aussage weder als wahr noch als falsch bezeichnen möchte.

Wir vermitteln hier theoretische Grundlagen für den formalen und rechnerischen Umgang mit vagen Eigenschaften. Wir bieten eine Einführung in die Theorie der Fuzzymengen und in die t-Norm-basierte Fuzzylogik. Anwendungen finden sich im Bereich der computergestützten Entscheidungsunterstützung sowie im Bereich von Regelsystemen; einige Beispiele werden skizziert.

Inhaltsverzeichnis

1 Vagheit

Welche Problematik verbirgt sich hinter dem Begriff der Vagheit? Wie kommt es zum Interesse an Vagheit explizit mit einbeziehenden Methoden? Der vorliegende Abschnitt soll erste Einblicke vermitteln.

1.1 Natürlichsprachliche und mathematische Aussagen

Jede Aufgabe, die mit Mitteln der elektronischen Datenverarbeitung angegangen werden soll, verlangt nach einem möglichst übersichtlichen Modell des betroffenen Gegenstandes. Wir fragen uns zunächst, welche grundsätzlichen Möglichkeit vorhanden sind, einen Gegenstand zu beschreiben, und wie man auf diese Weise zu einem passenden Modell gelangen kann.

Bevor wir auf die Möglichkeiten zu sprechen kommen, die die Mathematik bietet, fragen wir uns zunächst, wie ein Gegenstand allein mit Mitteln der natürlichen Sprache erfaßt werden kann. Wir ziehen nur die unterste sprachliche Ebene in Betracht, auf der unmittelbar auf die beobachtbare Umgebung Bezug genommen wird. Eine natürlichsprachliche Aussage dieses einfachen Typs bezieht sich auf Situationen, die man erlebt oder sich vorstellen kann.

Erfaßt wird nun in einer gegebenen Situation stets nur ein bestimmter, vom Sprecher ausgewählter Aspekt, welcher einer möglichen und in aller Regel nicht der einzigen Betrachtungsweise entspricht. Die Aussage benennt weiter eine strukturelle Eigenheit des Wahrgenommenen: Sie kann einem spezifischen, abgrenzbaren Gegenstand oder einer Menge gemeinsam betrachteter Gegenstände eine Eigenschaft zuordnen; sie kann alternativ eine vom Räumlichen unabhängige Wahrnehmung wiedergeben; und in beiden Fällen kann sie einen Zustand oder eine zeitliche Entwicklung bezeichnen. Vereinfachend wird im weiteren nur die Rede davon sein, daß mittels einer natürlichsprachlichen Aussage einem Gegenstand eine Eigenschaft zugeordnet wird.

Vor Augen halten kann man sich eines der drei folgenden einfachen Beispiele. Zu denken ist etwa daran, daß ein Gegenstand der Größe nach erfaßt wird; hierfür sind die Prädikate „klein sein“ oder „groß sein“ verwendbar. Weiter kann die Temperatur eines Gegenstandes in Betracht gezogen werden, erfaßbar durch Prädikate wie „kühl sein“, „erwärmt sein“. Schließlich ist denkbar, daß die räumliche Lage beschrieben wird, etwa mittels der Prädikate „aufrecht stehen“, „liegen“.

Allerdings ist es nicht der isoliert betrachtete Gegenstand, der die Aussage möglich macht. Natürlichsprachliche Aussagen sind nur im Verbund mit vorangegangenen Erfahrungen möglich. Wird über einen Gegenstand eine Aussage gemacht, heißt dies, den Gegenstand in Bezug zu setzen zu weiteren Gegenständen, die der derselben Kategorie zugeordnet werden, sich hinsichtlich des ausgewählten Aspektes vom vorliegenden aber unterscheiden.

Von einem Gegenstand G zu sagen, er sei klein, heißt, daß G in der Klasse von Gegenständen, denen G zugerechnet wird, größenmäßig im unteren Bereich liegt: Behauptet wird, daß G „deutlich kleiner als der Durchschnitt“ sei. Ein Objekt kann als erwärmt wahrgenommen werden, wenn klar ist, wie sich das Objekt nicht erwärmt anfühlen würde. Von der Lage eines Gegenstandes zu sprechen ist nur möglich, wenn weitere mögliche Lagen klar sind; ein an einer Wand befestigter Stab etwa kann rotiert werden, womit das Prädikat „aufrecht“ eine Lage bezeichnet, die sich von den übrigen abhebt.

Auf der Ebene der natürlichen Sprache sind die entscheidenden Punkte:

- Gegenstände werden hinsichtlich eines ausgewählten Aspektes (wie Größe, Temperatur, Lage, ...) betrachtet;
- sie werden als einer Klasse von Gegenständen zugehörig angesehen, die sich hinsichtlich des gewählten Aspektes von den betrachteten unterscheiden;
- und sie werden innerhalb der Klasse hinsichtlich des Aspektes relativ zueinander charakterisiert.

All dies gilt freilich nur für eine isoliert betrachtete Aussage; ein Charakteristikum der natürlichen Sprache ist ein fliegender Wechsel hin zu neuen Aspekten und zu neu angepaßten Klassen von Gegenständen.

Was mit der natürlichen Sprache erfaßt und verarbeitet werden kann, unterscheidet sich von dem, was mit mathematischen Mitteln geleistet werden kann, nicht grundsätzlich. Die mathematische Modellierung vollzieht den Übergang von der Wahl eines Aspektes, hinsichtlich dessen man Gegenstände des entsprechenden Typs beschreiben kann, zu einer Klasse in bezug auf den gewählten Aspekt variierender Gegenstände nach und macht ihn explizit. Die sich ergebende Struktur ist dann ein fixer Bezugsgegenstand, der das Ergebnis einer definierten Konstruktion ist und dann als Bezug der Argumentation festgeschrieben wird.

Die Konstruktion geht wie folgt. Vorgegeben sei ein gewisser Aspekt, unter dem wir einen gewissen Typ von Gegenständen betrachten können. Dieser gibt uns die Möglichkeit, Gegenstände zu vergleichen; damit ergibt sich, daß eine gewisse relative Eigenschaft φ auf je n Gegenstände zutrifft, worin im allgemeinen $n = 2$ oder $n = 3$ ist.

- Man beginne mit einer endlichen Menge S^0 von Gegenständen. Auf gewisse n-Tupel $(g_1, ..., g_n)$ von Elementen aus S^0 treffe die Eigenschaft φ zu. Es liegt uns somit eine endliche Struktur vor: die Menge S^0 zusammen mit der Relation

$$R_\varphi^0 = \{(g_1, \ldots, g_n) : \text{ auf } g_1, ..., g_n \in S^0 \text{ trifft } \varphi \text{ zu } \}.$$

- Außer den in S^0 enthaltenen seien weitere Gegenstände denkbar, die sich mit denen in S^0 hinsichtlich der Eigenschaft φ vergleichen lassen. Dann erweitern wir S^0 zu einer Menge S^1 und dehnen die Relation R_φ^0 auf S^1 zur Relation R_φ^1 aus.

- Wir wiederholen den letzten Schritt, sooft es möglich ist. Wir erhalten eine Folge $S^0, S^1, S^2, \ldots$, die abbrechen kann und ansonsten ad libitum fortläuft. Wir bilden die Vereinigungsmenge

$$S = \bigcup_i S^i$$

  und definieren die Relation R_φ auf S so wie für die einzelnen Teilmengen S^i. Wir erhalten im Ergebnis die Struktur

$$(S; R_\varphi),$$

  deren Grundbereich S abzählbar unendlich sein kann.

Im Fall des ersten Beispiels galt es, Gegenstände hinsichtlich ihrer Ausdehnung zu betrachten; jeder geeignete, etwa längliche, Gegenstand wird hinsichtlich seiner Größe betrachtet. Als relative Eigenschaft φ, die für je zwei Gegenstände g_1 und g_2 zutreffen kann, ergibt sich: „g_1 ist kleiner als g_2“. Für diese Eigenschaft gilt bekanntlich Folgendes: Von je zwei verschiedenen Gegenständen ist stets einer kleiner ist als der andere; zwischen je zwei Größen liegt je eine weitere; für jeden Gegenstand gibt es einen größeren. Hinzunehmen läßt sich die Vorstellung von einem punktförmigen Objekt ohne Ausdehnung, womit es eine kleinste Größe „null“ gibt.

Damit konstruieren wir die zugehörige mathematische Struktur wie folgt.

- Wir nehmen zunächst zwei Gegenstände verschiedener Größe; ersterer sei der Gegenstand mit Nullausdehnung, bezeichnet durch 0, und letzterer sei durch g_1 bezeichnet. Wir drücken die Kleiner-Relation durch

  $$0 < g_1$$

  aus. Unsere Anfangsstruktur besteht damit aus der Menge $S^0 = \{0, g_1\}$ und der Relation $R^0_\varphi = \{(0, g_1)\}$.

- Wir nehmen nun einen weiteren Gegenstand g_2 hinzu, der seiner Größe nach zwischen 0 und g_1 liegt, sowie g_3, welcher größer ist als g_1. Wir erhalten

  $$0 < g_2 < g_1 < g_3.$$

  Unsere Struktur ist nun zu folgender angewachsen: Der Grundbereich ist $S^1 = \{0, g_1, g_2, g_3\}$, und $R^1_\varphi = \{(0, g_2), (0, g_1), (0, g_3), (g_2, g_1), (g_2, g_3), (g_1, g_3)\}$.

- Der Vorgang läßt sich wiederholen und bricht nicht ab; im nächsten Schritt etwa verlängern wir unsere Kette zu

  $$0 < g_4 < g_2 < g_5 < g_1 < g_6 < g_3 < g_7.$$

  Wir bilden die Vereinigung S all der entstehenden endlichen Strukturen und erhalten die Struktur $(S; <)$. Diese heißt *dichte lineare Ordnung mit Anfangs- und ohne Endpunkt*. Man beachte, daß $(S; <)$ isomorph ist zu $(\mathbb{Q}^+; <)$, worin $\mathbb{Q}^+$ die positiven rationalen Zahlen sind und $<$ die natürliche Kleiner-Relation auf $\mathbb{Q}^+$ ist.

Wir wollen die folgende Sprachregelung vereinbaren. Es sei φ ein natürlichsprachliches Prädikat, d.h. ein einem gewissen Typ von Gegenständen zukommende Eigenschaft. Und es sei S die mathematische Struktur – bestehend aus dem Grundbereich S und einer gewissen Relation auf S –, die man φ gemäß dem beschriebenen Vorgehen zuordnen kann. Dann werden wir auf S als die φ *inhärente Struktur* Bezug nehmen.

1.2 Natürlichsprachliche Aussage im mathematischen Modell

Offensichtlich hängt die mathematische Modellbildung eng mit natürlichsprachlichen Konzepten zusammen; spezifiert man die Strukturen, die natürlichsprachlichen Ausdrücken inhärent ist, gelangt man zu den typischerweise in der Mathematik verwendeten. Dennoch sind mathematische Aussagen mit solchen der natürlichen Sprache nicht leicht in Deckung zu bringen.

In welchem Punkt unterscheiden sich mathematische Sätze von natürlichsprachlich Ausdrückbarem? Im Anschluß an die vorstehende Darstellung kann man es so sehen: Ein Element einer mathematischen Struktur, das auf einem anschaulichen Aspekt aufbaut, steht für einen hinsichtlich dieses Aspektes maximal bestimmten Gegenstand; der Gegenstand ist abgegrenzt von allen anderen, die sich hinsichtlich des Aspektes von ihm unterscheiden lassen, auch wenn die Unterscheidung nur gedanklich möglich und nicht praktisch umsetzbar ist. Durch Bezugnahme auf die mathematische Struktur sind sämtliche Erwägungen möglich, die den betreffenden Aspekt betreffen.

Auf der elementaren Ebene einer natürlichen Sprache hingegen wird wiedergegeben, was sich spontan einer Beobachtung zuordnen läßt; es geht um Eigenschaften, die wahrnehmbar sind. Auf eine maximale Spezifikation der Gegenstände kommt es dabei nicht an; erfährt eine Situation eine winzige Änderung, bleibt die Aussage in der Regel dieselbe, insbesondere naturgemäß immer dann, wenn die Änderung unterhalb der Wahrnehmungsschwelle liegt.

Eine positive rationale Zahl r modelliert eine Größe, die für jedes $n \geq 1$ abgrenzbar ist von $r - \frac{1}{2^n}$ und $r + \frac{1}{2^n}$. Die Zahl $\frac{1}{2^n}$ entsteht durch n-maliges gleichmäßiges Aufteilen der Länge 1, was in klarer Weise gedanklich nachvollziehbar ist. Insbesondere ist r echt kleiner als $r + \frac{1}{2^n}$. Zu sagen, daß Gegenstand A kleiner sei als Gegenstand B, ist hingegen etwas anderes: A ist klar erkennbar als der im Vergleich zu B kleinere. Und dies bleibt so, wenn A oder B geringfügig in der Größe verändert wird.

Es fragt sich, wie sich Überlegungen, die keine besondere Präzision benötigen, auf Grundlage der mathematischen Strukturen anstellen lassen. Tatsache ist, daß die potentielle Überbestimmtheit eines mathematischen Wertes in der Praxis in aller Regel keine Rolle spielt; eine rationale Zahl r läßt sich identifizieren mit einem Wert von ungefähr r.

Dennoch fragt man sich, das mathematische Modell S zur Hand, nach einem Umgang mit denjenigen natürlichsprachlichen Eigenschaften, von denen man ursprünglich ausgegangen ist. φ sei eine solche Eigenschaft und beziehe sich beispielsweise auf je einen einzelnen Gegenstand; S sei die zugehörige mathematische Struktur. Es mag sein, daß sich φ mittels S ohne Schwierigkeiten durch diejenige Teilmenge von S beschreiben läßt, deren Elemente φ erfüllen. In der Regel ist φ jedoch zu unspezifisch, als daß man eine genaue solche Teilmenge angeben könnte. Eher ist nur die umgekehrte Zuordnung möglich: von gewissen Elementen von S hin zur Eigenschaft. Damit gelangen wir zum Begriff der Vagheit:

Es sei φ ein natürlichsprachliches, etwa einstelliges, Prädikat und S die diesem inhärente mathematische Struktur. Gewissen Elementen der mathematischen Struktur S lasse sich klarerweise φ zuordnen, gewissen weiteren die

Negation ψ von φ. Es sei jedoch nicht möglich, alle Elemente von S auf diese Weise zuzuordnen, d.h. es verbleiben Elemente, denen sich weder φ noch ψ zuordnen läßt. In solchem Fall heißt φ ein **vager** *Begriff.*

Zu beachten ist hierbei: Keine der drei Teilmengen ist eindeutig bestimmt. Denn dies würde eine klare Abgrenzung derjenigen Werte bedingen, auf die ein Prädikat eindeutig zutrifft. Das Prädikat war aber als natürlichsprachlich vorausgesetzt, daher auf eine Wahrnehmung Bezug nehmend, die unter kleinen Änderungen invariant ist, wohingegen ein Wert der mathematischen Struktur im allgemeinen auf einer nicht abbrechenden Folge von Wahrnehmungen beruht.

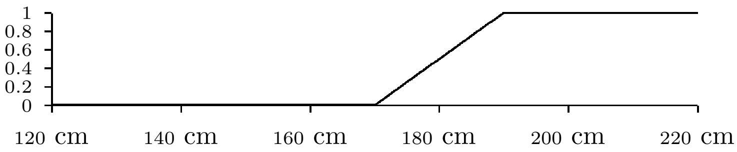
Die Grundbegriffe hinsichtlich längenmäßiger Ausdehnung sind schlicht „groß" und „klein". Geht es um die Körpergröße mitteleuropäischer Männer in Zentimetern, läßt sich beispielsweise allen Größen $g \in [0, 150]$ „klein" und allen $g \in [185, \infty)$ „groß" zuordnen. Die Werte im Intervall $(150, 185)$ bleiben ohne Zuordnung. Man beachte allerdings, daß die Grenzen 150 bzw. 185 willkürlich gesetzt sind; denn absolute Längen sind nicht wahrnehmbar und schon gar nicht mit beliebiger Präzision, weswegen sich natürlichsprachliche Begriffe auch nicht darauf beziehen können.

Dies alles kann man zum Anlaß nehmen, auf die Formalisierung natürlichsprachlicher Prädikate ein für allemal verzichten zu wollen. Allerdings gibt es doch guten Grund, sich anders zu entscheiden: Argumentationen auf Basis natürlicher Sprache sind rascher formuliert und leichter nachzuvollziehen. Die Frage ergibt sich, ob sich ein Prädikat wie „groß" im Rahmen des mathematischen Modells nicht doch sinnvoll wiedergeben läßt.

Das übliche Vorgehen ist das folgende: Ist φ ein Prädikat der natürlichen Sprache und S die diesem inhärente Struktur, wird jedem Element $e \in S$, auf das φ klar nicht zutrifft, eine 0, jedem $e \in S$, auf das φ klar zutrifft, eine 1 zugeordnet. Die verbleibenden Fälle sind die Grenzfälle von φ in bezug auf S. Jedem $e \in S$, für das das Prädikat nicht spezifiziert ist, wird sodann eine Tendenz zugeordnet, ausgedrückt durch einen Wert im reellen Einheitsintervall $[0, 1]$. Der Grad $0,5$ drückt Unentschiedenheit aus, und ein höherer Wert drückt einen tendenziell höheren Kompatibilitätsgrad aus. Eine Funktion von S nach $[0, 1]$ mit dieser Interpretation heißt Fuzzymenge über S.

Das Prädikat „groß" kann durch folgende Fuzzymenge modelliert werden.

Eine Fuzzymenge zur Modellierung des Prädikates „groß“



Zu beachten ist, daß die konkrete Gestalt dieser Funktion auf Willkür beruht. Aber eines ist erreicht: Es ist eine Verbindung zwischen einer mit natürlicher Sprache ausgedrückten Eigenschaft und der zugehörigen mathematischen Struktur geschaffen worden.

Während die Wahl der Fuzzymenge meist irrelevant ist, solange es um geringfügige Details des Kurvenverlaufs geht, ist zu beachten: Da es um natürlichsprachliche Begriffe geht, hängt die passende Fuzzymenge grundsätzlich in starkem Maße vom Zusammenhang ab, der nie außer acht gelassen werden kann. Allein dieser gibt vor, welches die Bezugsobjekte sind, und damit, welche Gegenstände der betreffenden Klasse die typischen sind und welche die typischen für die Negation sind.

1.3 Die philosophische Debatte über Vagheit

Wir bleiben noch eine Weile bei den grundlegenden Erörterungen. Vagheit ist nicht etwa ein einzelner Gegenstand am Rande der philosophischen Debatte; es handelt sich vielmehr um ein eigenständiges Gebiet mit reichhaltiger Literatur, innerhalb deren sich gar noch mehrere Schulen ausmachen lassen. Wir machen uns von den schwer verständlichen Theorien keine zu eigen, verweisen aber auf das tausendfach besprochene und unter Philosophen als letztlich ungelöst geltende sogenannte Sorites-Paradox:

Wenn ich von einem Haufen Sandkörner ein einziges Sandkorn hinfortnehme, habe ich weiterhin einen Haufen Sandkörner vor mir. Das heißt: Bilden n Körner einen Haufen, dann bilden auch $n - 1$ Körner einen Haufen. Außerdem bilden 10000 *Körner einen Haufen. Es folgt per Induktion:* 1 *Korn allein bildet ebenfalls einen Haufen.*

Wo liegt das Problem dieser widersinnigen Argumentation? Es liegt darin, daß die Eigenschaft, einen Haufen zu bilden, als hinsichtlich einer Menge von n Sandkörnern für jedes n anwendbar angenommen wird.

Das betrachtete natürlichsprachliche Prädikat H lautet „einen Haufen darstellen“ und dessen Negation „keinen Haufen darstellen“. Die betrachtete Situation ist beschrieben durch eine natürliche Zahl n, die Zahl der Sandkörner. Es gilt wie im letzten Abschnitt erläutert: Gewissen $n \in \mathbb{N}$ läßt sich H zuordnen, gewissen die Negation von H, aber nicht für alle gilt eine der beiden Möglichkeiten. Weiter ist eine eindeutige scharfe Grenze zu finden unmöglich, ab wann H nicht mehr eindeutig gilt, unmöglich; jedes natürlichsprachliche Prädikat hat einen gewissen Spielraum, d.h. ist invariant gegenüber kleinen Änderungen.

Was das Paradox so verwirrend macht, ist wohl die Vermengung natürlichsprachlicher mit formalsprachlicher Argumentation. Beide sind letztlich eben verschieden. Für die natürlichsprachliche Argumentation gilt die Orientierung an Prototypen. Man überlege, was der Satz

Bilden n Körner einen Haufen, dann bilden auch $n-1$ Körner einen Haufen.

eigentlich aussagt. *n Körner bilden einen Haufen*, heißt, daß n so groß ist, daß ein aus n Körnern bestehender Haufen eindeutigerweise einen Haufen bildet; man sieht mithin einen typischen Haufen vor dem geistigen Auge; folglich ist n invariant gegenüber Erniedrigung um 1. Auf diesen prototypischen Fall verweist der „Induktionsschritt“ – der aber im Verlaufe der Induktion für sämtliche n Verwendung findet.

Der „Induktionsschluß“ ist im Beispiel freilich einer formalsprachlichen Version nachempfunden, etwa der entsprechenden Regel der Peano-Arithmetik, der Theorie der natürlichen Zahlen. Um einen solchen anwenden zu können, muß ich von einer definierten Struktur ausgehen; in bezug auf diese müssen meine Voraussetzungen und meine Schlußregeln gelten und gilt dann erst sicher auch das Erschlossene. Jedoch gibt es keine Struktur, für die „Induktionsvoraussetzung“ und „-schluß“ stimmen. Man geht ja davon aus, daß für jedes n die Aussage $H(n)$ besagt, ob ein Haufen vorliegt oder nicht, und insbesondere davon, daß $H(1)$ falsch und $H(10000)$ richtig ist. Damit wird der Schluß schlicht inkonsistent.

Nachbesserung und Wiedergewinnung von Konsistenz ist mit der im vorigen Abschnitt erwähnten Methode möglich. Schließlich ist ja die implizite Annahme, daß für jedes $H(n)$ einen eindeutigen Wahrheitswert hat, unsinnig; etwas pragmatisch, aber immerhin weniger unsinnig läßt sich jedem n ein Wert $H(n) \in [0, 1]$ zuordnen, der besagt, zu welchem Grad eine Ansammlung von n Körnern ein Haufen sei, mit steigendem n ein größerer. Weiter ordnen wir auch dem „Induktionsschluß" einen Wahrheitswert zu, und zwar $\frac{n-1}{n}$. Und wir definieren, daß die Konjunktion zweier Aussagen, denen der Wert $1 - \varepsilon_1$ bzw. $1 - \varepsilon_2$ zukommt, den Wert $1 - \varepsilon_1 - \varepsilon_2$ erhält. Dann ist die Aussage

$$\begin{aligned} H(10000) \odot (H(10000) \to H(9999)) \odot \\ (H(9999) \to H(9998)) \odot \ldots \odot (H(2) \to H(1)) \end{aligned}$$

zwar stärker als $H(1)$, hat jedoch einen Wahrheitswert von lediglich $\frac{1}{10000}$; nähere Erklärungen zu diesem Vorgehen folgen im Abschnitt über Logik.

Sicherlich sind bei diesem Vorgehen alle Annahmen fraglich und ließen sich auf vielerlei Weisen variieren. Den Eigenheiten der verwendeten natürlichsprachlichen Begriffe werden sie aber in hinreichendem Maße gerecht.

1.4 Vagheit und Unsicherheit

Die Schwäche des Konzeptes der Fuzzymengen besteht darin, daß sich ein einzelner Wahrheitswert für sich genommen schlecht interpretieren läßt. Es hat in der Vergangenheit nicht an Versuchen gefehlt, eine Interpretation „nachzuliefern". Neben ernsthaften Ansätzen ist der Versuch schon fast notorisch, Fuzzywahrheitswerte als Wahrscheinlichkeiten zu interpretieren.

Wohl ist es möglich, losgelöst vom bislang beschriebenen Ansatz Wahrheitswerte in der Fuzzylogik wie auch immer zu begreifen und Fuzzylogiken beispielsweise als Logiken über Wahrscheinlichkeiten zu interpretieren. Auch in dieser Vorlesung wird von einer speziellen Situation die Rede sein, in der diese Interpretation die maßgebliche ist.

Im allgemeinen jedoch verbindet sich mit den Wahrheitswerten Vagheit, und Vagheit ist scharf von Unsicherheit zu trennen. Vage sind Eigenschaften, bei denen Grenzfälle auftreten. Bin ich mir hingegen über eine Aussage φ zum Grad $0,4$ unsicher, heißt dies, daß φ zutrifft oder nicht, ich mir hierüber aber nicht sicher bin und nur sagen kann, daß sich, wenn ich der Sache auf den Grund gehe, mit 40%-iger Wahrscheinlichkeit φ als zutreffend bestätigen läßt.

Weiter schließen sich Vagheit und Unsicherheit wechselseitig natürlich nicht aus; beides kann kombiniert auftreten. Ein vages Prädikat, d.h. eine Aussage mit nicht klar abgegrenztem Zutreffensbereich, kann im Grenzfall nicht klar für wahr oder falsch befunden werden, und es kann außerdem noch Unklarheit über den zu wählenden Wahrheitsgrad herrschen. Dies wird für uns jedoch nicht das Thema sein; hier geht es in allererster Linie um Vagheit.

2 Algebraische Theorie der Fuzzymengen

2.1 Fuzzymengen - Definition und Beispiele

In natürlicher Sprache formulierte Aussage der Art, daß ein gewisser Gegenstand eine gewisse Eigenschaft erfülle, spiegeln einen Sinneseindruck wider; an Information fließt ein, was sich mit einer einzelnen Beobachtung feststellen läßt. Das Standardbeispiel ist die Aussage „K. ist groß“, wenn wir den $2,20$-m-Menschen „K.“ vor uns sehen. Bedenkt man weiter, was sich über solcherlei Eindrücke hinausgehend feststellen läßt, ergibt sich meist, daß das Ausgesagte recht grob ist und sich präzisieren ließe. Im Standardbeispiel geht es um die Größe; zieht man sämtliche möglichen Größen in Betracht, die mit präziseren Beobachtungen, sprich mit Messungen, festzustellen sind, werden wir zu einer Struktur geführt, mit der sich sämtliche Größen ausdrücken und gegen die übrigen abgrenzen lassen. Unter Verwendung der positiven rellen Zahlen etwa lautet die Aussage „K. ist $2,20$ m groß“ und ist dann nicht mehr weiter präzisierbar.

Der Begriff der reellen Zahl und der Umgang mit reellen Zahlen ist uns wohlbekannt und bietet keine Schwierigkeiten. Der Fall ist aber häufig, daß Informationen nicht mit hinreichender Präzision vorliegen, trotzdem aber berücksichtigt werden sollen. Die Frage ist, wie sich grobe, insbesondere natürlichsprachliche Aussagen formal erfassen lassen. Unser Ziel ist es, das Rechnen mit groben Werten und das Schließen mit groben Aussagen zu ergründen.

Der Ansatz, der hier besprochen wird, sieht vor, an den gewohnten mathematischen Modellen festzuhalten, den jeweiligen Grundbereich der Modelle jedoch geeignet zu verallgemeinern. Die Grundfrage der Fuzzymengentheorie lautet: Wie kann ich grobe sprachliche Stellungnahmen hinsichtlich der feineren Struktur ausdrücken? Die allgemein verwendete, pragmatische Lösung ist oben angegeben worden: Wir verwenden Fuzzymengen über dem Grundbereich der mathematischen Struktur.

Definition 2.1 Das *reelle Einheitsintervall* ist die Menge

$$[0,1] = \{t \in \mathbb{R}\colon\ 0 \leq t \leq 1\}.$$

Eine Funktion $u\colon\ M \to [0,1]$ von einer beliebigen nichtleeren Menge M in das reelle Einheitsinterval heißt *Fuzzymenge* über M. Wir nennen M dann den *Bezugsbereich* von u.

Es bezeichne $\mathcal{F}(M)$ die Menge aller Fuzzymengen über M.

Über den Bezugsbereich M einer Fuzzymenge wird hier nichts vorausgesetzt; es darf sich um eine beliebige Menge handeln. Im typischen Fall ist M jedoch der Grundbereich einer mathematischen Struktur, mit der Gegenstände hinsichtlich einer Eigenschaft mit maximaler Präzision beschreibbar sind; M ist allermeist die Menge $\mathbb{R}$ der reellen Zahlen oder deren kartesisches Produkt, d.h. $\mathbb{R}^n$ für ein natürliches n. Fuzzymengen werden stets mit Berücksichtigung der dem Bezugsbereich unterliegenden Struktur gewählt; das obige Beispiel einer Fuzzymenge über $\mathbb{R}$ ist bereits ein Beispiel hierfür. Die Struktur von M ziehen wir in die Diskussion aber erst zu späterem Zeitpunkt ein; im Moment sei M beliebig.

Fuzzymengen über einem Bezugsbereich M lassen sich als eine Verallgemeinerung der einzelnen Elemente von M begreifen; dementsprechend sprechen wir im Fall $M = \mathbb{R}$ auch derweil von Fuzzymengen über M als von Fuzzyzahlen. Insbesondere ist jedes Element von M mit einer speziellen Fuzzymenge identifizierbar.

Definition 2.2 Es sei $m \in M$. Dann sei

$$\sigma_m \colon M \to [0,1], \quad x \mapsto \begin{cases} 1 & \text{falls } x = m, \\ 0 & \text{sonst} \end{cases}$$

der *Singleton* des Punktes m.

Es ist klar, daß die Zuordnung $M \to \mathcal{F}(M), \;\; m \mapsto \sigma_m$ eine Bijektion zwischen M und den Singletons der Elemente von M darstellt. Im Blick auf diese Zuordnung werden wir im weiteren M als eine Teilmenge von $\mathcal{F}(M)$ ansehen.

Genausogut können Fuzzymengen über M aber auch als verallgemeinerte Teilmengen von M angesehen werden; welche Sichtweise die passende ist, entscheidet die Anwendung.

Definition 2.3 Es sei $A \subseteq M$. Dann sei

$$\chi_A \colon M \to [0,1], \quad x \mapsto \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{sonst} \end{cases}$$

die *charakteristische Funktion* der Teilmenge A. Eine Fuzzymenge von der Form χ_A für ein $A \subseteq M$ heiße *scharf* oder auch *crisp*.

Wiederum ist $\mathcal{P}(M) \to \mathcal{F}(M), \quad A \mapsto \chi_A$ eine Bijektion zwischen der Potenzmenge von M und den crispen Fuzzymengen, und $\mathcal{P}(M)$ kann als Teilmenge von $\mathcal{F}(M)$ angesehen werden.

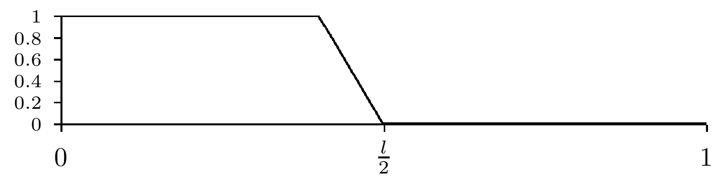
Wir erläutern nun die wichtigen Punkte zur Interpretation von Fuzzymengen anhand konkreter Beispiele.

Beispiel 1. Wir blicken seitlich auf eine Strecke mit den Randpunkten A und B, auf der sich eine Kugel befindet. Die Aussage φ laute: „Die Kugel befindet sich, von uns aus gesehen, im linken Bereich der Strecke $\overline{AB}$."

Das mathematische Modell ist das folgende. Ist l die Länge der Strecke, wird die Position der Kugel durch den Abstand des Kugelmittelpunktes zum Punkt A beschrieben, einen Wert $r \in [0, l]$ also.

Ist nun beispielsweise $r = \frac{1}{4}l$, trifft φ klarerweise zu; ist $r = \frac{3}{4}l$, ist φ klarerweise falsch. Ist aber r nur sehr geringfügig kleiner als $\frac{1}{2}l$, ist es nicht recht möglich zu sagen, φ treffe zu oder nicht. φ ist demgemäß eine typische Vagheit involvierende Aussage: Die Menge aller $r \in [0, l]$, für die φ stimmt, ist nicht scharf eingrenzbar.

Wir modellieren φ als Fuzzymenge über $[0, l]$: als eine Funktion u vom Intervall $[0, l]$ in die reellen Zahlen zwischen 0 und 1:

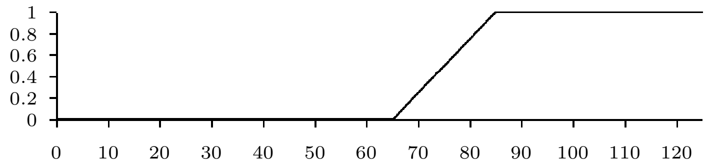


Die Interpretation geht wie folgt: Durch $r \in [0, l]$ mit $u(r) = 1$ beschriebene Positionen erfüllen φ klarerweise; durch $r \in [0, l]$ mit $u(r) = 0$ beschriebene Positionen erfüllen φ überhaupt nicht. In allen weiteren Fällen ist φ eigentlich nicht spezifiziert, drücken wir durch $u(r)$ aber eine Tendenz aus. Der Wert $u(r)$ quantifiziert, inwieweit die durch r gegebene Lage mit der Eigenschaft φ noch vereinbar ist.

Man beachte auch hierbei, daß die konkrete Ausgestaltung der Funktion u stark auf Willkür beruht; das Vorgehen ist pragmatisch. Auf der anderen Seite leistet u das, was es soll: in plausibler Weise die Eigenschaft „auf der linken Seite befindlich“ auszudrücken. Lediglich sind andere Möglichkeiten, u zu wählen, nicht ausgeschlossen.

Beispiel 2. K. sei eine Person; die Aussage ψ laute: „K. ist alt.“

Es sei A eine zeitliche Dauer in Jahren. Dann können wir, wie in Beispiel 1, für A eine Fuzzymenge ansetzen:

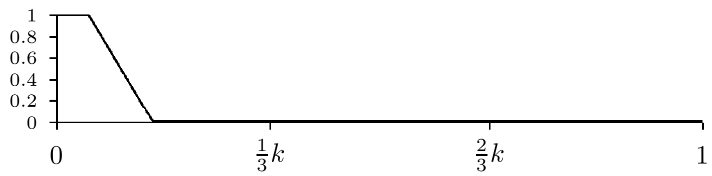


Hierbei muß jedoch klar sein, daß wir es mit einer Schwierigkeit zu tun haben, die in Beispiel 1 nicht aufgetreten ist: Natürlichsprachliche Aussagen sind vergleichend. In Beispiel 1 hatten wir alle zu betrachtenden Situationen explizit vorgegeben; einen äußeren Zusammenhang vorzugegeben war damit nicht nötig. Im allgemeinen muß vor der Wahl einer Fuzzymenge aber klar sein, auf welche Menge von Gegenständen sich eine behauptete Eigenschaft bezieht. Im vorliegenden Beispiel muß ein Personenkreis eingegrenzt werden. Andernfalls kann nur eine Standardwahl getroffen werden, wie im Bild geschehen – wir sind von der derzeitige Bevölkerung Mitteleuropas ausgegangen. Natürlich bleibt auch nach dieser Wahl die konkrete Ausgestalung von u immer noch willkürlich, im selben Sinn wie in Beispiel 1.

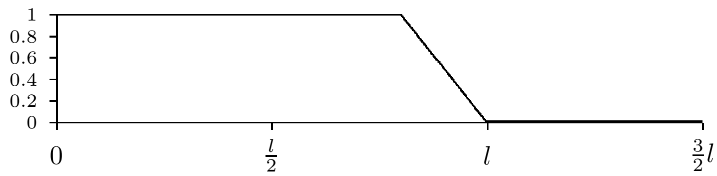
Beispiel 3. Es seien A und B zwei beliebige (materielle) Gegenstände. Die Aussage χ laute: „A und B liegen dicht beieiander.“

Wir sehen, daß wir in diesem Fall ohne weitere Abklärungen keine passende Fuzzymenge angeben können. Wiederum müssen Vergleichssituationen klar sein. Im vorliegenden Fall sollte klar sein, um welche Gegenstände es sich handelt und innerhalb welches Raumbereiches befindlich man sie sich vorstellt; es muß klar werden, wie man zu dem Urteil gelangt, daß sich A und B nahe sind. Erst dann läßt sich sagen, in welchen Fällen man eindeutig zu diesem Urteil gelangt und in welchen Fällen nie, und eine passende Fuzzymenge wählen.

Ein Minimum an Information wäre die folgende: A und B werden punktförmig angesehen und halten sich auf einer quadratischen Fläche der Kantenlänge k auf. Dann könnte „dicht beieinander“ einen Abstand von weniger als etwa $\frac{1}{10}k$ bedeuten, was zur folgenden Fuzzymenge Anlaß geben könnte:



Möglich ist weiter, daß man guten Grund hat, A und B als nahe zu bezeichnen; von diesem sollte ausgegangen werden, um festzustellen, welche Abstände als nah gelten und welche nicht. Geht es etwa darum, daß ein weiterer Gegenstand C mit der Ausdehnung l zwischen A und B hindurchpassen soll, könnte χ den Eindruck ausdrücken, daß der Abstand zwischen A und B eindeutig zu klein ist; man könnte dann vielleicht auf folgende Fuzzymenge kommen:



Beispiel 4. Es sei σ die Aussage: „P. mag Dvořák.“

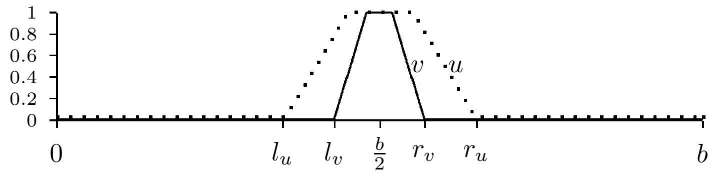
σ ordnet einer Person P. die Eigenschaft zu, ein Liebhaber Dvořáks zu sein. Diese Aussage ist etwas anderer Natur als die bisherigen; es sind keine raumzeitlichen Strukturen betroffen, nichts, was meßbar wäre, sondern vielmehr jemandes subjektive Empfindung. Dementsprechend gibt es auch keine Möglichkeit, die angesprochene Eigenheit P.s näher zu erforschen und zu präzisieren.

Nichtsdestoweniger läßt sich bezüglich einer Menge P von Personen, die über Dvořák eine Meinung haben, die Eigenschaft „Liebhaber Dvořáks“ erkunden. Es wird sich eine Teilmenge brennender Liebhaber und eine Teilmenge sich von solcher Musik belästigt Fühlender ergeben; und eine weitere Teilmenge von Personen, die allenfalls eine Tendenz äußern. Dieser Tatbestand ließe sich durch eine Fuzzymenge $u : P \to [0,1]$ modellieren. Der Unterschied zu Beispiel 1 liegt aber darin, daß die unterliegende Menge P strukturlos ist und insbesondere nicht die verschiedenen Ausprägungen der betrachteten Eigenschaft gesamthaft zusammenstellt.

Wir erachten dennoch auch σ als eine vage Aussage, die zur Diskussion paßt. Die σ zuzuordnende Fuzzymenge soll jedoch nur einelementig sein, d.h. nicht mehr als die Person P. enthalten; dieser ist der Wahrheitswert zuzuordnen, der dem Grad entspricht, zu dem P. σ zustimmt.

2.2 Die logisch-algebraische Struktur von Fuzzymengen

Wir beginnen ein weiteres Mal mit einem Beispiel. Die folgende Fuzzymenge modelliere die Aussage φ „Gegenstand G befindet sich im mittleren Bereich zwischen Punkt A und B“ sowie ψ „Gegenstand G befindet sich nahe der Mitte zwischen Punkt A und B“.



Der Bezugsbereich – die Punkte zwischen A und B, beschrieben durch den Abstand von A – durchläuft die einzelnen Möglichkeiten, die in Betracht gezogen werden – die möglichen Aufenthaltsorte von G. Die Möglichkeiten schließen sich wechselseitig aus, und es gilt eine von ihnen.

Wir geben weiter φ formal dadurch wieder, daß für jede der sich wechselseitig ausschließenden Möglichkeiten angegeben wird, wieweit diese mit φ vereinbar ist. Die φ modellierende Fuzzymenge u vertritt zunächst die Behauptung, daß sich G an einem der Orte $x \in [l_u, r_u]$ befindet; die übrigen Punkte bildet u auf 0 ab. Weiter wird durch die den in $[l_u, r_u]$ liegenden Punkten zugeordneten Wahrheitsgrade ausgedrückt, daß gewisse Möglichkeiten des Aufenthaltes von G der betrachteten Eigenschaft „im mittleren Bereich liegend“ perfekt entsprechen, andere hingegen nur begrenzt gut.

Die für uns relevante Sichtweise ist nun die folgende, in gewissem Sinne umgekehrte. Der Bezugsbereich durchläuft verschiedene alternative Möglichkeiten, und φ schließt jede der Möglichkeiten in mehr oder weniger entschiedenem Maße aus. Möglichkeiten mit Gewichtung 1 sind gar nicht ausgeschlossen; solche mit Gewichtung 0 sind ganz ausgeschlossen; die übrigen sind zu umso höherem Grade ausgeschlossen, je kleiner die zugeordnete Gewichtung ist.

Dies bedeutet, daß eine Erniedrigung der Wahrheitswerte die Aussage verschärft. Man betrachte die zweite Fuzzymenge v, die ψ modelliert und punktweise kleiner ist als φ. v gibt eine Aussage wieder, der gemäß weitere Aufenthaltsorte auszuschließen sind und für gewisse festzulegen ist, daß sie mit höherer Entschiedenheit auszuschließen sind, jeweils im Vergleich zu u. Mit anderen Worten vertritt v die stärkere Aussage als u, in Übereinstimmung damit, daß ψ offensichtlich inhaltlich stärker ist als φ.

Es folgt aus dieser Diskussion: Punktweise kleinere Fuzzymengen modellieren die stärkeren Aussagen. Weiter repräsentiert die Fuzzymenge, die alle Elemente des Bezugsbereiches auf 0 abbildet, die stärkste Aussage. Es handelt sich um den Widerspruch, da ja sämtliche mögliche Alternativen ausgeschlossen werden, und folglich um die Aussage „falsch“. Die schwächste Fuzzymenge ist diejenige, die alle Elemente auf 1 abbildet. Diese repräsentiert das Zutreffen als solches; sie schließt überhaupt nichts aus, schränkt auch keine der Möglichkeiten ein und modelliert dementsprechend die Aussage „wahr“.

Der algebraische Begriff, um die Struktur einer Menge von Aussagen mittels deren Stärke zu beschreiben, ist die des Posets.

Definition 2.4 Es sei $\leq$ eine binäre Relation auf einer Menge L, so daß für $a, b, c \in L$ Folgendes gilt:

(PO1) Es gilt $a \leq a$.

(PO2) Aus $a \leq b$ und $b \leq c$ folgt $a \leq c$.

(PO3) Falls $a \leq b$ und $b \leq a$, dann ist $a = b$.

Dann heiße $(L; \leq)$ eine *partiell geordnete Menge* oder kurz *Poset*.

Ein Element c eines Posets heiße das *Infimum* zweier Elemente a und b, falls $c \leq a$ und $c \leq b$ gilt und aus $x \leq a$ und $x \leq b$ stets $x \leq c$ folgt; wir schreiben in diesem Fall $c = a \wedge b$.

Analog heiße d das *Supremum* von a und b, falls $a \leq d$ und $b \leq d$ gilt und aus $a \leq x$ und $b \leq x$ stets $d \leq x$ folgt; wir schreiben in diesem Fall $c = a \vee b$.

Besitzen je zwei Elemente eines Posets ein Infimum und ein Supremum, heiße die Algebra $(L; \wedge, \vee)$ ein *Verband*.

Ferner heiße ein Element $0 \in L$ mit $0 \leq a$ für alle $a \in L$ *Nullelement* von L; ein Element $1 \in L$ mit $a \leq 1$ für alle $a \in L$ heiße *Einselement* von L. Enthält ein Verband ein Nullelement 0 und ein Einselement 1, heiße die Algebra $(L; \wedge, \vee, 0, 1)$ ein 0, 1-Verband.

Man beachte, daß infolge (PO3) je zwei Elemente a, b eines Posets höchstens ein Infimum und ein Supremum haben können; dies erst rechtfertigt die Einführung der Operationen $\wedge$ und $\vee$. In einem Verband ist weiter die unterliegende Ordnung aus der Operation $\wedge$ oder auch aus $\vee$ zurückgewinnbar: $a \leq b$ trifft zu genau dann, wenn $a \wedge b = a$, und ebenso genau dann, wenn $a \vee b = b$.

Wir fixieren von nun an, auch für die folgenden Teilabschnitte, eine Menge M als Bezugsbereich der betrachteten Fuzzymengen. Zur Veranschaulichung sollte man sich stets den Fall $M = \mathbb{R}$ vor Augen halten.

Definition 2.5 Für zwei Fuzzymengen $u, v \colon M \to [0,1]$ gelte $u \leq v$, falls

$$u(x) \leq v(x) \quad \text{für alle } x \in M$$

gilt. Weiter setzen wir

$$\begin{aligned} \bar{0} \colon M &\to [0,1], \quad x \mapsto 0, \\ \bar{1} \colon M &\to [0,1], \quad x \mapsto 1. \end{aligned}$$

Es ist klar, daß die Relation $\leq$ die Menge $\mathcal{F}(M)$ zur einem Poset macht. Kleinere Elemente von $\mathcal{F}(M)$ modellieren die stärkeren Aussagen; das kleinste Element $\bar{0}$ repräsentiert den Widerspruch; und das größte Element $\bar{1}$ steht für das Wahre.

Es geht im weiteren darum, zu untersuchen, welche Möglichkeiten bestehen, mehrere Fuzzymengen zu einer zusammenzufassen. Naheliegend ist als erstes die Feststellung, daß der Poset $\mathcal{F}(M)$ ein Verband ist.

Lemma 2.6 $\mathcal{F}(M)$ ist ein $0,1$-Verband. Für $u, v \colon M \to [0,1]$ ist

$$\begin{aligned} (u \wedge v)(x) &= u(x) \wedge v(x), \\ (u \vee v)(x) &= u(x) \vee v(x), \quad x \in M; \end{aligned} \tag{1}$$

ferner ist $\bar{0}$ das 0- und $\bar{1}$ das 1-Element.

Für Fuzzymengen u und v modelliert $u \wedge v$ die schwächste Aussage, die sowohl u als auch v impliziert. Analog modelliert $u \vee v$ die stärkste Aussage, die sowohl von u als auch von v impliziert wird.

Das Infimum eignet sich durchaus für die Rolle einer Konjunktion, ist aber nicht die einzige Möglichkeit, zwei durch Fuzzymengen modellierte Aussagen im Sinne einer Konjunktion zu einer einzelnen zusammenfassen. Und es ergibt sich weiter leider die Unmöglichkeit, für die Konjunktion eine Wahl zu treffen, die sich gegenüber anderen infolge übezeugender Eigenschaften auszeichnet. Wir tragen einige plausible Mindestanforderungen zusammen.

Es gilt zunächst die folgende Überlegung. Fuzzymengen ordnen ihrem Bezugsbereich Wahrheitsgrade zu, und der Bezugsbereich ist als Menge sich gegenseitig ausschließender Alternativen begreifbar. Wenn zwei Fuzzymengen über demselben Bezugsbereich logisch kombiniert werden sollen, liegt es daher nahe, jede der Alternativen getrennt zu behandeln und alle Alternativen in der gleichen Weise. Dies heißt, daß eine Funktion $\odot\colon\ [0,1]^2 \to [0,1]$ bestimmt werden sollte, gemäß der die Wahrheitswerte zweier Fuzzymengen Punkt für Punkt zu je einem Wahrheitswert zusammenzufassen sind. Ausgehend von einer Funktion $\odot\colon\ [0,1]^2 \to [0,1]$ kombinieren wir also zwei Fuzzymengen $u, v\colon\ M \to [0,1]$ wie folgt zu einer neuen, die wir durch $u \odot v$ bezeichnen:

$$(u \odot v)(x) = u(x) \odot v(x), \quad x \in M. \tag{2}$$

(Wenn wir wie hier Fuzzymengen punktweise verknüpfen, verwenden für die Verknüpfung von Fuzzymengen immer dasselbe Symbol wie für die Verknüpfung der Wahrheitswerte.) Wir sagen dann, daß die Operation $\odot$ auf der Menge der Fuzzymengen von der auf der Menge der Wahrheitswerte erklärten Operation $\odot$ *induziert* ist.

Unsere Aufgabe ist nun reduziert auf die Frage nach einer geeigneten auf den Wahrheitswerten operierenden binären Operation $\odot$. Da es eine kanonische Wahl nicht gibt, beschränken wir uns darauf, gewisse Eigenschaften einzufordern, und gelangen zum folgenden, zentralen Begriff der Fuzzymengentheorie.

Definition 2.7 Eine Funktion $\odot\colon\ [0,1]^2 \to [0,1]$ heißt *t-Norm*, falls Folgendes gilt:

(T1) $\odot$ ist assoziativ: $a \odot (b \odot c) = (a \odot b) \odot c$ für alle $a, b, c \in [0,1]$.

(T2) $\odot$ ist kommutativ: $a \odot b = b \odot a$ für alle $a, b \in [0,1]$.

(T3) 1 ist bezüglich $\odot$ neutral: $a \odot 1 = a$ für alle $a \in [0,1]$.

(T4) $\odot$ ist monoton: Aus $a \leq b$ folgt $a \odot c \leq b \odot c$ für $a, b, c \in [0,1]$.

Die drei bekanntesten t-Normen sind die folgenden. Die *łukasiewiczsche* t-Norm $\odot$ ist definiert gemäß

$$\odot_L \colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto \max\{a+b-1, 0\}. \tag{3}$$

Die *Produkt-t-Norm* ist das gewöhnliche Produkt reeller Zahlen:

$$\odot_P \colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto a \cdot b. \tag{4}$$

Und die *gödelsche* t-Norm ist schlicht das Infimum:

$$\odot_G \colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto a \wedge b. \tag{5}$$

Die Theorie der t-Normen ist umfangreich und Gegenstand der aktuellen Forschung. Wir geben einen Einblick im nachfolgenden Abschnitt.

Wir beschreiben weiter die Algebra, zu der die Menge aller Fuzzymengen durch Hinzunahme einer durch eine t-Norm induzierte Konjunktion wird.

Definition 2.8 Eine Algebra $(L; \odot, 1)$ heiße *kommutatives Monoid*, falls $\odot$ assoziativ und kommutativ ist und 1 bezüglich $\odot$ ein neutrales Element.

Eine Algebra $(L; \wedge, \vee, \odot, 0, 1)$ heiße *ℓ-Monoid*, falls Folgendes gilt:

(M1) $(L; \wedge, \vee, 0, 1)$ ist ein $0,1$-Verband.

(M2) $(L; \odot, 1)$ ist ein kommutatives Monoid.

(M3) $\odot$ ist monoton: Aus $a \leq b$ folgt $a \odot c \leq b \odot c$ für $a, b, c \in L$.

Aus der Definition ist sofort ersichtlich, daß eine Funktion $\odot \colon\ [0,1]^2 \to [0,1]$ genau dann eine t-Norm ist, wenn die Algebra $([0,1]; \wedge, \vee, \odot, 0, 1)$, worin sich $\wedge$ und $\vee$ auf die natürliche Ordnung der reellen Zahlen beziehen, ein ℓ-Monoid ist.

Für $\mathcal{F}(M)$ ergibt sich:

Lemma 2.9 *Es sei $\odot$ eine t-Norm. Auf der Menge der Fuzzymengen $\mathcal{F}(M)$ seien $\wedge$ und $\vee$ sowie $\bar{0}$ und $\bar{1}$ gemäß Definition* 2.5 *und die Operation $\odot$ gemäß* (2) *erklärt. Dann ist $(\mathcal{F}(M); \wedge, \vee, \odot, \bar{0}, \bar{1})$ ein ℓ-Monoid.*

Wir kommen zur nächsten grundlegenden Operation auf der Menge der Fuzzymengen: dem Residuum. Es handelt sich um eine Operation $\rightarrow\colon \mathcal{F}^2 \rightarrow \mathcal{F}$, die in erster Linie dazu dient, auszudrücken, daß eine Fuzzymenge stärker ist als eine andere, und in diesem Fall die Rolle einer Implikation spielt. Es seien $u, v \in \mathcal{F}(M)$; dann soll gelten:

$$u \leq v \quad \text{genau dann, wenn} \quad u \rightarrow v = \bar{1}, \tag{6}$$

d.h. genau dann, wenn $a \rightarrow b$ die Fuzzymenge „wahr" ist. Damit ist $u \rightarrow v$ auf $\bar{1}$ festgelegt, wenn $u \leq v$ gilt. $u \rightarrow v$ ist allerdings für alle $u, v \in \mathcal{F}(M)$ zu definieren; es kommt der folgende Ansatz zum Tragen:

Es seien u und v zwei Fuzzymengen. Dann repräsentiere $u \rightarrow v$ die schwächste Aussage, die zusammen mit u stärker ist als v.

Hierin bezieht sich „zusammen mit" auf eine vorher festgelegte Konjunktion; und daß es eine Fuzzymenge die stärkere Aussage modelliert, bedeutet, daß sie punktweise kleiner ist. Damit läßt sich der Ansatz durch folgende Definition präzisieren.

Definition 2.10 Es sei $\mathcal{F}(M)$ mit einer durch eine t-Norm induzierten Konjunktion $\odot$ ausgestattet und $a, b \in \mathcal{F}(M)$. Dann sei

$$u \rightarrow v = \max \{w \in \mathcal{F}(M) \colon\ u \odot w \leq v\},$$

vorausgesetzt, daß das Maximum existiert.

Es ergibt sich die Richtigkeit von (6) unmittelbar; ist $u \leq v$, so gilt $u \odot \bar{1} = u \leq v$, weshalb $u \rightarrow v$ definiert und gleich $\bar{1}$ ist; und aus $u \rightarrow v = \bar{1}$ folgt umgekehrt $u = u \odot \bar{1} \leq v$.

Nun ist allerdings $a \rightarrow b$ nicht notwendig für alle Paare von Fuzzymengen a und b definiert. Wir betrachten die Situation wiederum punktweise.

Definition 2.11 Es sei $\odot\colon\ [0,1]^2 \rightarrow [0,1]$ eine t-Norm. Dann sei für $a, b \in [0,1]$

$$a \rightarrow b = \max \{c \in \mathcal{F}(M) \colon\ a \odot c \leq b\},$$

vorausgesetzt, daß das Maximum existiert. Existiert das Maximum für alle $a, b \in [0,1]$, heißt $\rightarrow\colon\ [0,1]^2 \rightarrow [0,1]$ die zu $\odot$ gehörige *Residuumsfunktion* oder kurz das zu $\odot$ gehörige *Residuum*.

Für die Standard-t-Normen ergibt sich folgendes Bild; wir setzen Indizes wie in (3)–(5):

$$\to_L\colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto \begin{cases} 1-a+b & \text{für } a > b, \\ 1 & \text{für } a \leq b, \end{cases} \tag{7}$$

$$\to_P\colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto \begin{cases} \frac{b}{a} & \text{für } a > b, \\ 1 & \text{für } a \leq b, \end{cases} \tag{8}$$

$$\to_G\colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto \begin{cases} b & \text{für } a > b, \\ 1 & \text{für } a \leq b. \end{cases} \tag{9}$$

Eine t-Norm heißt im folgenden *linksstetig*, falls für jedes $a \in [0,1]$ die Funktion $(0,1] \to [0,1],\ \ x \mapsto x \odot a$ an jedem Punkt linksseitig stetig ist, d.h. $\lim_{x \nearrow a}(x \odot b) = a \odot b$ für $a, b \in [0,1]$ gilt. Wegen der Monotonie von $\odot$ heißt dies wiederum, daß für $a, a_0, a_1, ..., b \in [0,1]$ aus $a = \sup_i a_i$ gilt: $\sup_i(a_i \odot b) = a \odot b$.

Lemma 2.12 *Es sei* $\odot\colon\ [0,1]^2 \to [0,1]$ *eine t-Norm. Dann ist die Operation für jedes Paar* $a, b \in [0,1]$ *definiert, gdw* $\odot$ *linksstetig ist.*

Beweis. Es sei $\odot$ linksstetig und $a, b \in [0,1]$. Ist dann $c = \sup\ \{x \in [0,1] : x \odot a \leq b\}$, so folgt $c \odot a \leq b$; also existiert $a \to b$.

Umgekehrt sei die Operation $\to$ für alle Paare von Wahrheitswerten definiert und $a, a_1, a_2, \ldots, b \in [0,1]$ mit $a = \sup a_i$; zu zeigen ist $\sup_i(a_i \odot b) = a \odot b$. Es ist $\max\ \{x\colon\ x \odot b \leq \sup_i(a_i \odot b)\} \geq a$ und wegen der Monotonie von $\odot$ folglich $\sup_i(a_i \odot b) \leq a \odot b \leq \sup_i(a_i \odot b)$. Es folgt $\sup_i(a_i \odot b) = a \odot b$. $\square$

Damit ergibt sich, daß, wenn die Konjunktion von einer linksstetigen t-Norm induziert ist, auch auf der Menge aller Fuzzymengen die Operation $\to$ in allen Fällen definierbar ist.

Lemma 2.13 *Es sei* $\odot$ *eine linksstetige t-Norm mit zugehöriger Residuumsfunktion* $\to$*, und* $\mathcal{F}(M)$ *sei mit der von* $\odot$ *induzierten Konjunktion ausgestattet. Dann existiert* $u \to v$ *für alle* $u, v \in \mathcal{F}(M)$*, und es gilt*

$$(u \to v)(x) = u(x) \to v(x).$$

Beweis. Dies folgt unmittelbar daraus, daß die partielle Ordnung auf $\mathcal{F}$ punktweise erklärt ist. □

Wir können nun die auf einer linksstetigen t-Norm basierende Algebra $(\mathcal{F}(M); \wedge, \vee, \odot, \bar{0}, \bar{1})$ um die Operation $\rightarrow$ erweitern und gelangen zum zentralen algebraischen Begriff der Fuzzymengentheorie und Fuzzylogik.

Definition 2.14 Eine Algebra $(L; \wedge, \vee, \odot, \rightarrow, 0, 1)$ heiße *residuierter Verband*, falls Folgendes gilt:

(R1) $(L; \wedge, \vee, \odot, 0, 1)$ ist ein ℓ-Monoid.

(R2) Für $a, b \in L$ gilt

$$a \rightarrow b = \max \{c\colon\ a \odot c \leq b\}.$$

Man beachte, daß (R2) eine Existenzforderung darstellt; es besagt, daß $a \rightarrow b$ das größte Element c mit der Eigenschaft $a \odot c \leq b$ ist, daß insbesondere also ein solches Element existiert.

Die Bedingung (R2) läßt sich äquivalent wie folgt formulieren.

Lemma 2.15 *Es sei $(L; \wedge, \vee, \odot, 0, 1)$ ist ein ℓ-Monoid und $\rightarrow$ eine binäre Operation auf L. Dann ist $(L; \wedge, \vee, \odot, \rightarrow, 0, 1)$ ein residuierter Verband, gdw für alle $a, b, c \in L$ die folgende Bedingung erfüllt ist:*

$$a \odot b \ \leq c \quad \textit{genau dann, wenn} \quad a \ \leq b \rightarrow c. \tag{10}$$

Beweis. Es sei L ein residuierter Verband und $a, b, c \in L$. Dann ist $b \rightarrow c$ das größte Element x mit $x \odot b \leq c$. Also folgt aus $a \odot b \leq c$, daß $b \rightarrow c \geq a$. Wenn umgekehrt $a \leq b \rightarrow c$ gilt, so ist wegen der Monotonie $a \odot b \leq (b \rightarrow c) \odot b \leq c$. □

Wir gelangen nun zu einer weiteren wichtigen Operation auf $\mathcal{F}$: der Negation. Hier bieten sich zwei grundsätzliche Möglichkeiten; die erste baut auf das auf, was schon da ist; für eine Fuzzymenge a steht $\neg a$ für die schwächste Aussage, die zusammen mit a einen Widerspruch ergibt:

Definition 2.16 Es sei $\mathcal{F}$ ausgestattet mit der auf der linksstetigen t-Norm $\odot$ induzierten Konjunktion. Dann sei für $a \in \mathcal{F}(M)$

$$\neg a = a \rightarrow \bar{0}.$$

Sofern uns also die Operation $\rightarrow$ zur Verfügung steht, was die Regel ist, ist die Operation $\neg$ gemäß diesem Ansatz aus $\rightarrow$ definierbar. Wir geben an, was sich im Fall der Standard-t-Normen ergibt; wir setzen wieder entsprechende Indizes. Es gilt für $a \in [0,1]$

$$\neg_L\colon\ [0,1] \to [0,1], \quad a \mapsto 1-a, \tag{11}$$

$$\neg_P\colon\ [0,1] \to [0,1], \quad a \mapsto \begin{cases} 0 & \text{für } a > 0, \\ 1 & \text{für } a = 0, \end{cases} \tag{12}$$

$$\neg_G = \neg_P.$$

Die zur łukasiewiczschen t-Norm gehörige Negation ist diejenige, auf die man wohl käme, wenn man eine Negation unabhängig von einer Konjunktion zu definieren hätte. Es handelt sich um die Standardnegation, die auch im Fall zum Einsatz kommen kann, in denen man für die Konjunktion nicht $\odot_L$ verwendet.

Damit ist eine weitere, von der Konjunktion unabhängige Möglichkeit angesprochen, eine Negation zu definieren. Die Aufgabe ist ja, jeder Fuzzymenge $u\colon\ M \to [0,1]$ eine weitere Fuzzymenge $\sim u$ zuzuordnen, die das Gegenteil dessen modelliert, was u ausdrückt. Hierfür ist die punktweise angewandte Funktion $\neg_L$ die naheliegende Wahl; wir führen eine eigene Bezeichnung ein.

Definition 2.17 Die Funktion

$$\sim\colon\ [0,1] \to [0,1], \quad a \mapsto 1-a$$

heiße die *Standardnegation*. Für $u, v \in \mathcal{F}(M)$ definieren wir

$$(\sim u)(x) = \sim u(x), \quad x \in M.$$

Für die Standardnegation sind die folgenden Eigenschaften die charakteristischen.

Lemma 2.18 *Für die Funktion* $\sim\colon\ [0,1] \to [0,1]$ *gilt:*

(N1) $\sim$ *ist antiton: aus* $a \leq b$ *folgt* $\sim b \leq \sim a$ *für alle* $a, b \in [0,1]$*;*

(N2) $\sim$ *ist involutiv:* $\sim\sim a = a$ *für alle* $a \in [0,1]$.

Die letzte zu besprechende Operation ist schließlich die Disjunktion von Fuzzymengen. Wir führen eine solche nur dann ein, wenn wir auch die Standardnegation verwenden; die Disjunktion läßt sich dann in der bekannten Weise mittels Konjunktion und Standardnegation definieren.

Definition 2.19 Ist $\odot$ eine t-Norm und $\sim$ die Standardnegation, heißt

$$\oplus\colon\ [0,1]^2 \to [0,1],\quad (a,b) \mapsto {\sim}({\sim}a \odot {\sim}b)$$

die zu $\odot$ gehörige *t-Konorm*. Für $a, b \in \mathcal{F}(M)$ setzen wir

$$(u \oplus v)(x) = u(x) \oplus v(x) \quad \text{für } x \in M.$$

2.3 t-Normen

Eine t-Norm dient dazu, zwei Fuzzymengen im Sinne einer Konjunktion zu einer zusammenzufassen. Fuzzymengen werden punktweise verknüpft, und die t-Normen operieren auf der Menge $[0,1]$ der Wahrheitswerte; es handelt sich um Funktionen $\odot\colon\ [0,1]^2 \to [0,1]$ mit der Eigenschaft, daß $([0,1]; \wedge, \vee, \odot, 0, 1)$ ein ℓ-Monoid ist.

Leider existiert nicht nur keine einzelne Standard-t-Norm, sondern vielmehr eine solche Unzahl von t-Normen, daß derzeit keine Hoffnung auf eine Systematisierung besteht.

Wir halten zunächst einmal fest, daß zwei verschiedene t-Normen nicht notwendig als essentiell verschieden gelten sollten. Einem einzelnen, individuellen Wahrheitswert kommt in der Theorie der Fuzzymengen in der Regel keine Bedeutung zu, sofern es sich nicht um die Randwerte 0 oder 1 handelt; vielmehr kommt es auf die relative Lage von Wahrheitswerten an. Die Konstanten ebenso wie die relative Lage bleibt aber unter ordnungserhaltenden Selbstabbildungen des reellen Einheitsintervalls erhalten.

Definition 2.20 Eine Funktion $T\colon\ [0,1] \to [0,1]$ heißt *Ordnungsautomorphismus*, falls für alle $a, b \in [0,1]$ gilt: $a \leq b$, gdw $T(a) \leq T(b)$.

Zwei t-Normen $\odot_1\colon\ [0,1]^2 \to [0,1]$ und $\odot_2\colon\ [0,1]^2 \to [0,1]$ heißen *isomorph*, falls es einen Ordnungsautomorphismus T gibt, so daß

$$a \odot_2 b = T^{-1}(T(a) \odot_1 T(b))$$

für alle $a, b \in [0,1]$ gilt.

Isomorphe t-Normen werden im weiteren als nicht essentiell verschieden angesehen.

Wir entwickeln nun die Theorie der t-Normen unter Annahme der folgenden Bedingung; auf diesen Fall eingegrenzt ist eine Systematisierung möglich.

Definition 2.21 Eine t-Norm $\odot$ heiße *stetig*, falls $\odot$ als Funktion von $[0,1]^2$ nach $[0,1]$ stetig ist.

Die entscheidende algebraische Eigenschaft einer stetigen t-Norm ist die folgende.

Lemma 2.22 *Es sei $\odot$ eine stetige t-Norm. Dann gibt es für je zwei Elemente $a, b \in [0,1]$ mit $a \leq b$ ein $c \in [0,1]$ mit $a = b \odot c$.*

Insbesondere ist $a = b \odot (b \to a)$, und $b \to a$ ist das größte Element c mit $a = b \odot c$.

Beweis. Es sei $a \leq b$. Dann folgt aus $b \odot 0 = 0$ und $b \odot 1 = a$, daß es ein c mit $a = b \odot c$ gibt. Da in diesem Fall insbesondere $b \odot c \leq a$ gilt, ist $c \leq b \to a$; und es folgt $b \odot (b \to a) \leq a = b \odot c \leq b \odot (b \to a)$, also $a = b \odot (b \to a)$. □

Für das Verständnis der Wirkungsweise stetiger t-Normen machen wir uns die duale Sichtweise von Wahrheitswerten zu eigen, wie wir sie oben beschrieben haben: Ein Wahrheitswert kann als ein Wert verstanden werden, der bemißt, in welchem Maße eine Möglichkeit auszuschließen ist. Ordnet etwa die Fuzzymenge u dem Element m einer Bezugsmenge M den Wahrheitswert 1 zu, heißt dies, daß m mit der durch u modellierten Eigenschaft voll kompatibel ist; wir können dann auch sagen, daß m zum Grad 0, d.h. gar nicht ausgeschlossen wird. Analog können wir einen Wert $0,9$ dahingehend versteheen, daß m zum Grad $0,1$ ausgeschlossen wird. Ein Wert 0 schließlich bedeutet den Ausschluß zum Grad 1, d.h. den völligen Ausschluß.

Der Sinn einer Konjunktion liegt nun darin, zweierlei Informationen dieses Typs zusammenzufassen. Wird eine Möglichkeit von einer Aussage zu einem gewissen Grad s ausgeschlossen, von einer weiteren zu einem Grad t, sollte die Konjunktion beide Werte s und t in irgendeiner Weise kumulieren. Genau dies geschieht im Fall stetiger t-Normen in transparenter Weise.

Wir betrachten der Reihe nach die drei Beispielen des vorigen Abschnitts, für die offensichtlich Stetigkeit jeweils gegeben ist.

Im einfachsten Fall addieren sich die Ausschlußgrade einfach auf. Es sei $\odot_L$ die łukasiewiczsche t-Norm, d.h. wir haben

$$a \odot_L b = (a + b - 1) \vee 0$$

für $a, b \in [0, 1]$. Um diese Funktion besser zu begreifen, benennen wir die Wahrheitswerte um: Wir verschieben sie von $[0, 1]$ nach $[-1, 0]$. Es sei

$$R\colon [0, 1] \to [-1, 0], \quad a \mapsto a - 1,$$

und wir betrachten die mittels R transformierte Funktion auf $[-1, 0]$. Es ergibt sich

$$a \odot'_L b = R(R^{-1}a \odot_L R^{-1}b) = (a + b) \vee -1 \tag{13}$$

für $a, b \in [-1, 0]$. Mit anderen Worten addieren sich die – nunmehr negativen – Werte einfach auf, und ist die Summe kleiner -1, wird -1 genommen. Dies ist mit dem obigen Bild sich kumulierender Ausschlußgrade konsistent; ein Wert von $-0,4$ beispielsweise entspricht dem Wahrheitwert $0,6$ und damit einem Ausschlußgrad von $0,4$.

Die łukasiewiczsche t-Norm ist eine häufig verwendete t-Norm; wir wollen sie unter den t-Normen charakterisieren. Eine t-Norm zu charakterisieren heißt dabei: Eigenschaften zu benennen, so daß jede t-Norm mit diesen Eigenschaften zur łukasiewiczschen t-Norm isomorph ist.

Definition 2.23 Eine t-Norm $\odot$ heiße *archimedisch*, falls es für jedes Paar $a, b \in (0, 1)$ mit $a \leq b$ ein $n \geq 1$ gibt, so daß $b^n \leq a$. Hierbei bedeutet $b^n = \underbrace{b \odot \ldots \odot b}_{n\text{-mal}}$.

Weiter ist ein *Nullteiler* in bezug auf eine t-Norm $\odot$ ein Element $a \in (0, 1)$, so daß es ein $b \in (0, 1)$ gibt mit $a \odot b = 0$.

Lemma 2.24 *Es sei $\odot$ eine archimedische stetige t-Norm. Dann gilt:*

(i) *Für $a, b, c \in [0, 1]$ folgt aus $a \odot b = a \odot c > 0$. Dann gilt $b = c$.*

(ii) *Es gibt einen Nullteiler $a < 1$ in bezug auf $\odot$, gdw jedes Element $a < 1$ ein Nullteiler ist, gdw es für jedes Element $a < 1$ ein n mit $a^n = 0$ gibt.*

(iii) *Für jedes* $a \in [0,1)$ *gibt es ein größtes Element* $b \geq a$ *mit* $b^2 = a$. *Es gilt* $b > a$, *falls entweder* $a > 0$ *oder von* 1 *verschiedene Nullteiler in bezug auf* $\odot$ *existieren.*

Beweis. (i) Es gelte $a \odot b = a \odot c > 0$, und wir nehmen an, daß $b < c$. Dann gibt es wegen Lemma 2.22 ein $d < 1$ mit $c \odot d = b$. Folglich gilt $a \odot c = a \odot b = a \odot c \odot d$, d.h. $a \odot c = a \odot c \odot d^n$ für jedes n. Wegen der Archimedizität gilt jedoch $\bigwedge_n d^n = 0$. Also folgt $a \odot c = 0$ im Widerspruch zur Vorgabe.

(ii) Es sei $a < 1$ ein Nullteiler. Also $a \odot c = 0$ für ein $c > 0$. Ist $b < 1$ ein weiteres Element, gilt wegen der Archimedizität $b^n \leq a$ für ein hinreichend großes n; also $b \odot b^{n-1} \odot c = 0$. Es sei $m \geq 0$ maximal mit der Eigenschaft $b^m \odot c = 0$; dann ist $1 \leq m \leq n-1$. Dann ist $b \odot b^{m-1} \odot c = 0$ und $b^{m-t} \odot c > 0$. Also ist b ebenfalls ein Nullteiler. Damit ist die Äquivalenz der ersten beiden Aussagen bewiesen.

Wiederum sei $a < 1$ ein Nullteiler, also $a \odot c = 0$ für $c > 0$. Es sei n so groß, daß $a^{n-1} \leq c$; dann ist $a^n = 0$. Gilt umgekehrt $a^n = 0$, sei $m \leq n$ minimal mit der Eigenschaft $a^m = 0$. Dann ist $a \odot a^{m-1} = 0$ und $a^{m-1} > 0$, also a ein Nullteiler. Damit ist auch die Äquivalenz der ersten und letzten Aussage gezeigt.

(iii) Die Funktion $[0,1] \to [0,1],\ \ x \mapsto x^2$ ist stetig; und es gilt $0^2 = 0$ sowie $1^2 = 1$; somit gibt es ein größtes b mit $b^2 = a$. Klarerweise ist $b \geq a$.

Angenommen, $b = a > 0$. Dann wäre $a^2 = a$, im Widerspruch zu Teil (i).

Angenommen, $b = a = 0$, und $\odot$ besitze Nullteiler. Ersteres heißt, daß 0 das größte und folglich einzige Element x ist mit $x^2 = 0$. Letzteres heißt, daß es $c, d \in (0,1)$ mit $c \odot d = 0$ gibt, d.h. $c^2 = 0$, wenn etwa c das kleinere der beiden Elemente ist. Dies ist ein Widerspruch; also $b > 0 = a$. □

Theorem 2.25 *Die łukasiewiczsche t-Norm ist eine stetige t-Norm;* $\odot_L$ *ist archimedisch; und jedes Element von* $(0,1)$ *ist ein Nullteiler.*

Jede t-Norm mit diesen Eigenschaften ist zu $\odot_L$ *isomorph.*

Beweis. Die drei behaupteten Eigenschaften sind leicht anhand der Darstellung (13) zu verifizieren.

Es sei nun $\odot$ eine beliebige t-Norm, für die ebenfalls diese drei Eigenschaften gelten. Wir verlegen wiederum den Wertebereich ins Negative, d.h. wir betrachten die Funktion

$$\odot' \colon [-1, 0]^2 \to [-1, 0], \quad (a, b) \mapsto ((a + 1) \odot (b + 1)) - 1. \tag{14}$$

Wir setzen nun $a_0 = -1$; gemäß Lemma 2.24(iii) wählen wir a_1 als das größte Element mit $a_1^2 = a_0$; und so weiter. Dann erhalten wir eine Folge $-1 = a_0 < a_1 < a_2 < \ldots < 0$ mit $a_i = a_{i+1}^2$ für jedes $i = 0, \ldots$.

Wir behaupten: $\bigvee_i a_i = 0$. Beweis: Gäbe es ein $b < 0$ mit $b \geq a_i$ für alle i, wäre $b^{2^k} \geq a_i^{2^k}$ für alle k und i, also insbesondere $b^{2^k} \geq a_1 > -1$ für alle k, im Widerspruch zu Lemma 2.24(ii).

Wir behaupten weiter: Für jedes i gilt $-1 < a_i^{2^i-1} < a_i^{2^i-2} < \ldots < a_i < 0$. Beweis: Wir zeigen $-1 < a_i^{2^i-1}$; dann folgen die übrigen Ungleichungen aus Lemma 2.24(i). Angenommen sei $a_i^{2^i-1} = -1$. Dann wäre $(a_{i+1}^{2^i-1})^2 = -1$; andererseits ist $a_1 = a_{i+1}^{2^i}$ maximal mit der Eigenschaft $a_1^2 = -1$; es folgt $-1 < a_{i+1}^{2^i-1} = a_{i+1}^{2^i}$, im Widerspruch zu Lemma 2.24(i).

Wir setzen nun $D = \{a_i^m \colon 0 \leq m \leq 2^i\}$. Dann ist D dicht in $[-1, 0]$. Weiter gilt $m_1 \cdot a_{i_1} = m_2 \cdot a_{i_2}$, gdw $\frac{m_1}{2^{i_1}} = \frac{m_2}{2^{i_2}}$. Damit können wir die Funktion

$$T \colon D \to [-1, 0], \quad a_i^m \mapsto -\frac{m}{2^i}$$

definieren. Dann ist T eine ordnungserhaltende Bijektion zwischen zwei dichten Teilmengen von $[0, 1]$. Weiter gilt $T(a_{i_1}^{m_1}) \odot T(a_{i_2}^{m_2}) = (\frac{m_1}{2^{i_1}} + \frac{m_2}{2^{i_2}}) \vee -1$. Da sowohl $\odot$ als auch $\odot_L'$ stetig sind, folgt der Satz. $\square$

Es sei weiter $\odot_P$ die Produkt-t-Norm:

$$a \odot_P b = a \cdot b.$$

Obwohl es nicht den Anschein macht, liegen die Verhältnisse hier ähnlich wie bei der łukasiewiczschen t-Norm. Wir benennen ein weiteres Mal die Wahrheitswerte um; diesmal verwenden wir die Menge

$$\overline{\mathbb{R}^-} = \{r \in \mathbb{R} \colon r \leq 0\} \cup \{-\infty\},$$

worin $-\infty$ ein zusätzliches Element darstellt, für das $-\infty < r$ für alle $r \in \mathbb{R}^-$ gelte. Wir setzen

$$R\colon\ [0,1] \to \overline{\mathbb{R}^-}, \quad a \mapsto \begin{cases} \ln a & \text{für } a \in (0,1] \\ -\infty & \text{für } a = 0. \end{cases}$$

Für die transformierte Produkt-t-Norm ergibt sich

$$a \odot'_P b == R(R^{-1}a \odot_L R^{-1}b) = a + b$$

für $a, b \in \overline{\mathbb{R}^-}$, wobei wir $-\infty + a = a + (-\infty) = -\infty + (-\infty) = -\infty$ gesetzt haben. Bezüglich der neuen Wahrheitswertemenge wird also wiederum einfach aufaddiert; im Unterschied zur łukasiewiczschen t-Norm gibt wird dabei allerdings keine Grenze erreicht, solange es sich nicht um den Wahrheitswert für „falsch“ handelt.

Theorem 2.26 *Die Produktnorm $\odot_P$ ist eine stetige t-Norm; $\odot_P$ ist archimedisch; und kein Element von $(0,1)$ ist ein Nullteiler.*

Jede t-Norm mit diesen Eigenschaften ist zu $\odot_P$ isomorph.

Beweis. Der Beweis verläuft ähnlich wie im Fall der łukasiewiczschen t-Norm. Die Rolle, die dort vom Element -1 gespielt wird, wird jetzt von einem beliebigen von 0 verschiedenen Element übernommen. □

Lemma 2.24(ii) besagt, daß im Fall der Archimedizität entweder alle Elemente im offenen Intervall $(0,1)$ Nullteiler sind oder keines. Also folgt aus den Sätzen 2.25 und 2.26, daß es bis auf Isomorphie nur zwei verschiedene archimedische t-Normen gibt: die łukasiewiczsche und die Produkt-t-Norm.

Wir kommen schließlich zur gödelschen t-Norm.

Definition 2.27 Es sei $\odot$ eine t-Norm. Ein Element a mit $a \odot a = a$ heiße *idempotent* in bezug auf $\odot$.

Theorem 2.28 *In bezug auf die gödelsche t-Norm ist jedes Element idempotent.*

Jede t-Norm mit dieser Eigenschaft ist die gödelsche t-Norm.

Beweis. Für die gödelsche t-Norm $\odot_G$ gilt per Definition $a \odot_G a = a \wedge a = a$, d.h. jedes Element ist idempotent.

Ist umgekehrt $\odot$ eine t-Norm, in bezug auf die alle Elemente idempotent sind, dann gilt für $a, b \in [0,1]$ mit $a \leq b$, daß $a = a \odot 1 \geq a \odot b \geq a \odot a = a$, also $a \odot b = a$. Also ist in diesem Fall $\odot = \wedge = \odot_G$. $\square$

Eine allgemeine stetige t-Norm setzt sich in einfacher Weise aus den drei Standard-t-Normen zusammen. Die im folgenden Satz beschriebene Konstruktion heißt *ordinale Summe*.

Theorem 2.29 Es sei $\odot$ eine t-Norm. Dann gibt es abschlossene Intervalle $[l_i, r_i]$, $i \in I$, deren je zwei sich in maximal einem Punkt überschneiden, deren Vereinigung gleich $[0,1]$ ist und für die Folgendes gilt. Für jedes i ist $[l_i, r_i]$ unter $\odot$ abgeschlossen; ist dann T_i die lineare ordnungserhaltende Abbildung, die das Intervall $[l_i, r_i]$ auf das Intervall $[0,1]$ abbildet, ist

$$\odot_i \colon\ [0,1]^2 \to [0,1], \quad (a,b) \mapsto T_i(T_i^{-1}(a) \odot T_i^{-1}(b))$$

entweder isomorph zur łukasiewiczschen oder isomorph zur Produkt- oder gleich der gödelschen t-Norm. Weiter gilt für a, b, die nicht im selben Intervall $[l_i, r_i]$ liegen, daß $a \odot b = a \wedge b$.

Beweis. Es sei K die Menge aller Idempotente in bezug auf $\odot$. K ist, da $\odot$ stetig ist, eine abgeschlossene Teilmenge von $[0,1]$. Es seien $[0,1]$ überdeckende und sich allenfalls jeweils an nur Randpunkten überschneidende Intervalle $[l_i, r_i]$ so gewählt, daß $[l_i, r_i]$ mit K entweder nur die beiden Randpunkte gemeinsam hat oder aber ganz in K liegt.

Es sei eines der i gegeben. Ist $[l_i, r_i] \subseteq K$, ist die mittels T_i transformierte Funktion offenbar die gödelsche t-Norm.

Im anderen Fall sei $a, b \in [l_i, r_i]$. Aus $l_i \leq a, b \leq r_i$ folgt $l_i = l_i \odot l_i \leq a \odot b \leq r_i$, d.h. $a \odot b \in [l_i, r_i]$. Also ist $[l_i, r_i]$ unter $\odot$ abgeschlossen, und $\odot_i$ läßt sich definieren wie angegeben.

(T1), (T2) und (T4) gelten für $\odot_i$, weil diese Eigenschaften für $\odot$ gelten. Weiter sei $a \in [l_i, r_i]$; es sei $z \in [0,1]$ minimal mit der Eigenschaft $a = r_i \odot z$; dann ist $a \odot r_i = r_i \odot z \odot r_i = r_i \odot z = a$. Es folgt, daß für $\odot_i$ auch (T3) gilt. Somit ist klar, daß $\odot_i$ eine t-Norm ist, die zudem stetig ist.

Es sei nun $a \in [0,1]$ beliebig; wir zeigen, daß $\inf_n a^n$ in bezug auf $\odot$ idempotent ist. In der Tat gilt wegen der Stetigkeit $\inf_n a^n \odot \inf_n a^n = \inf_n(a^n \odot \inf_m a^m) = \inf_n \inf_m a^{n+m} = \inf_n a^n$.

Ist nun $l_i < a < b < r_i$, gilt $b^n \leq a$ für hinreichend großes $n \geq 1$. Denn wäre $a < b^n$ für alle n, so gälte $a \leq \bigwedge_n b^n$; $\bigwedge_n b^n$ ist aber ein Idempotent, deren es zwischen a und b keine gibt. Für die t-Norm $\odot_i$ folgt Archimedizität. Mithin ist $\odot_i$ entweder zur łukasiewiczschen oder zur Produkt-t-Norm isomorph. □

Eine einfache Folgerung hieraus besagt, welche Bedeutung der łukasiewiczschen t-Norm zukommt.

Theorem 2.30 *Es sei $\odot$ eine linksstetige t-Norm und $\rightarrow$ die zugehörige Residuumsfunktion. Dann sind sowohl $\odot$ als auch $\rightarrow$ stetig, gdw $\odot$ isomorph zur łukasiewiczschen t-Norm ist.*

Beweis. Es sei $\odot$ stetig. Dann ist $\odot$ gemäß der in Satz 2.29 beschriebenen Konstruktion aus łukasiewiczschen, Produkt- und gödelschen t-Normen zusammengesetzt. Wie sich im Hinblick auf (8) und (9) ergibt, besitzt $\rightarrow$ im zweiten Argument Unstetigkeitsstellen, falls dabei t-Normen der letzteren beiden Typen beteiligt sind. Also ist $\odot$ nur aus łukasiewiczschen t-Normen zusammengesetzt. Handelt es sich allerdings um mehr als eine, ist $[0,1] \rightarrow [0,1], \;\; a \mapsto a \rightarrow 0$ bei 0 unstetig. Die Behauptung folgt. □

3 Analytische Theorie der Fuzzymengen

3.1 Standardfuzzymengen

Vage Eigenschaften nehmen vornehmlich auf Strukturen Bezug, die sich mit reellen Zahlen beschreiben lassen. Daher sind die meisten in der Praxis verwendeten Fuzzymengen solche über $\mathbb{R}^n$ für ein $n \geq 1$. Unter diesen wiederum ist der Fall $n = 1$ der mit Abstand häufigste. Bislang haben wir Fuzzymengen als beliebige Funktionen von einer unstrukturierten Menge in das reelle Einheitsintervall erachtet; nunmehr definieren wir einen Typ Fuzzymengen über dem $\mathbb{R}^n$, der tatsächlich von Relevanz zu sein verspricht.

Im weiteren sei $n \geq 1$ fix. Im vorliegenden Teilabschnitt engen wir den Begriff der Fuzzymenge in geeigneter Weise ein.

Definition 3.1 Es sei $u\colon \mathbb{R}^n \to [0,1]$ eine Fuzzymenge. Dann heiße

$$\text{supp } u \;\stackrel{\text{def}}{=}\; \overline{\{x \in \mathbb{R}^n \colon\, u(x) > 0\}}$$

der *Träger* von u und

$$\ker u \;\stackrel{\text{def}}{=}\; \{x \in \mathbb{R}^n \colon\, u(x) = 1\}$$

der *Kern* von u.

Weiter heiße u *fuzzykonvex*, falls für je zwei Punkte $a, b \in \mathbb{R}^n$ und jedes $\lambda \in (0,1)$

$$u(\lambda a + (1-\lambda)b) \;\geq\; u(a) \wedge u(b) \tag{15}$$

gilt. Schließlich heiße u *von oben halbstetig*, falls für jedes $x \in \mathbb{R}^n$ und jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so daß $u(y) < u(x) + \varepsilon$ für alle $y \in \mathbb{R}^n$ in der δ-Umgebung von x.

u heiße eine *Standardfuzzymenge*, falls (i) der Träger von u beschränkt ist, (ii) der Kern von u nicht leer ist, (iii) u fuzzykonvex ist und (iv) u von oben halbstetig ist. Es sei $\mathcal{E}^n$ die Menge aller Standardfuzzymengen.

Der Leitgedanke ist, Fuzzymengen über dem $\mathbb{R}^n$ als verallgemeinerte Punkte des $\mathbb{R}^n$ zu betrachten. Bedingung (i) sagt demgemäß, daß eine Fuzzymenge außerhalb einer Umgebung eines Punktes verschwindet. Bedingung (ii) stellt sicher, daß innerhalb dieser Umgebung wenigstens ein Punkt als der unscharfen Menge voll zugehörig zu rechnen ist. Bedingung (iii) spiegelt das Bild, daß die Menge um einen Punkt herum gruppiert anzusehen ist. Bedingung (iv) stellt eine gewisse Glattheit der Fuzzymenge sicher. Hier könnte eigentlich mehr erwartet werden; man könnte Stetigkeit verlangen. Der Grund, nur von oben halbstetige Funktionen zuzulassen, liegt allein darin, daß man charakteristische Funktionen zur Verfügung haben möchte:

Lemma 3.2 *Es sei $A \subseteq \mathbb{R}^n$ nichtleer, konvex und kompakt. Dann ist χ_A eine Standardfuzzymenge.*

Beweis. Daß A kompakt ist, heißt, daß A eine beschränkte und abgeschlossene Teilmenge des $\mathbb{R}^n$ ist. Aus der Beschränktheit von A folgt Bedingung (i) der Definition einer Standardfuzzymenge.

Da A nichtleer ist, gilt auch (ii).

Weiter ist Gleichung (15) zu prüfen. Es sei $a, b \in \mathbb{R}^n$. Ist a oder b nicht in A, gilt (15) klarerweise; sind a und b beide in A, dann wegen der Konvexität von A auch die Verbindungslinie zwischen a und b, weswegen (15) auch in diesem Fall gilt.

Es verbleibt zu zeigen, daß χ_A bei $a \in \mathbb{R}^n$ von oben halbstetig ist. Ist $a \in A$, ist dies der Fall, weil χ_A ohnehin nur auf Werte ≤ 1 abbildet. Ist $a \notin A$, gibt es eine Umgebung von a, die außerhalb von A liegt und somit χ_A ganz auf 0 abgebildet wird. □

Für $n = 1$ ergibt sich folgendes Bild.

Lemma 3.3 *Es sei $u\colon \mathbb{R} \to [0, 1]$. Dann ist u eine Standardfuzzymenge, gdw die folgenden Eigenschaften erfüllt sind:*

(i) *Es gibt $l, r \in \mathbb{R}$, so daß $u(x) = 0$ ist für $x < l$ und $x > r$.*

(ii) *Es gibt ein $t \in \mathbb{R}$ mit $u(t) = 1$.*

(iii) *u besitzt an jedem Punkt den links- und rechtsseitigen Grenzwert und bildet auf einen Wert ab, der mindestens so groß wie jeder der beiden Grenzwerte.*

(iv) *Es gibt ein $t \in \mathbb{R}$, so daß u auf $(-\infty, t]$ monoton steigend und auf $[t, \infty)$ monoton fallend ist.*

Beweis. Es sei $u \in \mathcal{E}^n$. (i) und (ii) gelten aufgrund der entsprechenden Bedingungen an eine Standardfuzzymenge.

Weiter sei $t \in \mathbb{R}$ mit $u(t) = 1$. Für $a < b \leq t$ gilt $u(a) \leq u(b)$, da u fuzzykonvex ist; also ist u auf $(-\infty, t]$ monoton steigend. Ähnlich ist zu ersehen, daß u auf $[t, \infty)$ monoton fallend ist. Also gilt (iv).

Es sei $a \in \mathbb{R}$. Für ein hinreichend kleines $\delta > 0$ gilt: u ist auf $[a-\delta, a]$ monoton steigend oder fallend. Da u nach oben und unten beschränkt ist, folgt, daß der linkseitige Grenzwert von u am Punkt a existiert. Ähnlich folgt, daß auch rechtsseitige Grenzwert existiert. Weiter sei g der größere der beiden Grenzwerte; wir nehmen an, dies sei der linksseitige. Wäre nun $u(a) < g$, gäbe es für kein $\delta < g - u(a)$ ein $\varepsilon > 0$, so daß $u(x) < u(a) + \delta$ für alle $x \in [a - \varepsilon, a]$. Also muß $u(a) \geq g$ gelten, und (iii) ist gezeigt.

Umgekehrt erfülle u die Bedingung (i) bis (iv). Es gelten klarerweise Bedingungen (i) und (ii) der Definition einer Standardfuzzymenge. Weiter folgt aus (iii) leicht, daß u von oben halbstetig ist; und aufgrund von (iv) ist u offenbar u fuzzykonvex. $\square$

3.2 Die Schnitte einer Fuzzymenge

Um auf der Menge der Standardfuzzymenge in einfacher Weise arithmetische Operationen sowie eine Metrik einführen zu können, benötigen wir eine alternative Darstellung der Fuzzymengen.

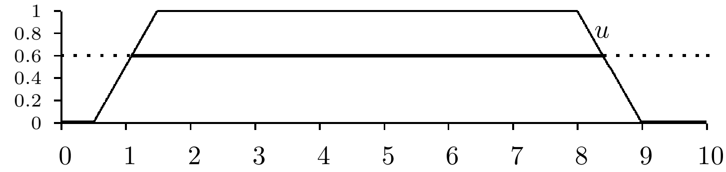
Definition 3.4 Es sei u eine Fuzzymenge über $\mathbb{R}^n$ und $\alpha \in (0, 1]$. Dann heiße die Menge

$$[a]_\alpha = \{p \in \mathbb{R}^n \colon\ u(p) \geq \alpha\}$$

der α-Schnitt von u.

Die Bezeichnung α-Schnitt weist auf das hin, was man sich vorzustellen hat: den Schnitt der horizontal liegenden Ebene auf Höhe α mit dem Graphen einer Fuzzymenge u; die Punkte des Graphen, die über die Ebene hinausragen, projizieren auf den α-Schnitt von u.

Der $0,6$-Schnitt dieser Fuzzymenge ist das Intervall $[1\frac{1}{10}, 8\frac{2}{5}]$:



Im Fall der Fuzzymenge u in diesem Beispiel ist ersichtlich, daß u eindeutig bestimmt ist, wenn für jedes $\alpha \in (0,1]$ der α-Schnitt bekannt ist. Dies ist für Standardfuzzymengen allgemein der Fall. Wir entwickeln in diesem Teilabschnitt den genauen Zusammenhang zwischen den Standardfuzzymengen einerseits und den zugehörigen Schnitten andererseits.

Lemma 3.5 *Es sei $u \in \mathcal{E}^n$ und $\alpha \in (0,1]$. Dann ist $[u]_\alpha$ eine kompakte, konvexe Teilmenge des $\mathbb{R}^n$.*

Insbesondere ist $[u]_\alpha$ im Fall $n = 1$ ein abgeschlossenes Intervall.

Beweis. $[u]_\alpha$ ist Teilmenge des Trägers von u und daher beschränkt. Weiter sei $p \notin [u]_\alpha$, also $u(p) < \alpha$. Da u von oben halbstetig ist, gilt dann für alle Punkte einer offenen Umgebung von p, daß u sie auf Werte $< \alpha$ abbildet. Also ist das Komplement von $[u]_\alpha$ offen, d.h. $[u]_\alpha$ selbst abgeschlossen. Damit ist gezeigt, daß $[u]_\alpha$ kompakt ist.

Es seien $a, b \in [u]_\alpha$. Dies heißt $u(a), u(b) \geq \alpha$. Da u fuzzykonvex ist, werden alle Punkte p auf der Verbindungslinie zwischen a und b ebenfalls auf Werte $u(p) \geq \alpha$ abgebildet. Also ist auch $u(p) \in [u]_\alpha$, d.h. $[u]_\alpha$ ist konvex. $\square$

Definition 3.6 Es bezeichne $\mathcal{K}^n$ die Menge aller kompakten, konvexen, nichtleeren Teilmengen des $\mathbb{R}^n$. Wir statten $\mathcal{K}^n$ mit der Teilmengenbeziehung $\subseteq$ aus.

Es ist klar, daß $\mathcal{K}^n$ mit dieser Vereinbarung zu einem Poset wird. Es handelt sich um einen Verband; das Infimum ist der Durchschnitt und das Supremum die konvexe Hülle der Vereinigung zweier Mengen aus $\mathcal{K}^n$. Allerdings gibt es weder ein kleinstes noch ein größtes Element in $\mathcal{K}^n$.

Definition 3.7 Für jedes $u \in \mathcal{E}^n$ sei

$$s_u \colon (0,1] \to \mathcal{K}^n, \quad \alpha \mapsto [u]_\alpha$$

die *Schnittfunktion* von u.

Theorem 3.8 *Die Schnittfunktion s_u einer Fuzzymenge $u \in \mathcal{E}^n$ hat folgende Eigenschaften:*

(i) *s_u ist monoton fallend: Aus $\alpha \le \beta$ folgt $[u]_\beta \supseteq [u]_\alpha$.*

(ii) *s_u ist beschränkt: Es gibt ein $A \in \mathcal{K}^n$ mit $[u]_\alpha \subseteq A$ für alle α.*

(iii) *s_u ist linksseitig ordnungsstetig: Für $\alpha \in (0,1]$ gilt $\bigcap_{\beta<\alpha}[u]_\beta = [u]_\alpha$.*

Weiter sei s eine monoton fallende, beschränkte, linksseitig ordnungsstetige Funktion von $(0,1]$ nach $\mathcal{K}^n$. Dann ist s die Schnittfunktion genau einer Fuzzymenge $u \in \mathcal{E}^n$. Für diese gilt

$$u(p) = \sup\, \{\beta \in (0,1] \colon p \in s(\beta)\}. \tag{16}$$

Beweis. Aus der Definition der Schnitte folgt unmittelbar, daß s_u monoton fallend ist. Daraus, daß u einen beschränkten Träger hat, folgt, daß s_u beschränkt ist. Und für $\alpha \in (0,1]$ gilt $\bigcap_{\beta<\alpha}[u]_\beta = \{p \in \mathbb{R}^n : u(p) \ge \beta \text{ für alle } \beta < \alpha\} = [u]_\alpha$.

Umgekehrt sei s eine monoton fallende, beschränkte, linksseitig ordnungsstetige Funktion von $(0,1]$ nach $\mathcal{K}^n$; wir definieren u gemäß (16).

Ist $p \notin s(\beta)$ für alle $\beta \in (0,1]$, dann ist $u(p) = \sup \emptyset = 0$; es folgt, daß u einen beschränkten Träger hat. Weiter ist $s(1)$ nach Definition nichtleer; für $p \in s(1)$ gilt $u(p) = 1$; also hat u einen nichtleeren Kern.

Es sei $a, b \in \mathbb{R}^n$ und $c = \lambda a + (1-\lambda)b$ für ein $\lambda \in (0,1)$. Es ist $a \in s(\beta)$ für alle $\beta < u(a)$ und $b \in s(\beta)$ für alle $\beta < u(b)$, also $a, b \in s(\beta)$ für alle $\beta < u(a) \wedge u(b)$. Wegen der Konvexität ist auch $c \in s(\beta)$ für alle $\beta < u(a) \wedge u(b)$ und folglich $u(c) \ge u(a) \wedge u(b)$. Somit ist u fuzzykonvex.

Es verbleibt zu zeigen, daß u von oben halbstetig ist. Es sei $p \in \mathbb{R}^n$ mit $u(p) < 1$ und $0 < \varepsilon < 1 - u(p)$. Es sei $\alpha = u(p) = \sup\{\beta \in (0,1] \colon p \in s(\beta)\}$; da s linksseitig ordnungsstetig ist, folgt $p \in s(\alpha)$, aber $p \notin s(\beta)$ für $\beta > \alpha$

und insbesondere für $\beta = \alpha + \varepsilon$. Da $s(\alpha + \varepsilon)$ abgeschlossen ist, liegt eine ganze Umgebung von p außerhalb $s(\alpha + \varepsilon)$. Also gilt für hinreichend kleines δ, daß $q \notin s(\alpha + \varepsilon)$, d.h. $u(q) < \alpha + \varepsilon$, für $d(p, q) < \delta$. $\square$

3.3 Lineare und metrische Struktur für die Standardfuzzymengen

Die Elemente der Standardfuzzymengen sind als verallgemeinerte Punkte des $\mathbb{R}^n$ aufzufassen. Insofern sollte $\mathcal{E}^n$ mit derselben Art von Struktur ausgestattet werden wie schon $\mathbb{R}^n$. $\mathbb{R}^n$ trägt zwei wichtige Strukturen: eine lineare und eine metrische.

Zur ersteren: Der $(\mathbb{R}^n; +, \cdot)$ ist ein linearer Raum über $\mathbb{R}$. Hierbei ist $+$ die Summe, definiert für zwei Elemente von $\mathbb{R}^n$ komponentenweise durch

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix};$$

und $\cdot$ ist die Multiplikation mit Werten $\lambda \in \mathbb{R}$:

$$\lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}.$$

Zur letzteren: Im $\mathbb{R}^n$ läßt sich eine Norm einführen, etwa die euklidische:

$$\left|\left| \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \right|\right| = a_1^2 + \ldots a_n^2;$$

Mittels der Norm $||\cdot||$ läßt sich eine Metrik $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definieren:

$$d(a, b) = ||a - b||\,,$$

worin $a, b \in \mathbb{R}^n$. Die Metrik d macht $\mathbb{R}^n$ zu einem metrischen Raum.

Im Fall $n = 1$ ist $+$ die gewöhnliche Addition, $\cdot$ die Multiplikation reeller Zahlen, und $d(a, b) = |a - b|$ ist der Absolutwert der Differenz zweier reeller Zahlen a und b.

Wir möchten die lineare und metrische Struktur nun von $\mathbb{R}^n$ auf $\mathcal{E}^n$ übertragen. Wir tun dies in zwei Schritten; im ersten Schritt betrachten wir die Menge der kompakten konvexen Teilmengen des $\mathbb{R}^n$, und im zweiten Schritt führen wir die relevanten Operationen schnittweise auf $\mathcal{E}^n$ ein.

Definition 3.9 Für je zwei Mengen $A, B \in \mathcal{K}^n$ setzen wir

$$A + B = \{a + b\colon\ a \in A,\ b \in B\};$$

für eine Menge $A \in \mathcal{K}^n$ und ein positives $\lambda \in \mathbb{R}$ setzen wir

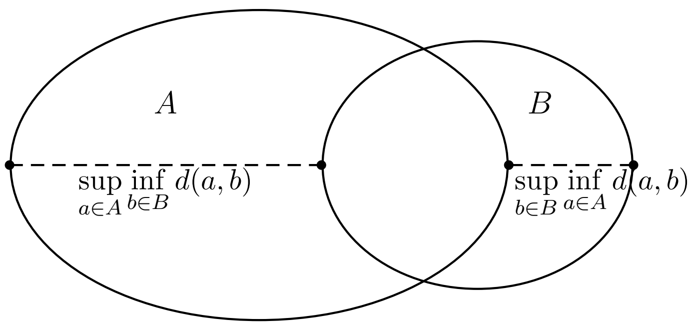
$$\lambda A = \{\lambda a\colon\ a \in A\}.$$

Weiter sei der *Hausdorff-Abstand* zwischen zwei Mengen für $A, B \in \mathcal{K}^n$ durch

$$d_H(A, B) = \sup_{a\in A} \inf_{b\in B} d(a, b) \ \vee\ \sup_{b\in B} \inf_{a\in A} d(a, b)$$

definiert.

Die elementweise erklärte Addition macht $\mathcal{K}^n$ nicht zu einem linearen Raum; es gibt bezüglich $+$ keine Inversen. Vielmehr heißt $\mathcal{K}^n$ zusammen mit der Addition $+$ und der Multiplikation mit positiven reellen Zahlen ein *positiver Konus.*



Der Hausdorffabstand zwischen den Mengen A und B
ist der linke als der größere der beiden angegebenen Werte.

Für die Hausdorff-Abstand ergibt sich im Fall $n = 1$ ergibt sich: Ist $a_1 \leq a_2$ und $b_1 \leq b_2$, dann ist $d([a_1, a_2], [b_1, b_2]) = |a_1 - b_1| \vee |a_2 - b_2|$.

Der Hausdorff-Abstand ist weiter eine Metrik: Für $A, B, C \in \mathcal{K}^n$ gilt $d(A, B) \geq 0$, $d(A, B) = d(B, A)$, $d(A, C) \leq d(A, B) + d(B, C)$ und $d(A, B) = 0$, gdw $A = B$. Wir erachten im weiteren $\mathcal{K}^n$ als metrischen Raum.

Die für $\mathcal{K}^n$ eingeführte Struktur läßt sich nun in einfachster Weise auf die Standardfuzzymengen übertragen, wenn man nur von der im vorhergehenden Abschnitt eingeführten Darstellung mittels Schnitten ausgeht.

Definition 3.10 Für Funktionen $s, t\colon (0,1] \to \mathcal{K}^n$ und positive reelle Zahlen λ definieren wir $s + t$ und λs gemäß

$$(s+t)(\alpha) = s(\alpha) + t(\alpha), \tag{17}$$
$$(\lambda s)(\alpha) = \lambda s(\alpha) \tag{18}$$

für $\alpha \in (0,1]$.

Lemma 3.11 *Es seien $s, t : (0,1] \to \mathcal{K}^n$ monoton fallende, beschränkte, linksseitig ordnungsstetige Funktionen. Dann ist auch $s+t$ sowie λs für $\lambda \geq 0$ monoton fallend, beschränkt und linksseitig ordnungsstetig.*

Beweis. Wir zeigen nur, daß $s + t$ linksseitig ordnungsstetig ist; die übrigen Behauptungen ergeben sich aus den Definitionen unmittelbar.

Zu zeigen ist $\cap_{\beta<\alpha}(s(\beta) + t(\beta)) = s(\alpha) + t(\alpha)$ für $\alpha \in (0,1]$. Die Relation „$\supseteq$" ist wegen der Monotonie von s und t klar. Umgekehrt sei $p \in s(\beta) + t(\beta)$ für alle $\beta < \alpha$; dies heißt, daß es für jedes $i = 2, 3, ...$ ein $x_i \in s(\frac{i-1}{i}\alpha)$ und ein $y_i \in s(\frac{i-1}{i}\alpha)$ gibt mit $p = x_i + y_i$. Dann hat die Folge $(x_i)_i$, die in einer kompakten Menge liegt, einen Häufungspunkt x, d.h. eine Teilfolge $(x_{i(n)})_n$ konvergiert gegen x. Angenommen sei, daß $x \notin s(\alpha)$; dann gäbe es ein n und eine offene Umgebung von x, die für jedes $m \geq n$ außerhalb $s(\frac{i(m)-1}{i(m)}\alpha)$ läge; also ist $x \in s(\alpha)$. Ähnlich folgt, daß $y = p - x \in t(\alpha)$ liegt, womit $p \in s(\alpha) + t(\alpha)$ gilt. $\square$

Damit wird die folgende Definition möglich.

Definition 3.12 Es sei $u, v \in \mathcal{E}^n$ und $\lambda \geq 0$. Dann sei $u+v$ diejenige Fuzzymenge, deren Schnittfunktion $s_u + s_v$ ist; und λu sei diejenige Fuzzymenge, deren Schnittfunktion λs_u ist.

Mit Bezug auf die ursprüngliche Darstellung von Fuzzymengen als Funktionen ins reelle Einheitsintervall lassen sich Summe und Vielfaches wie folgt darstellen; dies ist die ursprüngliche Definition von Zadeh:

Theorem 3.13 *Es seien $u, v \colon \mathbb{R}^n \to [0,1]$ Standardfuzzymengen und $\lambda > 0$. Dann gilt*

$$(u+v)(p) = \sup\{u(x) \wedge v(y) \colon x + y = p\}, \tag{19}$$

$$(\lambda u)(p) = u(\tfrac{1}{\lambda}p)\} \tag{20}$$

für $p \in \mathbb{R}^n$.

Nach demselben Prinzip läßt sich auf der Menge $\mathcal{E}^n$ der Standardfuzzymengen auch eine Metrik einführen:

Definition 3.14 Es seien $u, v \in \mathcal{E}^n$. Dann sei

$$d(u,v) = \sup_{\alpha \in (0,1]} d([u]_\alpha, [v]_\alpha)$$

Daß es sich hierbei in der Tat um eine Metrik handelt, folgt daraus, daß dies für die für Paare konvexer Mengen erklärte Funktion $d(\cdot,\cdot)$ der Fall ist.

4 Funktionen zwischen Fuzzymengen

4.1 Der Mamdani-Controller

Der Bereich, in dem Fuzzymengentheorie möglicherweise am erfolgreichsten zur praktischen Anwendung gelangt, ist die Regeltechnik. An einem einfachen Beispiel illustrieren wir die Vorzüge der Fuzzyregelungstechnik gegenüber den traditionellen Methoden und beschreiben an diesem den bekanntesten auf Fuzzytechnik beruhenden Controllertypen.

Wir beschreiben das Standardbeispiel, in leicht schematisierter Form: Die Aufgabe bestehe darin, einen mit einem Endpunkt aufliegenden und in einer vertikalen Ebene gehaltenen, sonst aber frei beweglichen Stab in vertikaler Lage zu halten, indem ein geeignetes Drehmoment aufgebracht wird. Aus Position und Geschwindigkeit des Stabes ist hierzu das Drehmoment zu bestimmen.

Traditionell wird ein solches System mit den Methoden der Newtonschen Mechanik modelliert. Es sei $\vartheta(t)$ die Winkelauslenkung des Stabes aus der Vertikalen zum Zeitpunkt t; dann ergibt sich für das Drehmoment

$$u(t) = c_1\ddot{\vartheta}(t) + c_2 \sin\vartheta(t),$$

worin c_1 und c_2 gewisse Konstanten sind. Die Aufgabe besteht darin, von gewissen Anfangswerten $\vartheta(0)$ und $\dot{\vartheta}(0)$ ausgehend die Funktion u so wählen, daß $\vartheta(t)$ und $\dot{\vartheta}(t)$ für $t \to \infty$ gegen 0 konvergieren.

Diese Aufgabe ist zwar wohl vergleichsweise einfach; immerhin sind aber zur mathematischen Formulierung des Problems gewisse physikalische und sind zur erfolgreichen Lösung noch einmal recht weitgehende mathematische Kenntnisse nötig.

Dem setzt die Fuzzytechnik den Hinweis entgegen, daß, sollte eine Person die Steuerung manuell durchführen, keinerlei derartige Kenntnisse gefordert sind. Vielmehr genügen eine Reihe von Hinweisen des Typs: „Ist der Stab nach rechts geneigt und bewegt sich in ebendiese Richtung, halte dagegen; wende dabei eine Kraft auf, die zur Geschwindigkeit des Stabs in vernünftigem Verhältnis steht.“ Interessanterweise läßt sich aus Anweisungen solches Typs in einfachster Weise eine Regelung konstruieren. Mindestens im erwähnten Fall soll die Sache genauso gut funktionieren wie mit der aufwendigeren traditionellen Methode.

Das Vorgehen ist das folgende. Position, Geschwindigkeit und Drehmoment des Stabes werden nicht mehr durch die exakten reellen Werte, sondern vielmehr mithilfe umgangssprachlicher Ausdrücke spezifiziert. „stark", „leicht nach links", „rechts geneigt", „in etwa in der Mitte": Dies sind fünf grobe Positionen, und nach demselben Muster sind Geschwindigkeit und Drehmoment zu klassifizieren. Sodann sind Regeln zu formulieren, die das Drehmoment in Abhängigkeit von Position und Geschwindigkeit festlegen, und zwar in bezug auf die paar vorgebenen Ausdrücke. Eine Regel könnte lauten: „Ist der Stab leicht nach links geneigt und bewegt sich langsam nach links, ist eine leichte Kraft nach rechts auszuüben."

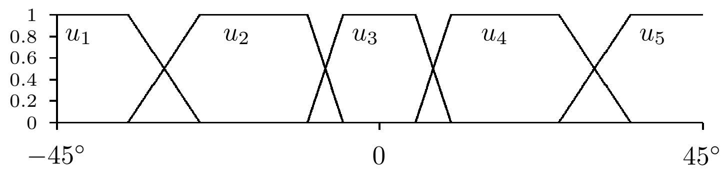
Einen Controller konstruiert man hieraus wie folgt. Wir modellieren zunächst die verwendeten Ausdrücke als Fuzzymengen. Dabei ist Folgendes zu beachten. Wir haben etwa im Fall der Position fünf Ausdrücke; beabsichtigt ist, daß mit diesen Ausdrücken eine grobe Aufteilung der Menge der möglichen Lagen des Stabes erreicht ist. Die Position ist durch das Intervall $[-45°, 45°]$ exakt beschreibbar; dieses Intervall grob in fünf Bereiche aufzuteilen bedeutet: Für jeden Ausdruck ist eine adäquate Fuzzymenge zu wählen in einer Weise, daß sich diese Mengen an jedem Punkt zu 1 aufaddieren.

Definition 4.1 Eine Menge $u_1, \ldots, u_n$ von Fuzzymengen über einer Menge M heiße *Teilung der Eins* über M, falls

$$u_1(x) + \ldots + u_n(x) = 1$$

für jeden Punkt $x \in M$ gilt.

In unserem Fall muß also eine aus fünf Fuzzymengen $u_1, ..., u_5$ bestehende Teilung der Eins über dem Intervall $[-45°, 45°]$ gewählt werden, so daß diese Fuzzymengen die sprachlichen Ausdrücke „stark nach links geneigt", ..., „stark nach rechts geneigt" gescheit modellieren, etwa:



In der gleichen Weise sind für die Geschwindigkeit Fuzzymengen $v_1, ..., v_5$ und für Drehmoment Fuzzymengen $w_1, ..., w_5$ zu wählen. Sind dann P, G,

D Symbole für Position, Geschwindigkeit, Drehmoment, können die Regeln nach folgendem Muster notiert werden:

(R_{ijk}) *Falls P gleich u_i und G gleich v_j ist, dann sei D gleich w_k.*

Wir beschreiben weiter die Wirkungsweise des Controllers. Der Controller erhält als Eingangsparameter die exakten Werte p und g für Position bzw. Geschwindigkeit und hat aufgrund der Regeln einen exakten Wert für das Drehmoment zu berechnen. Dies geschieht wie folgt in drei Schritten.

Erster Schritt. Für jede die Position des Stabes beschreibende Fuzzymenge u_i wird der Wahrheitswert $u_i(p)$ ermittelt, als derjenige Wert, der angibt, zu welchem Grad die Aussage u_i in der durch p beschriebenen Situation zutrifft. Diesen Vorgang nennt man eigenartigerweise *Fuzzifizierung* von p. Das gleiche geschieht mit g.

Man beachte, daß hierbei typischerweise einer oder zwei Fuzzymengen ein von 0 verschiedener Wert zugeordnet wird.

Zweiter Schritt. Es sei (R_{ijk}) eine wie vorstehend gegebene Regel. Den Fuzzymengen u_i und v_j, die in dieser Regel die Bedingung definieren, haben wir $u_i(p)$ bzw. $v_j(g)$ zugeordnet; und die Regel besagt, daß aus u_i und v_j die durch w_k dargestellte Aussage zu erschließen ist. Wie hierdurch nahegelegt wird, ordnen wir w_k den Wert $t_{ijk} = u_i(p) \wedge v_j(g)$ zu. Aus dem Paar w_k und t_{ijk} erzeugen wir die folgende Fuzzymenge:

$$f_{ijk} = \overline{t_{ijk}} \wedge w_k,$$

worin $\overline{t_{ijk}}$ die konstant auf t_{ijk} abbildende Fuzzymenge über dem Bezugsbereich von w_k sei. Gewissermaßen übernehmen wir also die Fuzzymenge w_k nicht voll, sondern abgeschwächt, dem Wert t_{ijk} entsprechend. Man beachte, daß, falls eine der Voraussetzung der Regel nicht erfüllt ist, d.h. entweder $u_i(p)$ oder $v_j(g)$ gleich 0 ist, sich die Nullfuzzymenge ergibt.

Sodann wird das Supremum aller aus den Regeln erhaltenen Fuzzymengen bestimmt:

$$e = \sup_{ijk} f_{ijk}$$

ist die Ergebnisfuzzymenge. Dieser zweite Schritt nennt sich *Inferenz.*

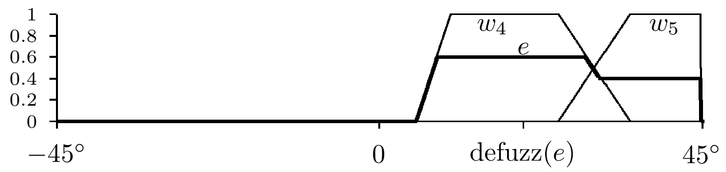
Dritter Schritt. Schließlich wird der Fuzzymenge e ein scharfer Wert defuzz(e) zugeordnet; es handelt sich um dasjenige Drehmoment, das effektiv zur Anwendung kommen soll. Nun drückt e ja gewichtete Möglichkeiten aus, die für die Wahl des Momentes in Frage kommen; insofern liegt es nahe, so etwa wie einen gewichteter Mittelwert zu bestimmen, und üblicherweise bestimmt man den Ausgangswert des Controllers wie folgt:

$$\operatorname{defuzz}(e) = \frac{\int xe(x)dx}{\int e(x)dx}, \tag{21}$$

wobei das Integral über den Wertebereich des Drehmomentes läuft, ein mit dem natürlichen Maß versehenes reelles Zahlenintervall.

Man nennt diesen Vorgang die *Defuzzifizierung*. Es gibt hierfür außer der angegebenen, der sogenannten Schwerpunktmethode, durchaus noch weitere Möglichkeiten.

Ist also beispielsweise der Stab in Position $-10°$ bei einer Geschwindigkeit von 0, liefert die Fuzzifizierung einen Grad von $0,6$ für „leicht nach links ausgelenkt“ und $0,4$ für „stark nach links ausgelenkt“, weiter von 1 für „sich nicht bewegend“. Gelten nun die Regeln, daß bei verschwindender Geschwindigkeit und leichter bzw. starker Auslenkung nach links schwach bzw. stark nach rechts zu steuern ist, wird der Fuzzymenge „leicht nach rechts drehen“ der Grad $0,6$, der Fuzzymenge „stark nach rechts drehen“ der Grad $0,4$ zugeordnet, um im Ergebnis erhält man eine Fuzzymenge folgender Form:



Das Ergebnis ist der aus e gemäß (21) bestimmte Wert.

4.2 Fuzzyrelationen

Bei einem Regler geht es um funktionale Zusammenhänge; aus gewissen Eingangs- sind Ausgangsdaten zu bestimmen. Wir gehen nun das Problem, wie Zusammenhänge zwischen Fuzzywerten darstellbar sind, in etwas allgemeinerer Form an.

Die Frage ist, wie sich Funktionen $\mathbb{R}^n \to \mathbb{R}$ geeignet verallgemeinern lassen für den Fall, daß man es mit Zusammenhängen zu tun hat, die nur für einige

Fuzzywerte beschrieben sind. Einfach sämtliche Funktionen von $\mathcal{F}(\mathbb{R})^n$ nach $\mathcal{F}(\mathbb{R})$ zu betrachten ist nicht die Lösung; deren gibt es zu viele.

Eine Funktion f von X nach Y ist andererseits als Menge der Paare $\{(x, f(x)) \colon x \in X\}$ aufzufassen, d.h. als eine spezielle Relation zwischen den Mengen X und Y. Eine Relation zwischen X und Y ist eine Teilmenge des kartesischen Produktes $X \times Y = \{(x, y) \colon x \in X, \ \ y \in Y\}$. Für den Begriff der Relation kann mit Leichtigkeit ein Fuzzyanalogon angegeben werden.

Definition 4.2 Eine *Fuzzyrelation* R zwischen Mengen X und Y ist eine Fuzzymenge über $X \times Y$.

Für das Bild eines Paares $(x, y) \in X \times Y$ unter einer Fuzzyrelation R schreiben wir im weiteren $R(x, y)$ und nicht $R((x, y))$.

Eine Fuzzyrelation R zwischen X und Y ist genauso zu interpretieren wie Fuzzymengen im allgemeinen. Modelliert wird ein Zusammenhang zwischen den beiden Mengen, und für jedes Paar (x, y) gibt $R(x, y)$ den Grad an, zu dem das gemeinsame Auftreten von x und y mit diesem Zusammenhang verträglich ist.

Im crispen Fall ist eine Relationen R_f zwischen Mengen X und Y, die von einer Funktion $f \colon X \to Y$ induziert ist, von sehr speziellem Typ: Für jedes $x \in X$ gibt es exakt ein $y \in Y$ mit $(x, y) \in R_f$. Man beachte, daß sich die Unterscheidung durch Funktionen induzierter und allgemeiner Relationen beim Übergang zu Fuzzymengen erübrigt, da Fuzzymengen Mehrdeutigkeit zulassen.

Wir überlegen nun, wie im Fall einer Fuzzyfall so etwas wie ein Funktionswert zu definieren ist. Ist eine crispe Relation R_f von einer Funktion $f \colon X \to Y$ induziert, enthält $\{y \in Y \colon (a, y)\}$ für jedes $a \in X$ genau ein Element, und zwar $f(a)$. Identifiziert man die Elemente von X und Y mit den Singletons, ist $\chi_{f(a)}(y) = \sup_{x \in X} \chi_a(x) \wedge R_f(x, y)$. Das Analogon des Funktionswertes im Fall einer Fuzzyrelation ergibt sich aus der Verallgemeinerung des letzteren Ausdrucks.

Im weiteren sei $\odot$ eine linksstetige t-Norm.

Definition 4.3 Es sei R eine Fuzzyrelation zwischen X und Y, und es sei u eine Fuzzymenge über X. Dann sei $u \circ R$ die durch

$$(u \circ R)(y) = \sup_{x \in X} (u(x) \odot R(x, y)), \quad y \in Y,$$

gegebene Fuzzymenge über Y, bezeichnet als das *Bild von R unter u*.

Gilt die durch u repräsentierte Aussage für den Wert $x \in X$ also zum Grad $u(x)$ und gilt der Zusammenhang zwischen x und einem Wert $y \in Y$ zum Grad $r = R(x, y)$, so soll also $(u \circ R)(y)$ zum Grad $\geq u(x) \odot r$ gelten; um den y zuzuordnenden Grad zu bestimmen, wird von letzterem Ausdruck das Supremum über alle $x \in X$ genommen.

Wie im Fall gewöhnlicher Funktionen ist eine Fuzzyrelation durch ihre Bilder eindeutig bestimmt.

Lemma 4.4 *Zwei Fuzzyrelationen R und S zwischen X und Y sind gleich, falls $u \circ R = u \circ S$ für alle $u \in \mathcal{F}(X)$.*

Beweis. Für $x \in X$ und $y \in Y$ gilt $R(x, y) = (\chi_x \circ R)(y)$. □

Weiter läßt sich die Hintereinanderschaltung zweier Funktionen verallgemeinern. Handelt es sich um eine crispe Relation R zwischen X und Y und eine weitere solche S zwischen Y und Z, ist $R \circ S = \{(x, z) :$ es gibt ein $y \in Y$ mit (x, y) und $(y, z)\}$.

Definition 4.5 Es sei R eine Fuzzyrelation zwischen X und Y, und S eine Fuzzyrelation zwischen Y und Z. Dann sei $R \circ S$ erklärt duch

$$(R \circ S)(x, z) = \sup_{y \in Y} (R(x, y) \odot S(y, z)), \quad x \in X, \ z \in Z.$$

Diese Definition leistet, was sie soll:

Lemma 4.6 *Es sei R eine Fuzzyrelation zwischen X und Y, und S sei eine Fuzzyrelation zwischen Y und Z. Eine Fuzzyrelation T zwischen X und Z ist gleich $R \circ S$ genau dann, wenn $u \circ T = (u \circ R) \circ S$ für alle Fuzzymengen u über X gilt.*

Beweis. Es sei $T = R \circ S$. Für $u \in \mathcal{F}(X)$ ist dann $(u \circ T)(z) = \sup_x (u(x) \odot T(x, z)) = \sup_x [u(x) \odot \sup_y (R(x, y) \odot S(y, z))] = \sup_x \sup_y [u(x) \odot R(x, y) \odot S(y, z)] = \sup_y [\sup_x (u(x) \odot R(x, y)) \odot S(y, z)] = \sup_y (u \circ R)(y) \odot S(y, z) = ((u \circ R) \circ S)(z)$ für alle z.

Umgekehrt sei $u \circ T = (u \circ R) \circ S$ für alle $u \in \mathcal{F}(X)$. Wir zeigen $(u \circ R) \circ S = u \circ (R \circ S)$, womit $T = R \circ S$ aus Lemma 4.4 folgt. Für $z \in Z$ gilt $((u \circ R) \circ S)(z) = \sup_y((u \circ R)(y) \odot S(y,z)) = \sup_y[\sup_x(u(x) \odot R(x,y)) \odot S(y,z)] = \sup_{x,y}[u(x) \odot R(x,y) \odot S(y,z)] = \sup_x(u(x) \odot \sup_y(R(x,y) \odot S(y,z)) = (u \circ (R \circ S)(z)$. □

4.3 Fuzzyrelationalgleichungen

Wir kommen zurück auf die Aufgabe, zwecks Realisierung eines Controllers eine Funktion zwischen reellen Variablen finden zu wollen, wenn nur bekannt ist, daß gewisse Fuzzymengen gewissen anderen zuzuordnen sind. Vor Augen haben wir das obige Beispiel: die Steuerungsregeln für die Balancierung eines Stabes waren von der Form „Wenn der erste Eingangswert gleich u und der zweite gleich v ist, dann soll der Ausgangswert gleich w sein" für Fuzzymenge u, v, w. Wir können die Eingangswerte eines Controller stets zu einer Variable mit Werten im $\mathbb{R}^n$, im Beispiel $\mathbb{R}^2$, zusammenfassen; dann liegen Zuordnungen

$$(u_1, w_1), \ldots, (u_n, w_n)$$

vor, worin $u_1, ..., u_n$ Fuzzymengen über dem Eingangswertebereich $X = \mathbb{R}^n$, im Beispiel $X = \mathbb{R}^2$, und $w_1, ..., w_n$ Fuzzymengen über dem Ausgangswertebereich Y, im Beispiel $Y = \mathbb{R}$, sind.

Im Blick auf den vorhergehenden Abschnitt drängt sich die Möglichkeit auf, diesen Zusammenhang durch eine Fuzzyrelation zwischen X und Y zu beschreiben. Dies führt zu folgendem Problem: Man bestimme eine Fuzzyrelation R zwischen X und Y, so daß

$$\begin{aligned} u_1 \circ R &= w_1 \\ &\vdots \\ u_n \circ R &= w_n \end{aligned} \tag{22}$$

gilt. Es handelt sich um ein System sogenannter *Fuzzyrelationalgleichungen*. Ein solches kann eine Lösung haben, muß aber nicht. Wir geben zwei Ansätze zur möglichen Lösung von (22) an.

Mittels der Relation R soll von u_i auf w_i geschlossen werden $i = 1, ..., n$; dies motiviert den folgenden Lösungsansatz. Es sei $\rightarrow$ das zu einer linksstetigen t-Norm $\odot$ gehörige Residuum. Wir definieren die Fuzzyrelation $\hat{R}$ zwischen X und Y durch

$$\hat{R}(x,y) \;=\; \bigwedge_{1\leq i\leq n} (u_i(x) \rightarrow w_i(y)). \tag{23}$$

Theorem 4.7 *Es sei $u_1,\ldots,u_n \in \mathcal{F}(X)$ und $w_1,\ldots,w_n \in \mathcal{F}(Y)$. Dann hat das System* (22) *eine Lösung genau dann, wenn die durch* (23) *definierte Fuzzyrelation $\hat{R}$ eine Lösung ist. In diesem Fall ist $\hat{R}$ die schwächste (d.h. größte) Lösung.*

Beweis. Die Fuzzyrelation R zwischen X und Y erfülle die Gleichungen (22). Zu zeigen ist: Dann gilt dies auch für $\hat{R}$.

Für jedes i gilt $u_i \circ R = w_i$, d.h. $\sup_x(u_i(x) \odot R(x,y)) = w_i(y)$ für $y \in Y$. Also ist $u_i(x) \odot R(x,y) \leq w_i(y)$ für alle $x \in X$ und $y \in Y$; aus der Residuumseigenschaft (10) folgt $R(x,y) \leq u_i(x) \rightarrow w_i(y)$; und da dies für jedes i der Fall ist, folgt $R(x,y) \leq \hat{R}(x,y)$.

Für $y \in Y$ und $1 \leq i \leq n$ folgt weiter $w_i(y) = \sup_x(u_i(x) \odot R(x,y)) \leq \sup_x[u_i(x) \odot \hat{R}(x,y)] \leq \sup_x[u_i(x) \odot (u_i(x) \rightarrow w_i(y))] \leq w_i(y)$; die letzte Ungleichung folgt aus der Definition von Operation $\rightarrow$. Also gilt $u_i \circ \hat{R} = w_i$, und die Behauptung ist bewiesen. $\square$

Im allgemeinen ist $\hat{R}$ jedoch nicht die einzige Lösung des Systems (22). Ein naheliegender weiterer Lösungsansatz ist die Fuzzyrelation R_{MA}, die sich offensichtlich an der Konstruktion des Graphen einer crispen Funktion orientiert und zu $\hat{R}$ in einem eigentümlichen Verhältnis steht:

$$R_{\mathrm{MA}}(x,y) \;=\; \bigvee_i (u_i(x) \odot w_i(y)).$$

Theorem 4.8 *Es seien $u_1,\ldots,u_n \in \mathcal{F}(X)$ normale Fuzzymengen, und es sei $w_1,\ldots,w_n \in \mathcal{F}(Y)$. Weiter gelte für alle $i,j = 1,\ldots n$*

$$\sup_{x\in X}(u_i(x) \odot u_j(x)) \;\leq\; \inf_{y\in Y}(w_i(y) \rightarrow w_j(y)).$$

Dann ist R_{MA} eine Lösung des Systems (22).

Beweis. Es sei $1 \leq i \leq n$; zu zeigen ist $u_i \circ R_{\mathrm{MA}} = w_i$, d.h.

$$\bigvee_j [\sup_x (u_i(x) \odot u_j(x)) \odot w_j(y)] = w_i(y)$$

für $y \in Y$. Da u_i normal ist, gibt es ein $x \in X$ mit $u_i(x) = 1$; es folgt $\bigvee_j [\sup_x (u_i(x) \odot u_j(x)) \odot w_j(y)] \geq \sup_x (u_i(x) \odot u_i(x)) \odot w_i(y) = w_i(y)$. Andererseits gilt laut Voraussetzung für jedes $x \in X$ und $y \in Y$, daß $u_i(x) \odot u_j(x) \leq w_j(y) \to w_i(y)$; also $\bigvee_j [\sup_x (u_i(x) \odot u_j(x)) \odot w_j(y)] \leq \bigvee_j [(w_j(y) \to w_i(y)) \odot w_j(y)] \leq w_i(y)$. □

Der springende Punkt ist, daß unabhängig davon, ob R_{MA} eine Lösung ist oder nicht und, falls ja, eine vernünftige, der Mamdani-Controller als auf R_{MA} basierend betrachtet werden kann. Es sei nämlich $e \in X$ der Eingabewert; dann ergibt sich für das Bild des Singletons von x unter R_{MA}:

$$\begin{aligned}(\chi_e \circ R_{\mathrm{MA}})(y) &= \sup_x (\chi_e(x) \odot \bigvee_i (u_i(x) \odot w_i(y))) \\ &= \bigvee_i (u_i(e) \odot w_i(y))\end{aligned}$$

für $y \in Y$; d.h. es handelt sich um das Supremum der Fuzzymengen $y \mapsto u_i(e) \odot w_i(y)$. Gehen wir vom Fall $\odot = \wedge$ aus, ist diese Fuzzymenge das Resultat der Inferenz im Mamdani-Controller; nach Defuzzifizierung ergibt sich der Ausgabewert.

4.4 Der Takagi-Sugeno-Controller

Auf einem etwas anderen und nicht mehr ganz rein fuzzytheoretischen Ansatz beruht der Takagi-Sugeno-Controller. Wiederum geht es darum, aus einem Eingangswert aus dem $\mathbb{R}^n$ einen Ausgangswert aus $\mathbb{R}$ zu bestimmen, und wiederum beruht der Controller auf einem Satz von Regeln, in deren Prämisse maximal n Fuzzymengen über dem $\mathbb{R}$ stehen. Im Unterschied zum Mamdani-Controller wird aber nicht auf Fuzzymengen über dem Ausgangswertebereich geschlossen; vielmehr wird auf eine Funktion $\mathbb{R}^n \to \mathbb{R}$ geschlossen. Zu einem gegebenen Eingangswert wird der Funktionswert gemäß jeder Regel errechnet und von den Resultaten gemäß Zutreffen der Prämissen das gewichtete Mittel gebildet.

Ein Takagi-Sugeno-Controller konstituiert sich also wie folgt. Es sei n die Zahl der reellen Eingangsparameter, und es sei für jedes $i = 1, ..., n$ $u_{i1}, \ldots, u_{ik_i}$ eine Teilung der Eins über $\mathbb{R}$. Gegeben seien weiter Funktionen $f_j \colon \mathbb{R}^n \to \mathbb{R}$, $j \in J$. Dann ist jede Regel (R_j), $j \in J$, von der Form

Ist der Eingangsparameter r_1 gleich $u_{1i_{j,1}}$ und ... und der Eingangsparameter r_n gleich $u_{ni_{j,n}}$, dann sei der Ausgangswert gleich $f_j(r_1, \ldots, r_n)$.

Aus den Eingangswerten $r_1, \ldots, r_n$ wird dann zunächst das Gewicht jeder Regel (R_j) gemäß

$$g_j = u_{ni_{j,1}}(r_1) \wedge \ldots \wedge u_{ni_{j,n}}(r_n)$$

berechnet und sodann der Ausgangswert gemäß der Formel

$$\frac{\Sigma_j g_j f_j(r_1, \ldots, r_n)}{\Sigma g_j}$$

bestimmt.

5 Die klassische Aussagenlogik

5.1 Aussagenlogik – der Ansatz

Das Thema dieses Berichtes ist es, zu erkunden, wie sich Eigenschaften eines für technische Zwecke zu beschreibenden Gegenstandes auch dann formal erfassen und für Rechnungen verwenden lassen, wenn sich ihr Zutreffen nicht in allen Einzelfällen klären läßt. Eine vage Eigenschaft zeichnet sich durch nicht beseitigbare Grenzfälle aus; sie läßt sich nicht auf die vollumfängliche Menge der möglichen Bezugsgegenstände anwenden, stattdessen nur auf eine Teilmenge, ohne daß genau klar wäre, wie sich diese vom Rest abgrenzen ließe.

Als kanonisches Modell für solche Fälle haben wir Fuzzymengen gewählt: Funktionen, die jedem Element eines Bezugsbereiches, d.h. in der Regel der Struktur, die den Gegenstand in einer präzisierten Form modelliert, einen Zutreffensgrad zuordnet, d.h. einen Wert aus dem reellen Einheitsintervall.

Die Struktur von Fuzzymengen haben wir in zweierlei Hinsicht erfaßt. Es ging zunächst um algebraisch-logische Aspekte; hier war wesentlich die Struktur betroffen, mit der die Menge der Wahrheitswerte ausgestattet ist oder ausgestattet werden kann. Zweitens ging es um analytische Aspekte; dabei war die Struktur des unterliegenden Bezugsbereichs maßgeblich, bei dem es sich in aller Regel um reelle Zahlen handelt.

Wir wollen im folgenden den ersteren Aspekt in aller Systematik weiterverfolgen: Wir betrachten Systeme vager Aussagen, modelliert durch Fuzzymengen und ausgestattet mit einer Struktur, die die wechselseitigen Zusammenhänge wiedergibt. Zur Verfügung haben wir eine Reihe von Verknüpfungen – wie Konjunktion, Disjunktion, Residuum, Negation – sowie Konstanten für „wahr“ und „falsch“. Die zu stellende Frage ist: Welche Aussagen, die mithilfe vorgegebener Verknüpfungen formulierbar sind, sind exakt die zutreffenden? Ein etwas allgemeineres Problem ist: Wenn ich von gewissen Zusammenhängen zwischen Aussagen ausgehe, d.h. gewisse mithilfe der Verknüpfungen ausdrückbare Aussagen als richtig annehme, was läßt sich erschließen?

Dabei bleibt der Bezug der modellierten Aussagen auf einen konkreten Gegenstand offen; insbesondere spielt die Struktur des den Fuzzymengen unterliegenden Bezugsbereiches fortan keine Rolle. Da zudem alle Verknüpfungen punktweise definiert sind, können wir alle Erörterungen auch als Erörterungen über die Wahrheitswerte begreifen.

Wir stellen zwei mögliche Systeme zum logischen Schließen vor. Zunächst sehen wir von der Möglichkeit, Vagheit mit einzubeziehen, ab; wir erörtern im vorliegenden Kapitel scharfe Aussagen, d.h. solche, die von crispen Teilmengen repräsentiert werden. Dies führt zur klassischen Aussagenlogik. Erst im nachfolgenden Kapitel befassen wir uns mit dem bekanntesten Beispiel einer Fuzzylogik: der łukasiewiczschen. Das ist diejenige Logik, die auf der łukasiewiczschen t-Norm und dem zugehörigen Residuum aufbaut.

5.2 Definition der klassischen Aussagenlogik

Wir gehen von der Aufgabe aus, einen Gegenstand beschreiben zu müssen; dazu wählen wir charakteristische Merkmale aus, hinsichtlich deren der Gegenstand variieren kann. Da wir uns mit rein logischen Zusammenhängen befassen wollten, werden wir nicht konkret; wir lassen die spezifischen Eigenarten der gewählten anschaulichen Aspekte außer acht. Was wir einzig festzulegen haben, ist die Art der Merkmale, mit denen wir es zu tun haben. In dieser Hinsicht treten wir im ersten Schritt hinter den zentralen Ansatz zurück, der für uns im Mittelpunkt steht: Wir ziehen Vagheit nicht in Betracht und erörtern den Fall, daß wir es mit einer Menge von Merkmalen zu tun haben, deren jedes stets eindeutig zutrifft oder nicht.

Wir geben uns eine feste Menge $\mathcal{P}$ von Aussagen vor, bezeichnet stets durch kleine griechische Buchstaben $\alpha, \beta, \varphi, \psi$ und so fort. Ein Gegenstand wird dadurch beschrieben, daß jeder dieser Aussagen ein Wahrheitswert zugeordnet wird: 1 für *zutreffend* oder 0 für *nicht zutreffend*. Es sei M die Menge all dieser Zuordnungen $z\colon \mathcal{P} \to \{0, 1\}$. Dann läßt sich jede Aussage $\alpha \in \mathcal{P}$ mit der Teilmenge $v(\alpha) = \{z \in M : z(\alpha) = 1\}$ von M identifizieren. Damit ist die Situation die folgende: Gegeben ist eine fixe Menge M, die den möglichen Variationen des betrachteten Gegenstandes entspricht, und ein System von Teilmengen von M, die die Ja-Nein-Aussagen über den Gegenstand repräsentieren.

Unser Ziel besteht nun darin, die allgemeingültigen wechselseitigen Beziehungen zwischen solcherlei Aussagen festzustellen. Hierzu führen wir Verknüpfungen zwischen diesen ein. Eine Ja-Nein-Aussage kann negiert werden; ein Paar zweier Aussagen kann und- sowie oderverknüpft werden. Mit Blick auf die Identifikation einer Aussage α mit der Teilmenge $v(\alpha)$ heißt Negation die Bildung des Komplements, Undverknüpfung die des Durchschnitts und Oderverknüpfung die der Vereinigung: $v(\neg\alpha) = \complement v(\beta) = M \setminus v(\beta)$, $v(\alpha \wedge \beta) = v(\alpha) \cap v(\beta)$ und $v(\alpha \vee \beta) = v(\alpha) \cup v(\beta)$. Hinzu kommt die Implikation $\alpha \to \beta$; diese soll die schwächste Aussage sein, die zusammen mit α β impliziert, also der größten Menge entsprechen, deren Durchschnitt mit α kleiner ist als β; es ergibt sich $v(\alpha \to \beta) = \complement v(\alpha) \cup v(\beta)$. Schließlich gibt es die zwei speziellen Aussagen 1 und 0 für „wahr" bzw. „falsch", die der ganzen Menge M bzw. der leeren Menge $\emptyset$ entsprechen.

Die genannten Operationen sind nicht alle voneinander unabhängig; wir führen zunächst nur die Verknüpfungen $\wedge$ und $\to$ sowie die Konstante 0 ein.

Definition 5.1 Die Aussagen der klassischen Aussagenlogik, kurz **KL**, sind die durch folgende Regeln zustandegekommenen Zeichenketten:

(i) $\varphi_0, \varphi_1, \ldots$ sowie 0 sind Aussagen, und zwar die sogenannt *atomaren*;

(ii) mit α und β sind auch $\alpha \wedge \beta$ und $\alpha \to \beta$ Aussagen, wobei α und β, falls nicht atomar, in Klammern zu schließen sind.

Es sei $\mathcal{P}$ die Menge aller Aussagen von **KL**.

Wir erweitern den Vorrat an Aussagen um die weggelassenen Verknüpfungen; zu beachten ist allerdings, daß es sich im folgenden formal nur um Abkürzungen handelt.

Definition 5.2 In **KL** stehe

$$\begin{aligned} \alpha \vee \beta &\text{ für } \neg(\neg\alpha \wedge \neg\beta), \\ \neg\alpha &\text{ für } \alpha \to 0, \\ 1 &\text{ für } \neg 0. \end{aligned}$$

Aussagen werden durch Teilmengen repräsentiert; wir führen die zugehörige Struktur ein.

Definition 5.3 Es sei M eine Menge. Für $A, B \subseteq M$ sei $A \cap B$ der Durchschnitt und $A \cup B$ die Vereinigung von A und B. Weiter sei

$$\begin{aligned} A \to B &= \{x \in M\colon \text{ wenn } x \in A, \text{ dann auch } x \in B\} \\ &= \{x \in M\colon x \notin A \text{ oder } x \in B\}. \end{aligned}$$

das relative Pseudokomplement von A bezüglich B. Schließlich sei $\bar{0} = \emptyset$ die leere und $\bar{1} = M$ die ganze Menge.

Weiter sei L eine Menge von Teilmengen von M, die $\emptyset$ und $\bar{1}$ enthalte und unter $\cap$, $\cup$ und $\rightarrow$ abgeschlossen sei. Dann heiße die Struktur

$$(L; \cap, \cup, \rightarrow, \bar{0}, \bar{1})$$

eine *Algebra von Teilmengen der Menge M* oder kurz eine *Teilmengenalgebra*.

Man beachte, daß eine Teilmengenalgebra $L \subseteq \mathcal{P}(M)$ durch die Operationen $\cap$ und $\rightarrow$ sowie die Konstante $\emptyset$ bereits so weit bestimmt ist, daß die übrigen Struktur definierbar wird. Das Mengenkomplement einer Menge A ist nämlich $\complement A = A \rightarrow \emptyset$, womit $A \cup B = \complement(\complement A \cap \complement B)$ sowie $\bar{1} = \complement\emptyset$ wird.

Teilmengenalgebren dienen dazu, Aussagen der klassischen Aussagenlogik zu modellieren:

Definition 5.4 Es sei $(L; \cap, \cup, \rightarrow, \bar{0}, \bar{1})$ eine Algebra von Teilmengen einer Menge M. Es sei v eine Funktion von der Menge $\mathcal{P}$ der Aussagen nach L derart, daß

$$\begin{aligned} v(\alpha \wedge \beta) &= v(\alpha) \cap v(\beta), \\ v(\alpha \rightarrow \beta) &= v(\alpha) \rightarrow v(\beta), \\ v(0) &= \bar{0} \end{aligned}$$

für alle α und β. Dann heißt v eine *Belegung* der Aussagen in L, und wir sprechen von L als einem *Modell* der Aussagen von **KL**.

Lemma 5.5 *Es sei v eine Belegung der Aussagen von* **KL** *in einer Teilmengenalgebra $(L; \cap, \cup, \rightarrow, \emptyset, \bar{1})$. Dann gilt*

$$\begin{aligned} v(\alpha \vee \beta) &= v(\alpha) \cup v(\beta), \\ v(\neg\alpha) &= \complement v(\alpha), \\ v(1) &= \bar{1}. \end{aligned}$$

Beweis. Unmittelbare Folge der Definitionen. □

Wir interessieren uns für die grundsätzlich geltenden Zusammenhänge zwischen Aussagen – dafür, welche Aussagen stärker sind als welche anderen. Es seien α und β zwei Aussagen und v eine Belegung durch Teilmengen einer Menge M; dann wird α durch $A = v(\alpha)$ und β durch $B = v(\beta)$ repräsentiert. Daß α stärker ist als β, bedeutet, daß α durch die kleinere Menge repräsentiert wird, daß also $A \subseteq B$ ist. Nun gilt, wie man sich leicht überzeugt, folgender Zusammenhang:

$$A \subseteq B \text{ gilt genau dann, wenn } A \to B = \bar{1}.$$

Hierbei ist $\bar{1}$ die ganze Grundmenge und repräsentiert die Aussage „wahr“. Um die relativen Stärken von Aussagen unter einer gegebenen Belegung festzustellen, genügt es folglich, die „wahren“ Aussagen auszusondern, d.h. die durch das Einselement $\bar{1}$ modellierten.

Uns geht es weiter um stets geltende Zusammenhänge, d.h. solche, die von einer speziellen Belegung unabhängig sind. Dies führt schließlich zur Frage, welche Aussagen unabhängig von ihrer Belegung stets wahr sind, d.h. welchen Aussagen in jedem Modell das Einselement zugeordnet wird.

Definition 5.6 Eine Aussage α von **KL** heiße *gültig unter einer Belegung* $v\colon \mathcal{P} \to L$, falls $v(\alpha) = \bar{1}$.

Eine Aussage α heiße *gültig*, in Zeichen $\models \alpha$, falls α unter jeder Belegung gültig ist.

Um die Gültigkeit einer Aussage zu prüfen, braucht man nicht sämtliche Teilmengenalgebren heranzuziehen. Die einfachste nichttriviale Teilmengenalgebra ist die zweielementige: Es sei $M = \{m\}$, d.h. einelementig; und es sei $L = \mathcal{P}(M) = \{\emptyset, M\} = \{\bar{0}, \bar{1}\}$. Dann ergibt sich für die Operationen:

A	B	$A \cap B$	$A \cup B$	$A \to B$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{1}$
$\bar{1}$	$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{0}$
$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$

Wie ersichtlich gleichen diese Verknüpfungen der Und- und Oderverknüpfung sowie der Implikation bezogen auf die beiden klassischen Wahrheitswerte. Wir definieren nun eine Modifikation der Logik **KL**, eine Aussagenlogik, die auf den Wahrheitswerten 0 und 1 basiert.

Definition 5.7 Die Aussagen der Logik $\mathbf{KL}_{\{0,1\}}$ seien die gleiche wie die von **KL**. Wir bezeichnen die Menge dieser Aussagen weiterhin mit $\mathcal{P}$.

Weiter heiße die Menge $\{0, 1\}$, ausgestattet mit den nachstehend definierten Operation $\wedge$, $\vee$ und $\rightarrow$ sowie den Konstanten 0 und 1 die *klassische Wahrheitswertealgebra*.

s	t	$s \wedge t$	$s \vee t$	$s \rightarrow t$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	1

Eine *Belegung* der Aussagen von $\mathbf{KL}_{\{0,1\}}$ sei eine Abbildung v von $\mathcal{P}$ in die klassische Wahrheitswertealgebra, so daß

$$\begin{aligned} v(\alpha \wedge \beta) &= v(\alpha) \cap v(\beta), \\ v(\alpha \rightarrow \beta) &= v(\alpha) \rightarrow v(\beta), \\ v(0) &= 0 \end{aligned}$$

für alle $\alpha, \beta \in \mathcal{P}$ gilt. Eine Aussage von $\mathbf{KL}_{\{0,1\}}$ sei *gültig*, falls $v(\alpha) = 1$ für jede Belegung v der Aussagen.

Wir können die Logiken **KL** und $\mathbf{KL}_{\{0,1\}}$ als gleich betrachten:

Theorem 5.8 *Eine Aussage φ ist gültig in* **KL** *genau dann, wenn sie gültig in* $\mathbf{KL}_{\{0,1\}}$ *ist.*

Beweis. Wir zeigen: φ ist gültig genau dann, wenn sie unter jeder Belegung in der zweielementigen Teilmengenalgebra gültig ist. Letzteres ist offensichtlich äquivalent zur Gültigkeit von φ in $\mathbf{KL}_{\{0,1\}}$.

Ist φ gültig, dann auch in der zweielementigen Teilmengenalgebra.

Es sei φ nicht gültig. Es sei $v\colon \mathcal{P} \rightarrow L$ eine Belegung in einer Algebra von Teilmengen einer Menge M, und es gelte $v(\varphi) < \bar{1}$; das bedeutet $x \notin v(\varphi)$ für ein $x \in M$. Dann ist $w\colon \mathcal{P} \rightarrow \{\emptyset, \{x\}\}$, indem wir setzen $w(\alpha) = v(\alpha) \cap \{x\}$, ebenfalls eine Belegung, und zwar in einer zweielementigen Teilmengenalgebra. Es gilt $w(\varphi) = \emptyset$, d.h. φ ist unter w nicht gültig. $\square$

Die Frage ist nun, wie sich die gültigen Aussagen systematisch bestimmen lassen.

Definition 5.9 Die *Axiome* von **KL** sind für beliebige Aussagen α, β, γ die folgenden:

(K1) $[(\alpha \to \beta) \wedge (\beta \to \gamma)] \to (\alpha \to \gamma)$,

(K2) $(\alpha \wedge \beta) \to \alpha$,

(K3) $(\alpha \wedge \beta) \to (\beta \wedge \alpha)$,

(K4) $[(\alpha \to \beta) \wedge (\alpha \to \gamma)] \to [\alpha \to (\beta \wedge \gamma)]$,

(K5a) $[(\alpha \wedge \beta) \to \gamma] \to [\alpha \to (\beta \to \gamma)]$,

(K5b) $[\alpha \to (\beta \to \gamma)] \to [(\alpha \wedge \beta) \to \gamma]$,

(K6) $0 \to \alpha$,

(K7) $[(\alpha \to \beta) \wedge ((\alpha \to 0) \to \beta)] \to \beta$.

Es sei Φ eine Menge von Aussagen; Φ heiße dann *Theorie* von **KL**. Ein *Beweis* der Aussage α aus Φ sei eine Sequenz $\gamma_1, \ldots, \gamma_n$, $n \geq 1$, dergestalt, daß den Abschluß die Aussage $\gamma_n = \alpha$ bildet und daß γ_m, $1 \leq m \leq n$, entweder aus Φ stammt oder Axiom von **KL** ist oder sich Aussagen γ_i und $\gamma_i \to \gamma_m$ unter den vorhergehenden befinden. Eine Aussage α heiße *beweisbar* aus Φ, in Zeichen $\Phi \vdash \alpha$, falls es einen Beweis von α aus Φ gibt. α heiße *beweisbar*, in Zeichen $\vdash \alpha$, falls es beweisbar aus $\emptyset$ ist.

Die in einer Theorie enthaltenen Aussagen werden im weiteren ebenfalls als deren *Axiome* bezeichnet. – Auf die Beweisregel, gemäß deren aus α zusammen mit $\alpha \to \beta$ die Aussage β erschlossen werden kann, wird allgemein als *Modus ponens*, abgekürzt (MP), Bezug genommen.

Um einen kleinen Eindruck davon zu vermitteln, in welcher Weise innerhalb **KL** Herleitungen funktionieren, beweisen wir einige Grundzusammenhänge.

Lemma 5.10 *Es seien α, β, γ Aussagen von* **KL**.

(i) $\vdash \alpha \to \alpha$.

(ii) $\vdash 1$ *und* $\vdash \alpha \to 1$.

(iii) *Aus* $\vdash \alpha \to \beta$ *und* $\vdash \beta \to \gamma$ *folgt* $\vdash \alpha \to \gamma$.

(iv) *Aus* $\vdash \beta$ *und* $\vdash \gamma$ *folgt* $\vdash \beta \wedge \gamma$.
Aus $\vdash \alpha \to \beta$ *und* $\vdash \alpha \to \gamma$ *folgt* $\vdash \alpha \to \beta \wedge \gamma$.

Beweis. Wir kehren die Reihenfolge der Beweise wegen gegenseitiger Abhängigkeiten um.

(iv) Es gelte $\vdash \beta$ und $\vdash \gamma$. Wegen Teil (i) gilt $(\beta \wedge \gamma) \to (\beta \wedge \gamma)$, und mit (K5) und (MP) leiten wir hieraus $\beta \to (\gamma \to (\beta \wedge \gamma))$ ab. Mit den Voraussetzungen und zweimaliger Anwendung von (MP) ergibt sich $\vdash \beta \wedge \gamma$.

Weiter sei $\alpha \to \beta$ und $\alpha \to \gamma$ beweisbar. Mit dem Vorstehenden, (K4) und (MP) ergibt sich $\vdash \alpha \to \beta \wedge \gamma$.

(iii) Es gelte $\vdash \alpha \to \beta$ und $\vdash \beta \to \gamma$. Damit gilt gemäß Teil (iv) $\vdash (\alpha \to \beta) \wedge (\beta \to \gamma)$. Wegen (K1) und mit (MP) folgt $\vdash \alpha \to \gamma$.

(ii) Wir zeigen $\alpha \to 1$, d.h. $\alpha \to (0 \to 0)$. Aus (K3), (K2) und Teil (iii) folgt, daß $(\alpha \wedge 0) \to 0$ beweisbar ist. Wegen (K5) ist $[(\alpha \wedge 0) \to 0] \to [\alpha \to (0 \to 0)]$ beweisbar. Somit folgt die Behauptung aus (MP).

Ist weiter α eine irgendeine beweisbare Aussage, folgt aus (MP) weiter $\vdash 1$.

(i) Wie in Teil (ii) sehen wir, daß $1 \to (\alpha \to \alpha)$ beweisbar ist. Es folgt aus 1 und (MP), daß $\vdash \alpha \to \alpha$. $\square$

Die Frage ist, ob in **KL** genau das beweisbar ist, was in **KL** gültig ist. Wir müssen zweierlei zeigen: Erstens darf nur Gültiges beweisbar sein; man spricht in diesem Fall davon, daß der Kalkül *korrekt* ist. Zweitens soll alles beweisbar sein, was gültig ist; in dem Fall heißt der Kalkül *vollständig*.

Der erste Punkt ist rasch erledigt:

Theorem 5.11 **KL** *ist korrekt: Ist* φ *in* **KL** *beweisbar, ist* φ *in* **KL** *gültig.*

Beweis. Es sei $v\colon \mathcal{P} \to L$ eine Belegung der Aussagen von **KL**. Wir müssen zeigen, daß (i) alle Axiome unter v gültig sind und (ii) falls α und $\alpha \to \beta$ unter v gültig ist, auch β unter v gültig ist.

Ad (i). Wir beschränken uns auf das Axiom (K2); zu zeigen ist $v((\alpha \wedge \beta) \to \alpha) = \bar{1}$. Es gilt gemäß Definition 5.4 $v((\alpha \wedge \beta) \to \alpha) = (v(\alpha) \cap v(\beta)) \to v(\alpha) = \mathsf{C}(v(\alpha) \cap v(\beta)) \cup v(\alpha) = \mathsf{C}v(\alpha) \cup \mathsf{C}v(\beta) \cup v(\alpha) = \bar{1}$.

Ad (ii). Es sei $v(\alpha) = v(\alpha \to \beta) = \bar{1}$. Dann ist $v(\alpha) \to v(\beta) = \bar{1}$ und folglich $v(\alpha) \subseteq v(\beta)$. Also $v(\beta) = \bar{1}$. □

Der Beweis der Vollständigkeit von **KL** ist aufwendiger und folgt im nächsten Abschnitt.

Wir schließen den Abschnitt mit einem beweistheoretischen Satz.

Theorem 5.12 *Die Logik* **KL** *erfüllt die Deduktionseigenschaft: Für jede Theorie* Φ *und Aussagen* φ *und* ψ *gilt* $\Phi \cup \{\varphi\} \vdash \psi$, *gdw* $\Phi \vdash \varphi \to \psi$.

Beweis. Es gelte $\Phi \cup \{\varphi\} \vdash \psi$, und es sei $(\gamma_1, \ldots, \gamma_n)$ der dazugehörige Beweis. Für $i = 1, \ldots, n$ sei induktiv gezeigt: $\Phi \vdash \varphi \to \gamma_i$; daraus folgt die Behauptung.

Ist γ_i ein Axiom von **KL** oder aus Φ oder gleich φ, ist dies klar. Es sei $j < i$ und γ_i das Ergebnis des auf γ_j und $\gamma_j \to \gamma_i$ angewendeten Modus ponens. Da dann nach Annahme $\Phi \vdash \varphi \to \gamma_j$ und $\Phi \vdash \varphi \to (\gamma_j \to \gamma_i)$ gilt, folgt $\Phi \vdash \varphi \to (\gamma_j \wedge (\gamma_j \to \gamma_i))$ und damit wegen $\vdash (\gamma_j \wedge (\gamma_j \to \gamma_i)) \to \gamma_i$ $\Phi \vdash \varphi \to \gamma_i$.

Gilt umgekehrt $\Phi \vdash \varphi \to \psi$, folgt mittels des Modus ponens $\Phi \cup \{\varphi\} \vdash \psi$. □

5.3 Die Vollständigkeit der klassischen Aussagenlogik

Gegeben sei eine Theorie von **KL**; wir betrachten die Menge $\mathcal{P}$ von Aussagen von **KL**, indem wir je zwei Aussagen, die von Φ als äquivalent beweisbar sind, identifizieren. Dabei bedeutet Äquivalenz, daß eine Aussage stärker ist als die andere und umgekehrt:

Definition 5.13 Es sei Φ eine Theorie von **KL**. Wir setzen $\varphi \leftrightarrow_\Phi \psi$, falls $\Phi \vdash \varphi \to \psi$ und $\Phi \vdash \psi \to \varphi$ gilt.

Wir zeigen als erstes, daß $\leftrightarrow_\Phi$ eine Äquivalenzrelation ist.

Definition 5.14 Es sei X eine Menge. Dann heißt eine zweistellige Relation $\sim$ auf X eine *Äquivalenzrelation*, falls für alle $a, b, c \in X$ gilt: (i) $a \sim a$; (ii) wenn $a \sim b$, so auch $b \sim a$; (iii) wenn $a \sim b$ und $b \sim c$, so auch $a \sim c$.

In diesem Fall heißt die Teilmenge $[a]_\sim = \{x \in X\colon x \sim a\}$, $a \in X$, die *Äquivalenzklasse* von a. Die Menge $[X]_\sim = \{[a]\colon a \in X\}$ aller Äquivalenzklassen heißt *Quotient* von X bezüglich $\sim$.

Die fundamentale Eigenschaft von Äquivalenzrelationen ist im folgenden Lemma ausgedrückt:

Lemma 5.15 *Es sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Dann sind je zwei Äquivalenzklassen entweder gleich oder leeren Durchschnitts; und die Vereinigung aller Äquivalenzklassen ist X.*

Umgekehrt sei $(X_i)_{i \in I}$ eine Partition einer Menge X, d.h. es gelte $X_i \cap X_j = \emptyset$ für $i \neq j$ sowie $\bigcup_i X_i = X$. Definieren wir $x \sim y$, falls $x, y \in X_i$ für ein i, ist $\sim$ eine Äquivalenzrelation.

In der beschriebenen Weise wird eine Eins-zu-Eins-Entsprechung zwischen den Äquivalenzrelationen auf einer Menge X und den Partionen von X hergestellt.

Lemma 5.16 *Es sei Φ eine Theorie von* **KL**. *Dann ist $\leftrightarrow_\Phi$ eine Äquivalenzrelation auf der Menge $\mathcal{P}$ aller Aussagen von* **KL**.

Beweis. Es gilt $\vdash \alpha \to \alpha$ gemäß Lemma 5.10(i). Die Relation $\leftrightarrow_\Phi$ ist nach Konstruktion symmetrisch. Weiter ist $\leftrightarrow_\Phi$ gemäß Lemma 5.10(iii) auch transitiv. $\square$

Definition 5.17 Es sei X eine Menge, $n \geq 1$ und $f\colon X^n \to X$ eine n-stellige Operation auf X; weiter sei $\sim$ eine Äquivalenzrelation auf X. Dann heißt $\sim$ mit f *kompatibel*, falls für $a_1, ..., a_n, a'_1, ..., a'_n \in A$ aus $a'_1 \sim a_1, \ldots, a'_n \sim a_n$ $f(a'_1, ..., a'_n) \sim f(a_1, ..., a_n)$ folgt.

In diesem Fall heißt

$$f\colon [X]_\sim{}^n \to [X], \quad ([a_1]_\sim, ..., [a_n]_\sim) \mapsto [f(a_1, ..., a_n)]$$

die von f auf $[X]_\sim$ *induzierte* Operation.

Man beachte, daß sich die auf einem Quotienten induzierte Operation nur deswegen definieren läßt, weil die Äquivalenzrelation mit der Operation kompatibel ist.

Lemma 5.18 *Es sei* Φ *eine Theorie von* **KL**. *Dann ist* $\leftrightarrow_\Phi$ *mit den Operationen* $\wedge$ *und* $\rightarrow$ *kompatibel. Folglich ist für Aussagen* α *und* β *definierbar*

$$\begin{aligned} [\alpha]_\Phi \wedge [\beta]_\Phi &\stackrel{\text{def}}{=} [\alpha \wedge \beta]_\Phi, \\ [\alpha]_\Phi \rightarrow [\beta]_\Phi &\stackrel{\text{def}}{=} [\alpha \rightarrow \beta]_\Phi, \\ [\alpha]_\Phi \vee [\beta]_\Phi &\stackrel{\text{def}}{=} [\alpha \vee \beta]_\Phi \\ \neg[\alpha]_\Phi &\stackrel{\text{def}}{=} [\neg\alpha]_\Phi. \end{aligned}$$

Beweis. Wenn Φ $\alpha' \leftrightarrow \alpha$ beweist, folgt $(\alpha \wedge \beta) \rightarrow \alpha'$ und $(\alpha \wedge \beta) \rightarrow \beta$, also $(\alpha \wedge \beta) \rightarrow (\alpha' \wedge \beta)$, und ähnlich ergibt sich $(\alpha' \wedge \beta) \rightarrow (\alpha \wedge \beta)$. Es folgt die Kompatibilität von $\leftrightarrow_\Phi$ mit $\wedge$. Diejenige mit $\rightarrow$ folgt mithilft der Transitivität von $\rightarrow$, d.h. (K1). □

Definition 5.19 Es sei Φ eine Theorie von **KL**. Wir setzen

$$\mathcal{L}_\Phi \stackrel{\text{def}}{=} \{[\varphi]_\Phi \colon \varphi \text{ ist Aussage von } \mathbf{KL}\}$$

sowie $\mathbf{0} \stackrel{\text{def}}{=} [0]_\Phi$ und $\mathbf{1} \stackrel{\text{def}}{=} [1]_\Phi$. Dann heiße $(\mathcal{L}_\Phi; \wedge, \rightarrow, \mathbf{0}$ die *Lindenbaumalgebra* von **KL** zur Theorie Φ.

Im weiteren werden wir die Lindenbaumalgebren von **KL** algebraisch charakterisieren. Wir ziehen zunächst nur die partielle Ordnung in Betracht.

Lemma 5.20 *Es sei* $(\mathcal{L}_\Phi; \wedge, \rightarrow, \mathbf{0})$ *die Lindenbaumalgebra einer Theorie* Φ *von* **KL**. *Dann ist* $(\mathcal{L}_\Phi; \wedge, \vee, \mathbf{0}, \mathbf{1})$ *ein* $0,1$-*Verband. Für die partielle Ordnung gilt:*

$$[\alpha]_\Phi \leq [\beta]_\Phi, \quad \textit{falls} \ \ \Phi \vdash \alpha \rightarrow \beta. \tag{24}$$

Beweis. Da $\alpha \rightarrow \alpha$ beweisbar ist, ist $\leq$ reflexiv. Da $\Phi \vdash \alpha \rightarrow \beta$ und $\Phi \vdash \beta \rightarrow \alpha$ $\Phi \vdash (\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$ und damit definitionsgemäß $\alpha \leftrightarrow_\Phi \beta$ impliziert, ist $\leq$

antisymmetrisch. Da schließlich aus $\Phi \vdash \alpha \to \beta$ und $\Phi \vdash \beta \to \gamma$ $\Phi \vdash \alpha \to \gamma$ folgt, ist $\leq$ auch transitiv und somit eine partielle Ordnung auf $\mathcal{L}_\Phi$.

Weiter gilt $\Phi \vdash \alpha \wedge \beta \to \alpha$ und $\Phi \vdash \alpha \wedge \beta \to \beta$; und es impliziert $\Phi \vdash \gamma \to \alpha$ und $\Phi \vdash \gamma \to \beta$, daß $\Phi \vdash \gamma \to \alpha \wedge \beta$. Es folgt, daß $[\alpha \wedge \beta]_\Phi$ das Infimum von $[\alpha]_\Phi$ und $[\beta]_\Phi$ ist. Ähnlich ist zu ersehen, daß $[\alpha \vee \beta]_\Phi$ deren Supremum ist.

Wegen $0 \to \alpha$ und $\alpha \to 1$ ist $\mathbf{0}$ kleinstes und $\mathbf{1}$ größtes Element von $\mathcal{L}_\Phi$. □

Definition 5.21 Eine Struktur $(L; \wedge, \vee, \to, 0, 1)$ heißt boolesche Algebra, falls das folgende gilt:

(B1) $(L; \wedge, \vee, 0, 1)$ ist ein $0, 1$-Verband.

(B2) Für $a, b \in L$ gilt:

$$a \to b \;=\; \max \{c\colon\ a \wedge c \leq b\}.$$

Mit anderen Worten: Eine boolesche Algebra ist ein residuierter Verband, in dem die monodiale Operation $\odot$ gleich dem Infimum $\wedge$ ist.

Theorem 5.22 *Es sei* $(\mathcal{L}_\Phi; \wedge, \to, \mathbf{0})$ *die Lindenbaumalgebra einer Theorie* Φ *von* **KL**. *Dann ist* $(\mathcal{L}_\Phi; \wedge, \vee, \to, \mathbf{0}, \mathbf{1})$ *eine boolesche Algebra.*

Beweis. Gegeben sei eine Theorie Φ. Gezeigt war mit Lemma 5.18, daß $\mathcal{L}$ ein $0, 1$-Verband ist. Im Hinblick auf Lemma 2.15 verbleibt zu zeigen, daß für Aussagen α, β, γ

$$[\alpha] \wedge [\beta] \leq [\gamma], \text{ gdw } [\alpha] \leq [\beta] \to [\gamma]$$

gilt. Die linke Seite heißt, daß $\Phi \vdash \alpha \wedge \beta \to \gamma$; die rechte, daß $\Phi \vdash \alpha \to (\beta \to \gamma)$ gilt. Infolge (K4) sind diese beiden Aussagen äquivalent. □

Für boolesche Algebren gilt der folgende Darstellungssatz, dessen Beweis wir auslassen.

Theorem 5.23 *Jede boolesche Algebra ist einer Teilmengenalgebra isomorph.*

Damit können wir die Vollständigkeit von **KL** zeigen.

Theorem 5.24 *Die Logik* **KL** *ist vollständig: Ist* φ *gültig, so auch beweisbar.*

Beweis. Es sei $(\mathcal{L}; \wedge, \rightarrow, \mathbf{0})$ die zur Theorie $\emptyset$ gehörende Lindenbaumalgebra von **KL**. Dann definiert

$$v\colon \mathcal{P} \rightarrow \mathcal{L}, \quad \alpha \mapsto [\alpha]$$

eine Belegung der Aussagen von **KL**.

Es sei φ eine nicht beweisbare Aussage von **KL**. Dies heißt, daß $\varphi \notin \mathbf{1}$, daß also $v(\varphi) = [\varphi] < \mathbf{1}$ gilt. Die Lindenbaumalgebra ist aber eine boolesche Algebra und folglich einer Teilmengenalgebra isomorph. Es folgt, daß φ nicht gültig ist. □

5.4 Ein Beispiel: Cadiag-1

An Beispielen einer Anwendung der klassischen Aussagenlogik gibt es sicher unzählige; wir nehmen eines aus der Medizin: Cadiag-1, das System zur Diagnoseunterstützung in einer frühen Version. In Cadiag-1 gibt es Aussagen zweierlei Sorten:

$$\begin{aligned} &\sigma_1, \sigma_2, \ldots \text{ für die Symptome,} \\ &\delta_1, \delta_2, \ldots \text{ für die Diagnosen.} \end{aligned}$$

Ein σ_i zu behaupten bedeutet, daß bei einem Patienten das i-te Symptom vorliegt; δ_j anzunehmen bedeutet, daß für den Patienten die j-te Diagnose zu stellen ist.

Aus den symptomalen Aussagen lassen sich mittels $\wedge, \vee, \neg$ zusammengesetzte Aussagen bilden. Zwischen je einer solchen, etwa σ, und je einer Diagnose, etwa δ, sind dann Zusammenhänge der folgenden Art Inhalt der Wissenbasis von Cadiag-1:

(i) σ ist obligatorisch und beweisend für δ;

(ii) σ ist fakultativ und beweisend für δ;

(iii) σ ist obligatorisch, aber nicht beweisend für δ;

(iv) σ ist fakultativ und nicht beweisend für δ;

(v) σ schließt δ aus.

σ und δ lassen sich nun in der naheliegenden Weise als Aussagen von **KL** begreifen. Die genannten Zusammenhänge können dann durch folgende Formeln wiedergegeben werden:

(i) $\sigma \rightarrow \delta$ und $\delta \rightarrow \sigma$;

(ii) $\sigma \rightarrow \delta$;

(iii) $\delta \rightarrow \sigma$;

(iv) drückt keinen logischen Zusammenhang aus;

(v) $\sigma \rightarrow \neg\delta$.

Dann entspricht dasjenige, was sich in **KL** über die Diagnosen $\delta_1, \ldots$ beweisen läßt, demjenigen, was Cadiag-1 als sicheres Ergebnis liefert.

Diese Darstellung ist natürlich vereinfacht; Cadiag-1 ist in Wahrheit komplexer. Außer der Aussagenlogik kommen weitere Methoden zur Anwendung, die beispielsweise dazu dienen, zumindest Vermutungen hinsichtlich der Diagnosen auszumachen.

6 Fuzzylogik

6.1 Definition der łukasiewiczschen Logik

Wir gehen nunmehr von der Aufgabe aus, einen Gegenstand beschreiben zu wollen, indem auch solche Eigenschaften berücksichtigt werden, deren Zutreffen oder Nichtzutreffen nicht in jedem Fall festzustellen ist.

Wir gehen in möglichst enger Analogie zum Fall der klassischen Aussagenlogik vor. Wir geben uns wiederum eine Menge $\mathcal{P}$ von Aussagen vor und beschreiben den Gegenstand dadurch, daß wir jeder der Aussagen einen Wahrheitswert zuordnen; dies ist nunmehr ein Wert im reellen Einheitsintervall $[0, 1]$. Ist M die Menge aller Variationen des Gegenstandes und α eine der Aussage, ist damit jedem $m \in M$ ein Wahrheitswert zugeordnet; α ist durch diese Fuzzymenge repräsentiert. Die Situation ist nunmehr also diejenige, daß uns ein System von Fuzzymengen über einem festen Bezugsbereich gegeben ist.

Das Ziel besteht wiederum darin, die wechselseitigen Beziehung zwischen Aussagen dieses Typs festzustellen. An Verknüpfungen verwenden wir eine Konjunktion $\odot$ und ein Residuum $\rightarrow$; hinzu kommt die Konstante 0 für die Aussage „falsch“. Wir wählen die gemäß (3) definierte łukasiewiczsche t-Norm sowie das zugehörige, gemäß (7) definierte Residuum $\rightarrow_L$. Den Index L lassen wir im weiteren weg.

Definition 6.1 Die Aussagen der łukasiewiczen Aussagenlogik, kurz **ŁL**, sind die durch folgende Regeln zustandegekommenen Zeichenketten:

(i) $\varphi_0, \varphi_1, \ldots$ sowie 0 sind Aussagen, und zwar die sogenannt *atomaren*;

(ii) mit α und β sind auch $\alpha \odot \beta$ und $\alpha \rightarrow \beta$ Aussagen, wobei α und β, falls nicht atomar, in Klammern zu schließen sind.

Es sei $\mathcal{P}$ die Menge aller Aussagen von **ŁL**.

Wir führen die folgenden Abkürzungen ein:

Definition 6.2 In **ŁL** stehe

$$\begin{aligned} \alpha \wedge \beta & \text{ für } \alpha \odot (\alpha \to \beta), \\ \neg\alpha & \text{ für } \alpha \to 0, \\ \alpha \vee \beta & \text{ für } \neg(\neg\alpha \wedge \neg\beta), \\ 1 & \text{ für } \neg 0. \end{aligned}$$

Die Aussagen von **ŁL** sollen durch Fuzzymengen modelliert werden; die benötigte Struktur ist diesmal die folgende.

Definition 6.3 Es sei L eine Menge von Fuzzymengen über dem Bezugsbereich M. L enthalte die gemäß (2.5) definierten Fuzzymengen $\bar{0}$ und $\bar{1}$; L sei abgeschlossen unter von der łukasiewiczschen t-Norm induzierten Operation $\odot$ sowie der vom łukasiewiczschem Residuum induzierten Operation $\to$; und L enthalte mit je zwei Fuzzymenge deren gemäß (1) berechnetes punktweises Infimum und Supremum. Dann heiße die Struktur

$$(L; \wedge, \vee, \odot, \to, \bar{0}, \bar{1})$$

eine *łukasiewiczsche Algebra von Fuzzymengen* oder kurz eine *Fuzzymengenalgebra.*

Eine Fuzzymengenalgebra $L \subseteq \mathcal{F}(M)$ ist durch die Operationen $\cap$ und $\to$ sowie die Konstante $\emptyset$ bereits eindeutig bestimmt. Es ist $\neg u = u \to \bar{0}$ die Standardnegation von $u \in L$; für $u, v \in L$ gilt damit

$$\begin{aligned} u \wedge v &= u \odot (u \to v), \\ u \vee v &= \neg(\neg u \wedge \neg v). \end{aligned}$$

Aussagen werden durch Fuzzymengen wie folgt repräsentiert.

Definition 6.4 Es sei $(L; \cap, \cup, \to, \bar{0}, \bar{1})$ eine Algebra von Teilmengen einer Menge M. Es sei v eine Funktion von der Menge $\mathcal{P}$ der Aussagen nach L derart, daß

$$\begin{aligned} v(\alpha \odot \beta) &= v(\alpha) \odot v(\beta), \\ v(\alpha \to \beta) &= v(\alpha) \to v(\beta), \\ v(0) &= \bar{0} \end{aligned}$$

für alle α und β. Dann heißt v eine *Belegung* der Aussagen in L, und wir sprechen von L als einem *Modell* der Aussagen von **ŁL**.

Lemma 6.5 *Es sei v eine Belegung der Aussagen von* **ŁL** *in einer Fuzzymengenalgebra* $(L; \cap, \cup, \rightarrow, \bar{0}, \bar{1})$. *Dann gilt*

$$\begin{aligned} v(\alpha \wedge \beta) &= v(\alpha) \wedge v(\beta), \\ v(\alpha \vee \beta) &= v(\alpha) \vee v(\beta), \\ v(\neg\alpha) &= \complement v(\alpha), \\ v(1) &= \bar{1}. \end{aligned}$$

Beweis. Unmittelbare Folge der Definitionen. □

Wie im Fall der klassischen Aussagenlogik interessieren wir uns für die „wahren" Aussagen unter den Elementen von $\mathcal{P}$, d.h. für diejenigen, denen unter jeglicher Belegung das Einselement $\bar{1}$ zugewiesen wird.

Definition 6.6 Eine Aussage α von **ŁL** heiße *gültig unter einer Belegung* $v \colon \mathcal{P} \rightarrow L$, falls $v(\alpha) = \bar{1}$.

Eine Aussage α heiße *gültig*, in Zeichen $\models \alpha$, falls α unter jeder Belegung gültig ist.

Um die Gültigkeit einer Aussage zu prüfen, braucht man nicht sämtliche Fuzzymengenalgebren heranzuziehen; wiederum genügt es, eine Algebra über einer einelementigen Menge zu nehmen. Wir definieren diejenige Modifikation von **ŁL**, die auf Wahrheitswerten aufbaut statt auf Fuzzymengen.

Definition 6.7 Die Aussagen der Logik $\mathbf{ŁL}_{[0,1]}$ seien die gleichen wie die von **ŁL**. Die Menge all dieser sei weiterhin $\mathcal{P}$.

Das reelle Einheitsintervall $[0, 1]$, ausgestattet mit den Verbandsoperationen $\wedge$ und $\vee$ bezüglich der natürlichen Ordnung, der łukasiewiczschen t-Norm $\odot$, dem zugehörigen Residuum und den Konstanten 0 und 1 heißt łukasiewiczsche Wahrheitswertealgebra.

Eine *Belegung* der Aussagen von $\mathbf{ŁL}_{[0,1]}$ sei eine Abbildung v von den Aussagen in die łukasiewiczsche Wahrheitswertealgebra, so daß

$$\begin{aligned} v(\alpha \odot \beta) &= v(\alpha) \odot v(\beta), \\ v(\alpha \to \beta) &= v(\alpha) \to v(\beta), \\ v(0) &= 0 \end{aligned}$$

Eine Aussage von $\mathbf{ŁL}_{[0,1]}$ sei *gültig*, falls $v(\alpha) = 1$ für jede Belegung v.

Wiederum sind **ŁL** und $\mathbf{ŁL}_{[0,1]}$ als gleich zu betrachten:

Theorem 6.8 *Eine Aussage φ ist gültig in* **ŁL** *genau dann, wenn sie gültig in* $\mathbf{ŁL}_{[0,1]}$ *ist.*

Beweis. Es sei φ gültig. Dann ist φ insbesondere unter jeder Belegung gültig, die in die Algebra aller Fuzzymengen über einem einelementigen Bezugsbereich abbildet. Also ist φ gültig in $\mathbf{ŁL}_{[0,1]}$.

Es sei φ nicht gültig. Es sei $v : \mathcal{P} \to L$ eine Belegung in einer Algebra von Fuzzymengen über einer Menge M, und es gelte $v(\varphi) < \bar{1}$; das bedeutet $v(\varphi)(x) < 1$ für ein $x \in M$. Dann ist $w \colon \mathcal{P} \to [0,1]$, indem wir setzen $w(\alpha) = v(\alpha)(x)$ eine Belegung in der łukasiewiczschen Wahrheitswertealgebra. Es gilt $w(\varphi) < 1$, d.h. φ ist in $\mathbf{ŁL}_{[0,1]}$ nicht gültig. □

Die Frage ist wiederum, wie sich die in **ŁL** gültigen Aussagen bestimmen lassen.

Definition 6.9 Die *Axiome* von **ŁL** sind für beliebige Aussagen α, β, γ die folgenden:

(L1) $[(\alpha \to \beta) \odot (\beta \to \gamma)] \to (\alpha \to \gamma)$,

(L2) $(\alpha \odot \beta) \to \alpha$,

(L3) $(\alpha \odot \beta) \to (\beta \odot \alpha)$,

(L4) $(\alpha \odot (\alpha \to \beta)) \to (\beta \odot (\beta \to \alpha))$,

(L5a) $[(\alpha \odot \beta) \to \gamma] \to [\alpha \to (\beta \to \gamma)]$, (L5b) $[\alpha \to (\beta \to \gamma)] \to [(\alpha \odot \beta) \to \gamma]$,

(L6) $0 \to \alpha$,

(L7) $((\alpha \to 0) \to 0) \to \alpha$.

Es sei Φ eine Menge von Aussagen; Φ heiße dann *Theorie* von **ŁL**. Ein *Beweis* der Aussage α aus Φ sei eine Sequenz $\gamma_1, \ldots, \gamma_n$, $n \geq 1$, dergestalt, daß den Abschluß die Aussage $\gamma_n = \alpha$ bildet und daß γ_m, $1 \leq m \leq n$, entweder aus Φ stammt oder Axiom von **ŁL** ist oder sich Aussagen γ_i und $\gamma_i \to \gamma_m$ unter den vorhergehenden befinden. Eine Aussage α heiße *beweisbar* aus Φ, in Zeichen $\Phi \vdash \alpha$, falls es einen Beweis von α aus Φ gibt. α heiße *beweisbar*, in Zeichen $\vdash \alpha$, falls es beweisbar aus $\emptyset$ ist.

Die in Lemma 5.10 für **KL** bewiesenen Grundzusammenhänge gelten auch für **ŁL**, so etwa:

Lemma 6.10 *Es seien α, β, γ Aussagen von* **ŁL**.

(i) $\vdash \alpha \to \alpha$.

(ii) $\vdash 1$ *und* $\vdash \alpha \to 1$.

(iii) *Aus* $\vdash \alpha \to \beta$ *und* $\vdash \beta \to \gamma$ *folgt* $\vdash \alpha \to \gamma$.

Beweis. Die Beweise gehen ähnlich wie im Fall von **KL**; siehe Lemma 5.10. □

Bemerkenswerter ist, was in **ŁL** im Gegensatz zu **KL** nicht mehr gilt. So ist in **ŁL** das *tertium non datur* in der Form $\alpha \vee \neg\alpha$ nicht beweisbar. Insbesondere ist eine Fallunterscheidung gemäß α und $\neg\alpha$ nicht möglich; aus $\vdash \alpha \to \beta$ und $\vdash \neg\alpha \to \beta$ läßt sich $\vdash \beta$ nicht erschließen.

Wir fragen nun, ob in **ŁL** genau das beweisbar ist, was in **ŁL** gültig ist. Zu zeigen, daß der Kalkül **ŁL** korrekt ist, bereitet wiederum keine Probleme:

Theorem 6.11 **ŁL** *ist korrekt: Ist φ in* **ŁL** *beweisbar, ist φ in* **ŁL** *gültig.*

Beweis. Das Vorgehen ist wie im Beweis von Satz 5.11. □

Andeutungen, wie der Beweis der Vollständigkeit von **ŁL** verläuft, folgen im nächsten Abschnitt.

Wir schließen mit dem Analogon von Satz 5.12.

Theorem 6.12 *Die Logik* **ŁL** *erfüllt die lokale Deduktionseigenschaft: Für jede Theorie* Φ *und Aussagen* φ *und* ψ *gilt* $\Phi \cup \{\varphi\} \vdash \psi$*, gdw es ein* $n \geq 1$ *mit* $\Phi \vdash \varphi^n \to \psi$ *gibt.*

Beweis. Der Beweis verläuft wie der von Satz 5.12. □

6.2 Die Vollständigkeit der łukasiewiczschen Logik

Wie im Fall der klassischen Aussagenlogik betrachten wir die Menge $\mathcal{P}$ von Aussagen von **ŁL**, indem wir je zwei Aussagen, die in **ŁL** als äquivalent beweisbar sind, identifizieren.

Definition 6.13 Es sei Φ eine Theorie von **ŁL**. Wir setzen $\varphi \leftrightarrow_\Phi \psi$, falls $\Phi \vdash \varphi \to \psi$ und $\Phi \vdash \psi \to \varphi$ gilt.

Lemma 6.14 *Es sei* Φ *eine Theorie von* **ŁL**. *Dann ist* $\leftrightarrow_\Phi$ *auf der Menge* $\mathcal{P}$ *aller Aussagen von* **ŁL** *eine Äquivalenzrelation, die mit den Operationen* $\odot$ *und* $\to$ *kompatibel. Folglich ist für Aussagen* α *und* β *definierbar*

$$\begin{aligned} [\alpha]_\Phi \odot [\beta]_\Phi &\stackrel{\text{def}}{=} [\alpha \odot \beta]_\Phi, \\ [\alpha]_\Phi \to [\beta]_\Phi &\stackrel{\text{def}}{=} [\alpha \to \beta]_\Phi, \end{aligned}$$

Beweis. Der Beweis geht ähnlich wie für die Lemmata 5.16 und 5.18. □

Definition 6.15 Es sei Φ eine Theorie von **ŁL**. Wir setzen

$$\mathcal{L}_\Phi \stackrel{\text{def}}{=} \{[\varphi]_\Phi \colon \varphi \text{ ist Aussage von } \mathbf{ŁL}\},$$

sowie $\mathbf{0} \stackrel{\text{def}}{=} [0]_\Phi$ und $\mathbf{1} \stackrel{\text{def}}{=} [1]_\Phi$. Dann heiße $(\mathcal{L}_\Phi; \odot, \to, \mathbf{0}$ die *Lindenbaumalgebra* von **KL** zur Theorie Φ.

Im weiteren werden wir die Lindenbaumalgebren von **ŁL** algebraisch charakterisieren. Wir ziehen zunächst nur die partielle Ordnung in Betracht.

Lemma 6.16 *Es sei* $(\mathcal{L}_\Phi; \odot, \rightarrow, \mathbf{0})$ *die Lindenbaumalgebra einer Theorie* Φ *von* **ŁL**. *Dann ist* $(\mathcal{L}_\Phi; \wedge, \vee, \mathbf{0}, \mathbf{1})$ *ein* $0,1$-*Verband. Für die partielle Ordnung gilt:*

$$[\alpha]_\Phi \leq [\beta]_\Phi, \quad \textit{falls} \;\; \Phi \vdash \alpha \rightarrow \beta. \tag{25}$$

Beweis. Ähnlich wie für Lemma 5.20. □

Der passende Typ von Algebra ist nunmehr der folgende.

Definition 6.17 Eine Struktur $(L; \wedge, \vee, \odot, \rightarrow, 0, 1)$ heißt MV-Algebra, falls L ein residuierter Verband ist, in welchem

$$\neg\neg a = a$$

für alle $a \in L$ gilt.

Mit anderen Worten: Eine MV-Algebra ist ein residuierter Verband, in dem die aus dem Residuum definierte Negation involutiv ist.

Theorem 6.18 *Es sei* $(\mathcal{L}_\Phi; \odot, \rightarrow, \mathbf{0})$ *die Lindenbaumalgebra einer Theorie* Φ *von* **ŁL**. *Dann ist* $(\mathcal{L}_\Phi; \wedge, \vee, \odot, \rightarrow, \mathbf{0}, \mathbf{1})$ *eine MV-Algebra.*

Beweis. Daß es sich um einen residuierten Verband handelt, folgt wie in 5.22.

Zu zeigen verbleibt, daß $((\alpha \rightarrow 0) \rightarrow 0) \rightarrow \alpha$ und $\alpha \rightarrow ((\alpha \rightarrow 0) \rightarrow 0)$. Ersteres folgt aus (L7); letzteres aus (L5). □

Die Vollständigkeit von **ŁL** zeigen ist aufwendig; wir geben nur einige Hinweise.

Theorem 6.19 *Die Logik* **ŁL** *ist vollständig: Ist* φ *gültig, so auch beweisbar.*

Beweis. (Hinweise.) Es sei $(\mathcal{L}; \wedge, \rightarrow, \mathbf{0})$ die zur Theorie $\emptyset$ gehörende Lindenbaumalgebra von **ŁL**. Dann definiert

$$v \colon \mathcal{P} \rightarrow \mathcal{L}, \quad \alpha \mapsto [\alpha]$$

eine Belegung der Aussagen von **ŁL**.

Es sei φ eine nicht beweisbare Aussage von **ŁL**. Dies heißt, daß $\varphi \notin \mathbf{1}$, daß also $v(\varphi) = [\varphi] < \mathbf{1}$ gilt. Die Lindenbaumalgebra ist eine MV-Algebra; man kann zeigen, daß, wenn eine Formel nicht in allen MV-Algebren den Wert 1 erhält, es eine Belegung in der łukasiewiczschen Wahrheitswertealgebra gibt, unter der sie nicht den Wert 1 erhält. Es folgt, daß φ nicht gültig ist. $\square$

6.3 Ein Beispiel: Cadiag-2

Wir besprechen ein Beispiel aus der Medizin, in welchem Fuzzylogik zur Anwendung gelangt: Cadiag-2, den Nachfolger von Cadiag-1. Es handelt sich allerdings nicht um die łukasiewiczsche Logik und auch sonst keine Standardlogik; die zu verwendende Logik kommt aber der Gödellogik recht nahe, die sich von der łukasiewiczschen in der Wahl der t-Norm unterscheidet: Es wird für die Konjunktion das Minimum verwendet.

Wir können hier nur eine sehr grobe Beschreibung sowohl des Systems Cadiag-2 als auch der Logik bieten; es geht allein darum, zugrundeliegende Ideen zu vermitteln. Wie in Cadiag-1 gibt es Aussagen zweierlei Typs:

$$\sigma_1, \sigma_2, \ldots \text{ für die Symptome,}$$
$$\delta_1, \delta_2, \ldots \text{ für die Diagnosen.}$$

Im Unterschied zu Cadiag-1 handelt es sich nunmehr aber um Aussagen, denen graduelle Wahrheitswerte zugeordnet werden. Dabei drückt ein einem Symptom zugeordneter Wahrheitswert wie gehabt ein graduelles Zutreffen von σ aus, der Vagheit von σ entsprechend. Demgegenüber drückt ein einer Diagnose δ zugeordnete Wahrheitswert eine mögliche Unsicherheit aus.

Aus den symptomalen Aussagen lassen sich wiederum mittels $\wedge$ und $\sim$ zusammengesetzte Aussagen bilden. Zwischen je einer solchen, etwa σ, und je einer Diagnose, etwa δ, sind dann Zusammenhänge der folgenden Art festgeschrieben:

(i) σ impliziert δ mit der Beweiskraft $t \in (0, 1]$;

(ii) σ schließt δ aus;

(iii) $\neg\sigma$ schließt δ aus.

Im Fall (i) ist folgendes ausgesagt: Trifft σ voll zu, gilt δ mit einer Sicherheit von t; dabei bedeutet 1 die volle Sicherheit über das Zutreffen von δ und 0 den Ausschluß von δ.

Die Aussagen über Symptome und Diagnosen lassen sich als Aussagen einer Fuzzylogik begreifen. Die Sprache dieser Logik enthält eine Konjunktion $\odot$, interpretiert durch die gödelsche t-Norm; eine Implikation $\rightarrow$; eine Negation, interpretiert durch die Standardnegation; und Konstanten $\bar{t}$, und zwar für jedes rationale $t \in [0, 1]$ eine.

Die genannten Zusammenhänge lassen sich – vereinfacht – durch folgende Formeln wiedergeben:

(i) $\bar{t} \rightarrow (\sigma \rightarrow \delta)$;

(ii) $\sigma \rightarrow \neg\delta$;

(iii) $\neg\sigma \rightarrow \neg\delta$;

Wenn nun ein von Cadiag-2 errechnetes Ergebnis von der Form „Diagnose δ liegt mit Wahrscheinlichkeit t vor“ ist, kann in der Fuzzylogik eine Aussage der Form

$$\bar{t} \rightarrow \delta$$

bewiesen werden.