Geisteswissenschaft

Bernhard Brunnsteiner

Fuzzy Logik und Wahrscheinlichkeitslehre als Theorien der Ungewissheit

Diplomarbeit

GRIN

Bibliografische Information der Deutschen Nationalbibliothek:

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

Impressum:

ISBN: 9783638004008

Dieses Buch bei GRIN:

https://www.grin.com/document/88166

Bernhard Brunnsteiner

Fuzzy Logik und Wahrscheinlichkeitslehre als Theorien der Ungewissheit

GRIN Verlag

Fuzzy Logik und Wahrscheinlichkeitslehre als Theorien der Ungewissheit

Diplomarbeit

zur Erlangung des akademischen Grades
eines Magisters der Philosophie
an der Geisteswissenschaftlichen Fakultät
der Karl-Franzens-Universität Graz

Bernhard Brunnsteiner

am Institut für Philosophie

Graz, 2007

Inhaltsverzeichnis

Vorwort

„The true logic of the world is in the calculus of probabilities."[1]

Diese Worte vom englischen Physiker James Clerk Maxwell aus dem 19. Jahrhundert entsprechen einem Weltbild, dass bereits Erfahrungen mit Ungewissheiten in der Natur gemacht hat, die sich nur mehr durch stochastische Begriffe formulieren lassen. Die Wahrscheinlichkeitstheorie wurde so seit ihrer Entstehung im 17. Jahrhundert weiterentwickelt und axiomatisiert; sie präsentiert sich heute als reichhaltige Theorie zur Modellierung und Formalisierung von Ungewissheit in verschiedenen Formen.

Im Jahre 1965 veröffentlichte der Elektrotechniker Lofti Zadeh einen Artikel namens „Fuzzy Sets"[2] indem er die fuzzy Theorie begründete und erstmals das Konzept einer unscharfen Menge vorstellte. Diese Theorie unscharfer Mengen erweist sich heute als nützliches Instrument, Vagheit als spezielle Form von Ungewissheit zu modellieren. Bei der Begründung der fuzzy Theorie durch Zadeh motivierte ihn dabei die Vorstellung diese neue Theorie für die Regelungstechnik nutzbar zu machen und Expertenwissen, das durch umgangssprachliche und oftmals vage Regeln ausgedrückt wird, für Entscheidungen in technischen Bereichen zu verwenden.[3]

Die Anwendungsgebiete der Wahrscheinlichkeitslehre und der fuzzy Theorie lassen als Theorien der Ungewissheit Überschneidungen zu, und so gerieten von Beginn an Vertreter der Wahrscheinlichkeitslehre und Anhänger der fuzzy Theorie miteinander in Konflikt und polemisierten gegeneinander. So schreibt beispielsweise der kanadische Mathematiker William Kahan fuzzy Logik sei das Kokain der Wissenschaft und sei gefährlich und falsch.[4]

Wahrscheinlichkeitslehre und fuzzy Theorie versuchen nicht nur beide Ungewissheit zu modellieren, sie weisen auch auf anderen Ebenen starke Ähnlichkeiten auf. Beides sind Theorien, deren Elemente mit Maßen im Wertebereich zwischen 0 und 1 gemessen werden: In der Wahrscheinlichkeitstheorie werden Ereignissen Grade der Wahrscheinlichkeit zwischen 0 und 1 zugeordnet. In der Theorie der unscharfen Mengen werden Elementen einer Menge Zugehörigkeitsgrade zwischen 0 und 1 zugeordnet. Weiters gibt es Versuche sowohl

[1] Hajek S 362
[2] Vgl. Zadeh 1965
[3] Vgl. Avenhaus / Seising S 270
[4] Vgl. Kosko 1995 S 13

Zugehörigkeitsgrade als auch Wahrscheinlichkeiten als Quotienten von positiven Eigenschaften durch mögliche Eigenschaften darzustellen.[5]

In diesem Sinne ist es durchaus berechtigt nach Zusammenhängen zwischen der fuzzy Theorie und der Wahrscheinlichkeitslehre zu suchen, wie es in dieser Arbeit geschehen soll: Es wird Vagheit als spezielle Form von Ungewissheit näher analysiert werden und mittels der fuzzy Theorie zu modellieren versucht. Weiters wird zu Beginn dieser Arbeit eine grundlegende Einführung in die Theorie unscharfer Mengen gegeben, die in späterer Folge durch eine weitere Interpretation unscharfer Mengen erweitert werden wird. Auf diese Interpretation aufbauend werden einige Maße und Theoreme eingeführt, mit deren Hilfe gezeigt werden wird, dass die Wahrscheinlichkeitslehre innerhalb der fuzzy Theorie deduzierbar ist.

Im letzten Kapitel wird durch verschiedene Argumentationen erläutert, warum fuzzy Theorie und Wahrscheinlichkeitslehre nicht dasselbe sein können.

In den letzten Jahren hat die fuzzy Theorie in technischen Bereichen wie der Regelungstechnik seine Nützlichkeit bewiesen. In vielen technischen Produkten kommt sie zum Einsatz und ist daher auch Gegenstand vieler Studienrichtungen technischer Universitäten. Es lohnt sich jedoch auch die philosophische Betrachtung dieser Theorie, wie es hier geschehen soll. Auf Details aus den Ingenieurswissenschaften wurde deshalb hier bewusst verzichtet und eine rein logisch-philosophische Analyse gewählt.

[5] Vgl. Wang 1993 S 8f

1 Strukturelle Einführung in die fuzzy Theorie

> *„Eine Menge ist eine Zusammenfassung bestimmter, wohlunterschiedlicher Dinge unserer Anschauung oder unseres Denkens, welche Elemente der Menge genannt werden, zu einem Ganzen*[6]

Mit diesen Worten definierte Georg Cantor 1895 den Begriff einer Menge und schuf damit die Mengenlehre, welche sich als äußerst fruchtbares Mittel zur Konsolidierung der Mathematik erwiesen hat. Doch schon bald erwies sich der obige Mengenbegriff als zu allgemein und es zeigte sich, dass dieser Widersprüche – wie die Russellsche Antinomie – innerhalb der Theorie zulässt. Als Ausweg wählte man einen axiomatischen Zugang, welche den Vorteil hat, nicht genau beschreiben zu müssen, was eine Menge ist – eine Menge ist einfach ein Objekt, das sämtliche Mengenaxiome erfüllt. Ein weiterer Ausweg könnte eine mehrwertige Logik sein, in der sich der Widerspruch nicht ergibt, wie im Kapitel 3.5.1 ersichtlich gemacht werden wird.[7]

Der Prozess der Mengenbildung als Zusammenfassung verschiedener Elemente ist jedem von der Schulzeit an geläufig. Man bildet Mengen, indem man deren Elemente aufzählt oder eine bestimmende Eigenschaft (Prädikat) anführt, die diejenigen Elemente herausfiltert, die diese Eigenschaft erfüllen. Adolf Fraenkel, Mitbegründer des gängigen Axiomensystems von Zermelo-Fraenkel, bemerkt dazu folgendes:

> *„Eine Menge M ist definiert oder „existiert", sobald von jedem beliebigen Ding feststeht, ob es Element von M ist oder nicht."*[8]

Daraus ergibt sich, dass für je zwei gegebene Objekte x und y feststeht, ob x Element von y ist oder nicht; man schreibt: $x \in y$ bzw. $x \notin y$. Eine dritte Möglichkeit wird aufgrund des Prinzips des ausgeschlossenen Dritten in der klassischen Logik ausgeschlossen. Formal lässt sich eine Menge M also schreiben als eine Zusammenfassung verschiedener Elemente, für die eine bestimmte Eigenschaft F erfüllt ist, d.h. $M = \{x \mid F(x)\}$. Durch die binäre Logik, auf die die Mengenlehre zurückgreift, gilt stets: x erfüllt F oder x erfüllt F nicht. Es lässt sich somit eine Menge M^C definieren, die all diejenigen Elemente enthält, welche die bestimmende Eigenschaft für M nicht erfüllen und das Komplement von M genannt wird; man schreibt M^C

[6] Deiser 2002 S 13
[7] Vgl. Gottwald 1989 S 291ff
[8] Fraenkel 1927 S 2

$= \{x \mid \neg F(x)\}$. Es gilt also stets x ist entweder in M oder in M^C enthalten, aber nicht in beiden zugleich, also $M \cap M^C = \{\}$ und $M \cup M^C = X$ für die Grundmenge X.[9]

In der klassischen Mengenlehre gelten also Analoga zu Prinzipien der klassischen Logik, wie dem Satz vom Widerspruch und dem Prinzip vom ausgeschlossenen Dritten, die beide auf das Zweiwertigkeitsprinzip der binären Logik beruhen, das besagt, dass jede Aussage genau einen der beiden Wahrheitswerte wahr oder falsch besitzt[10]. Das Prinzip vom ausgeschlossenen Dritten lässt sich folgendermaßen formalisieren: $A \vee \neg A$ ist eine Tautologie; hingegen bedeutet der Satz vom Widerspruch: $A \wedge \neg A$ ist eine Kontradiktion. Diese zwei fundamentalen Prinzipien gehen auf Aristoteles zurück, der in seiner Metaphysik bemerkt:

> *„Unter den Prinzipien des Beweisens verstehe ich die gemeinsamen Grundsätze, auf Grund deren man überall einen Beweis führt, z.B. den Grundsatz, daß man notwendig jegliches entweder bejahen oder verneinen muß, und daß es unmöglich ist, daß eines und dasselbe zugleich sei und nicht sei…“*[11]

Dies ist ein wesentliches Charakteristikum der klassischen Mengenlehre und Logik und es zeigt sich hier ein wesentlicher Unterschied zum Begriff der unscharfen Mengen in der Fuzzy Theorie. Während klassische Mengen ein Element entweder enthalten oder nicht, ist die Zugehörigkeit eines Elements zu einer fuzzy Menge kontinuierlich darstellbar. Über so genannte Zugehörigkeitsfunktionen, welche im folgenden Kapitel noch genauer erläutert werden, lässt sich ein Element zu einer Menge mit jedem beliebeigen Wert zwischen 0 und 1 zuordnen. Eine unscharfe Menge stellt somit eine Erweiterung des klassischen Mengenbegriffes dar, da dieser im ersteren enthalten ist, wenn man die Eigenschaft „ist Element von“ mit der Zugehörigkeit 1 und deren Gegenteil mit der Zugehörigkeit 0 versieht.

Mit der fuzzy Theorie gelingt es nun, die Abgrenzung der Elemente einer Menge von denjenigen, die keine Elemente dieser Menge sind, kontinuierlich zu ziehen. Gerade diese Möglichkeit, ein Objekt einer Menge nicht ganz oder gar nicht zuordnen zu müssen, sondern in seiner Zugehörigkeit zur Menge Zwischenwerte annehmen zu können, verleiht der Fuzzy Theorie die Möglichkeit die natürliche Welt adäquater zu beschreiben.

Da in der klassischen Logik seit Aristoteles das Prinzip vom ausgeschlossenen Dritten angenommen wird, ist diese kontinuierliche Grenzziehung eine völlig neue Sichtweise für Logiker. Sie erlaubt, Alltagsbegriffe besser modellieren zu können, da eine mathematische Definition dieser Begriffe den Alltagsvorstellungen oft widerspricht. In der klassischen Logik

[9] Vgl. Hasse 1989 S 14f
[10] Vgl. Malinkowsi 1993 S 309
[11] Aristoteles Metaphysik S 40

ist ein Mensch entweder schön oder er ist es nicht – dies scheint unserer Auffassung von Schönheit zu widersprechen. Mit Hilfe der fuzzy Logik lassen sich differenziertere Zuweisungen von Schönheit treffen, bei denen man nicht Menschen als ideal schön oder nicht einstufen muss.

Der polnische Logiker Jan Lukasiewicz, der als einer der ersten das Konzept einer zweiwertigen Logik verließ und eine mehrwertige Logik definierte, die mit dem aristotelischen Satz vom Widerspruch bricht, beschreibt den Moment der Aufgabe dieses Postulats folgendermaßen:

> *„Wenn ich mich nicht irre, so nähert sich uns der d r i t t e Moment in der Geschichte des Satzes vom Widerspruch, der alte Versäumnisse behebt. In der Entwicklung der Logik kommt dieser Zeitpunkt ebenso notwendig, wie es notwendig in der Entwicklung der Geometrie der Zeitpunkt der Revision des Parallelen-Axioms war. Aristoteles hat die A n f ä n g e der Logik geschaffen, und jeder Anfang ist unvollkommen.“*[12]

Lukasiewicz tritt für eine umfangreiche Diskussion ein, die überprüfen soll, ob der Satz vom Widerspruch oder das Prinzip vom ausgeschlossenen Dritten wirklich evident und unumgänglich sind oder nicht:

> *„Erst dann wird sich zeigen, welchen Stellenwert der Satz vom Widerspruch unter den anderen logischen Regeln einnimmt, worauf sich seine Geltung und sein Wert gründen, wie weit seine Anwendbarkeit reicht; dann wird es klar werden, ob dieser Satz wirklich der höchste von allen ist und als Grundstein für unsere gesamte Logik angesehen werden kann, oder ob man ihn auch umwandeln, beziehungsweise gar, ohne ihn zu berücksichtigen, ein System einer n i c h t a r i s t o t e l i s c h e n Logik entwickeln kann, so wie durch die Umwandlung des Parallelen-Axioms ein System der n i c h t e u k l i d i s c h e n Geometrie entstand.“*[13]

Jan Lukasiewicz hat somit zu Beginn des 20. Jahrhunderts erstmalig in der Geschichte des abendländischen Denkens den Weg der zweiwertigen Logik des Aristoteles verlassen und bereits selbst ein System mehrwertiger Logik definiert und somit einen Grundstein für die spätere Entwicklung der fuzzy Theorie gelegt.

[12] Lukasiewicz 1993 S 5
[13] Lukasiewicz 1993 S 6f

1.1 Beispiel klassischer und unscharfer Mengen

Betrachtet man die Bevölkerung eines Staates, stellt sich oft die Aufgabe diese in zwei Gruppen zu teilen, etwa in eine der Erwachsenen und eine der Kinder. Der juristische Weg zur Lösung dieser Aufgabe geschieht über die Einführung der Begriffe voll- und minderjährig. Man teilt nun die Bevölkerung in die Menge der Volljährigen, das ist also die Menge all derjeniger, die bereits das 18. Lebensjahr überschritten haben, und die Menge der Minderjährigen, also alle, die noch keine 18 Jahre alt sind. Das mag in vielen Situationen durchaus sinnvoll sein, doch ergibt sich auch die skurrile Situationen einen 17jährigen, der einen Tag vor seinem 18. Geburtstag steht, als minderjährig zählen zu müssen, während er am Tag darauf zu den Volljährigen zählt. Seine rechtliche Situation hat sich somit von einem zum anderen Tag gewandelt, während sein subjektives Empfinden während dieser zwei Tage sich wohl kaum geändert hat. Diese Mengeneinteilung basiert auf dem klassischen Mengenbegriff; von jedem Menschen steht zweifelsfrei fest, ob er zur Menge der Volljährigen oder der der Minderjährigen zählt.

Mittels der unscharfen Mengen kann man nun diesen abrupten Wechsel vermeiden, da man die Zugehörigkeit zur Menge der Erwachsenen kontinuierlich gestalten kann. Durch die Definition einer Zugehörigkeitsfunktion ist es etwa möglich, 14jährigen eine Zugehörigkeit von z.B. 0.2, 18jährigen von 0.8 und 40jährigen von 1 zuzuweisen. Dabei ergibt sich, dass hier die 18jährigen eine Zugehörigkeit von 0.8 zu den Erwachsenen, zeitgleich aber eine Zugehörigkeit von 0.2 zu den Kindern haben.

Beide Zugänge zur Einteilung haben ihre Vorzüge und beide haben ihre Nachteile. Bei dem Zugang über die unscharfen Mengen stellt sich etwa die Frage, wie man zu den genauen Werten der Zugehörigkeit gelangt. Warum wählt man 0.8 und nicht etwa 0.7998 als Zugehörigkeit des 18jährigen? Auf diese Vor- und Nachteile der Fuzzy Theorie wird nun im Laufe dieser Arbeit eingegangen werden und es wird versucht werden, zu klären, inwieweit diese der klassischen Logik über- bzw. unterlegen ist.

2 Formale Einführung in die fuzzy Theorie

Die Definition einer unscharfen Menge geht auf den Elektrotechniker Lofti Zadeh zurück, der 1965 den Aufsatz „Fuzzy Sets“ publizierte, in dem er erstmals das Konzept der fuzzy Theorie präsentierte.[14] Schon eingangs wurde erklärt:

> *„A fuzzy set is a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one.“*[15]

Im vorhergehenden Kapitel wurde angedeutet, dass in einem gewissen Sinne die unscharfe Mengenlehre als Erweiterung der klassischen angesehen werden kann; hier sei nun angemerkt, dass die Definition einer unscharfen Menge jedoch ihrerseits bereits das Konzept der klassischen Menge voraussetzt, welche in ihrer Definition eingeht.

Voraussetzung einer unscharfen Menge ist eine klassische, scharfe Grundmenge X, deren Elemente herangezogen werden. Für die Elemente dieser Grundmenge wird nun eine Zugehörigkeitsfunktion μ definiert, welche die Zugehörigkeit der einzelnen Elemente zur neuen fuzzy Menge angibt. Dann ist die Menge aller Zahlenpaare (x; μ(x))

$$\mathrm{A} := \{ (x; \mu(x) | x \in X, 0 \leq \mu(x) \leq 1 \in \Re\}$$

eine unscharfe Menge. X ist dabei die Grundmenge und μ die Zugehörigkeitsfunktion, welche reelle Werte annimmt; konventionell beschränkt man μ auf Werte zwischen 0 und 1, wobei ein höherer Wert jeweils einer höheren Zugehörigkeit entspricht.

Die Aufstellung der Zugehörigkeitsfunktion ist ein subjektiver Akt und wird in der Praxis nicht eindeutig bestimmt sein. Aufgrund der sprachlichen Unschärfe der Charakterisierung einer fuzzy Menge kann nur ein unscharfes Kriterium bestimmt werden, welches den Elementen $x \in X$ die entsprechenden μ Werte zuordnet.[16]

Darin besteht jedoch keinesfalls ein großer Nachteil gegenüber anderen mathematischen Theorien, die in der Praxis angewandt werden. Wahrscheinlichkeitsverteilungen sind ebenfalls bis zu einem gewissen Grad beliebig und subjektiv gewählt, da man die exakten Wahrscheinlichkeitswerte – falls solche überhaupt existieren – nicht berechnen kann. Auch in mathematischen Modellen mittels Differentialgleichungen werden Annahmen getroffen und die Gleichungen werden hinsichtlich einer leichten Lösbarkeit adaptiert. Die Forderung nach einem exakt definierten Verfahren zur Bestimmung der Zugehörigkeitswerte einer fuzzy

[14] Vgl. Zadeh 1965
[15] Zadeh 1965 S 338
[16] Vgl Bothe 1995 S 27f

Menge erscheint somit als übertrieben. Die Festlegung auf bestimmte Zugehörigkeitswerte muss also – wie bei anderen Theorien – nicht eindeutig sein. In der Anwendung zeigt sich darüber hinaus, dass der Einfluss der exakten Werte der Zugehörigkeitsfunktion gering ist.[17]

Innerhalb der unscharfen Mengenlehre können bekannte Begriffe der klassischen Mengenlehre eingeführt werden; so z.B. die unscharfe leere Menge oder die unscharfe Potenzmenge.[18] Ferner werden im Folgenden Operationen auf fuzzy Mengen definiert.

2.1 Operationen auf unscharfen Mengen

Für eine fuzzy Menge A mit einer Zugehörigkeitsfunktion μ definiert man das Komplement von A durch

$$A^C := \{(x; \lambda(x) \mid x \in X, \lambda(x) = 1 - \mu(x)\}$$

Man erkennt leicht, dass $(A^C)^C = A$ gilt. Anhand der Definition erkennt man ebenso, die bereits erwähnte Eigenschaft unscharfer Mengen, Elemente mit dem mengentheoretischen Komplement gemein haben zu können, also $A \cap A^C \neq \{\}$ und $A \cup A^C \neq X$ für eine unscharfe Menge $A \subseteq X$

Für $A,B \subseteq X$ mit Zugehörigkeitsfunktionen μ_A bzw. μ_B definiert man nun

$A \cap B$ mittels $\mu_{A \cap B} = \min\{\mu_A(x), \mu_B(x)\}$ (Durchschnitt) und

$A \cup B$ mittels $\mu_{A \cup B} = \max\{\mu_A(x), \mu_B(x)\}$ (Vereinigung)

Linguistische Verknüpfungen wie „nicht“, „und“ bzw. „oder“ werden in Analogie zur Booleschen Algebra durch die mengentheoretischen Operatoren Komplement, Durchschnitt bzw. Vereinigung realisiert. Ebenso übertragen sich die de Morganschen Regeln auf die unscharfen Mengen.[19]

2.2 Unscharfe Teilmengen

In der klassischen Mengenlehre ist eine Menge A in B enthalten ($A \subseteq B$), falls sämtliche Elemente von A auch Elemente von B sind. Äquivalent dazu ist die Charakterisierung von $A \subseteq B$ durch $A \in 2^B$, also dass A in B enthalten ist, falls A Element der Potenzmenge von B ist.

[17] Vgl. Avenhaus / Seising S 279
[18] Vgl ebenda S 29ff
[19] Vgl Bothe 1995 S 38ff

Natürlich kann man nun auch für unscharfe Mengen die Inklusionsrelation definieren: Seien A und B $\subseteq$ X nun fuzzy Mengen mit Zugehörigkeitsfunktionen μ_A bzw. μ_B, dann gilt:

$$A \subseteq B \text{ genau dann, falls } \mu_A(x) \leq \mu_B(x) \text{ für alle } x \in A.$$[20]

Die Inklusionsrelation $\subseteq$ ist hierbei eine binäre zweiwertige Relation: Eine fuzzy Menge ist Teilmenge einer anderen oder sie ist es nicht – ein Drittes gibt es nicht. Das Konzept der Teilmengigkeit unscharfer Mengen ist also selbst scharf und nicht fuzzy. Im Gegensatz zu diesem Inklusionsbegriffs Zadehs wird in Kapitel 4.2 die Beziehung $\subseteq$ als unscharfe Relation eingeführt werden.

2.3 Fuzzy Logik

In der bisherigen Diskussion wurde ersichtlich, dass das Konzept einer fuzzy Menge allein auf der Annahme beruht, dass die Beziehung $\in$ (Element von) nicht zweiwertig ist, sondern vielmehr jeden beliebigen Wert zwischen 0 (falsch) und 1 (wahr) annehmen kann. In der Theorie der unscharfen Mengen wird von $\in$ als zweistelligem Prädikat nicht verlangt, dass ihm nur wahr und falsch als Wahrheitswerte zugeordnet werden kann; alle anderen Aussagen und Prädikate bleiben hingegen zweiwertig.[21]

Ersetzt man nun die zweiwertige Logik durch eine unendlich-wertige Logik, ergibt sich nicht nur an das Prädikat $\in$ die Forderung der Zulassung sämtlicher Zahlen aus dem Einheitsintervall als Wahrheitswerte, sondern darüber hinaus an alle logischen Junktoren. Gibt es unendlich viele Wahrheitswerte, kann keine Tabelle mehr – wie im Falle der zweiwertigen Logik – als Definition der Semantik dienen. Zur Bestimmung der Bewertung der logischen Junktoren wird nun eine Abbildung von $[0,1]\times[0,1]\rightarrow[0,1]$ definiert, die jedem Wahrheitswertepaar, bestehend aus den Wahrheitswerten der Teilsätze des Aussagenpaares, den entsprechenden Wahrheitswert der zusammengesetzten logischen Aussage zuordnet.[22]

Wie es in der klassischen Logik einen engen Zusammenhang zur klassischen Mengenlehre gibt, so ergibt sich auch in der Theorie der unscharfen Mengen ein enger Zusammenhang zum unendlich-wertigen Logiksystem der fuzzy Logik. Dieser Zusammenhang ergibt sich vermöge der Deutung des Enthaltenseingrades des Elements $x \in A$

[20] Vgl. Zadeh 1965 S 340
[21] Vgl. Kruse / Gebhardt / Klawonn 1993 S 55
[22] Vgl. Kruse / Gebhardt / Klawonn 1993 S 55f

($\mu_A(x)$) als Wahrheitswert. Hierzu definiert man das mehrwertige zweistellige Prädikat $\in$ durch die Festlegung

$$|x \in A| := \mu_A(x),$$

wobei die Schreibweise von $|x \in A|$ durch Betragsstriche hier den Wahrheitswert bedeutet. Aus dieser Definition ergeben sich ferner folgende Zusammenhänge

$$|x \in A \cap B| = \mu_{A \cap B}(x)$$

$$|x \in A \cup B| = \mu_{A \cup B}(x)$$

$$|\neg\, x \in A| = \mu_A{}^C(x).$$ [23]

Hiermit ergibt sich nun der Zusammenhang der Theorie der fuzzy Mengen und der fuzzy Logik selbst.

[23] Vgl. Gottwald 1989 S 301f

3 Vagheit als spezielle Form der Ungewissheit

3.1 Die Paradoxie des Sorites

Eubulides von Milet soll im vierten Jahrhundert v. Chr. den so genannten Sorites eingeführt haben; das griechische Wort sorós bedeutet Haufen. Sorites steht heute noch allgemein für Paradoxien, die durch Verwendung des Kettenschlusses auftreten[24]; der Kettenschluss ist eine wiederholte Anwendung des Schlusses A→B und B→C, ergo A→C. Beim Problem des Sorites wird wesentlich davon Gebrauch gemacht, dass als gültige Antwort auf eine Entscheidungsfrage nur „Ja“ oder „Nein“ zugelassen wird. Umgangssprachliche Zwischenformen wie „Naja“, „eher ja“ oder „eher nein“ sind als Antworten nicht möglich.

Dann präsentiert sich der Sorites in folgender Form:

Frage: Bildet ein Korn einen Haufen?

Antwort: Nein.

Frage: Bilden zwei Körner einen Haufen?

Antwort: Nein.

Frage: Verwandelt also die Zufügung nur eines einzelnen Kornes etwas in einen Haufen?

Antwort: Nein.

Also bildet keine Anzahl an Körnern einen Haufen.[25]

Der Kern des Problems besteht darin, dass „Haufen“ kein klar definierter Begriff ist. Somit macht auch eine exakt bestimmte Grenze, die bestimmt, ab wie vielen Körnern etwas ein Haufen wird, keinen Sinn. Genauso wenig macht es Sinn, auf die Frage „Ab wie wenigen Haaren trägt man eine Glatze?“ eine exakte numerische Antwort zu geben.

Historisch gesehen ist das Paradoxon des Sorites eines der ersten Überlegungen zum Problem der Vagheit der Begriffe.[26] Über zweitausend Jahre später formuliert der deutsche Mathematiker und Philosoph Gottlob Frege in den Grundgesetzen der Arithmetik folgendermaßen:

> *„Eine Definition eines Begriffes (möglichen Prädikats) muss vollständig sein, sie muss für jeden Gegenstand unzweideutig bestimmen, ob er unter den Begriff falle (ob das*

[24] Vgl. Sainsbury 1993 S 39
[25] Vgl Buldt S 48ff
[26] Vgl. Buldt S 41ff

Prädikat mit Wahrheit von ihm ausgesagt werden könne) oder nicht. Es darf also keinen Gegenstand geben, für den es nach der Definition zweifelhaft bliebe, ob er unter den Begriff fiele, wenn es auch für uns Menschen bei unserem mangelhaften Wissen nicht immer möglich sein mag, die Frage zu entscheiden. Man kann das bildlich so ausdrücken: der Begriff muss scharf begrenzt sein ... Einem unscharf begrenzten Begriffe würde ein Bezirk entsprechen, der nicht überall eine scharfe Grenzlinie hätte, sondern stellenweise ganz verschwimmend in die Umgebung überginge."[27]

Klassische Gesetze der Aussagenlogik, wie das Gesetz vom ausgeschlossenen Dritten oder das Prinzip der Kontraposition, benötigen das obige Postulat, da diese wichtigen Gesetze der binären Logik sonst an Gültigkeit verlören.[28] Frege schreibt hierzu:

„Das Gesetz vom ausgeschlossenen Dritten ist ja eigentlich nur in anderer Form die Forderung, dass der Begriff scharf begrenzt sei."[29]

Ein sprachlicher Ausdruck ist also vage, falls die Grenzen seiner Anwendbarkeit nicht scharf definiert sind; falls seine Anwendbarkeit nicht in jeder Situation grundsätzlich zweifelsfrei geklärt ist, das heißt prinzipiell zweifelsfrei geklärt werden kann.[30] Ob man seine Anwendbarkeit – bei mangelhaftem menschlichen Wissen – tatsächlich klären kann, ist nicht entscheidend.

Begriffe wie „Haufen" sind also vage und nicht exakt bestimmt. Qualitative Begriffe wie „rot", „schön" oder „groß" werden ebenfalls als vage eingestuft werden, wenn man über sie nachsinnt. Doch wie ist das bei quantitativen Termini? Ist „Meter" oder „Sekunde" ein präziser Begriff? Bertrand Russell schreibt in seinem 1923 verfassten Aufsatz „Vagueness", dass auch solch quantitative Termini vage sind, da sie nur unzureichend präzise definiert werden können. Wenn der Meter als ein gewisser Abstand zwischen zwei Marken eines bestimmten Stabes in Paris mit einer gewissen Temperatur definiert wird, dann enthält auch diese Definition Vagheit, denn die Marken sind keine Punkte, und die Temperatur kann nicht exakt bestimmt werden. Deshalb sind – so folgert Russell – alle nicht logischen Begriffe vage; insbesondere sind „wahr" und „falsch" vage. Doch aus dieser Vagheit der Begriffe „Wahrheit" und „Falschheit" folgt implizit auch die Vagheit der logischen Begriffe, die über diese definiert werden.[31]

Geht man davon aus, dass alle Begriffe mehr oder weniger vage sind, stellt sich die Frage nach der Vagheit des Begriffs „Vagheit". Mit der obigen Argumentation macht es

---

[27] Frege 1966 S 69
[28] Vgl. Buldt S 66f
[29] Frege 1966 S 69
[30] Vgl. Luzzati 1999
[31] Vgl. Russell 1923

wenig Sinn, davon auszugehen, dass „vage“ exakt definierbar und die Menge aller vagen Begriffe eine klar zu bestimmende Menge sei. Dies führt zu folgender Unterscheidung: Vagheit erster Stufe befasst sich mit der Vagheit der Begriffe, wie es in dem vorangegangenen Absatz getan wurde. Die Vagheit zweiter Stufe befasst sich mit der Vagheit des Begriffs „Vagheit“ selbst. Diese Vagheit zweiter Stufe ist umstritten; die eine Seite hält sie für unumgänglich und sogar als konstitutives Element echter Vagheit, die andere Seite meint, sie umgehen zu können.

Kehrt man zurück zur Paradoxie des Sorites, lässt sich die Frage nach der Existenz der Vagheit zweiter Stufe veranschaulichen:

Man ist sich gemeinhin einig, dass ein Korn alleine kein Haufen ist. Ebenso wird man sich einig sein, dass 100.000 Körner ein Haufen sind. Von nun an stehe das Prädikat $H(n)$, zur kompakteren Schreibweise, für die Eigenschaft „n Körner sind ein Haufen“. Es gilt also $H(100.000)$ und $\neg H(1)$. Da der Begriff „Haufen“ vage ist, lässt sich keine scharfe Grenze zwischen Haufen-Sein und Nicht-Haufen-Sein ziehen, das heißt, es gibt keine natürliche Zahl ω mit $1<\omega<100.000$, sodass $H(n)$ für $n>\omega$ und $\neg H(n)$ für $n<\omega$.

Aufgrund der Vagheit von „Haufen“ gibt es also stets einen Bereich an $\xi\in N$, für die nicht exakt bestimmbar ist, ob $H(\xi)$ oder $\neg H(\xi)$ gilt. Verneint man nun die Vagheit zweiter Stufe, geht man von der Existenz zweier natürlicher Zahlen ω_1, ω_2 aus, die diesen Bereich der Unsicherheit exakt eingrenzen.

Es gilt also für $\xi<\omega_1$:	$\neg H(\xi)$
$\omega_1<\xi<\omega_2$:	Es ist unsicher, ob ξ Körner ein Haufen sind.
$\omega_2<\xi$:	$H(\xi)$

Dieser abrupte Wechsel zwischen sicherem Haufen-Sein und unsicherem Haufen-Sein ist allerdings genauso wenig zufriedenstellend wie ein abrupter Wechsel zwischen Haufen-Sein und Nicht-Haufen-Sein. Eine Vagheit zweiter Stufe scheint also, wenn nicht evident, so doch wenigstens naheliegend zu sein. [32]

[32] Vgl. Buldt S 47f

3.2 Drei Arten der Vagheit

Kehren wir nun zur Vagheit erster Stufe zurück. Es gibt drei verschiedene Möglichkeiten, Vagheit aufzufassen: erstens als eine ontologische Vagheit. Deutet man auf einen Menschen und behauptet man, das sei Herr Q, dann stellt sich die Frage, ob das auch Herr Q ist, wenn ihm ein paar Haare ausgegangen sind, oder sich sonst eine kleine Veränderung an ihm ereignet hätte. Ab wann ist dann Herr Q nicht mehr Herr Q? Dies ist eine wichtige Frage und motiviert eine weitere interessante Fragestellung: Gibt es vage Gegenstände? Wenn man auf einen Berg in der Umgebung deutet und fragt, wo denn der Berg beginnt und die Ebene aufhört, erkennt man, dass Berge als Teil der Wirklichkeit vage Begrenzungen haben. Daraus muss jedoch nicht folgen, dass der Berg an sich tatsächlich vage ist – es könnte auch nur das Wort „Berg“ vage sein.[33]

Als zweite Art der Vagheit ergibt sich die sprachliche: „Herr Q ist dick“ ist eine Aussage, die sich nicht durch empirische Daten messen lassen wird; Dicksein ist an sich nicht exakt bestimmbar und befindet sich in einem prädikativen Halbschatten, kann also weder durch empirische noch durch begriffliche Untersuchungen bestimmt werden (Russell sprach anstelle von prädikativen Halbschatten von einer Penumbra).

Eine dritte Art der Vagheit ist von sprachpragmatischer Natur: Nicht die Sprache ist vage, sondern ihr Gebrauch. Die Sprache als Abbild der Welt ist nach dieser Auffassung in der Lage die Welt so zu beschreiben, wie sie ist; geht man also von keiner ontologischen Vagheit aus, ergibt sich auch keine Vagheit der Sprache. Was vage ist, ist der Gebrauch der Sprache, da der Mensch aus Zeit- und Energiegründen die sprachlichen Möglichkeiten nicht voll ausschöpfen kann und die Sprache nur unvollständig anwendet. So wäre z.B. eine Unterscheidung des Terminus „Haufen“ in beliebig viele unterschiedliche Termini für Ein-Korn-Haufen, 2-Körner-Haufen,…, 1-Million-Körner-Haufen, etc. theoretisch denkbar, um die Problematik im Paradoxon des Sorites zu vermeiden. Denkt man an das bekannte Gerücht, dass Eskimos angeblich über zwanzig verschiedene Wörter für unser deutsches „weiß“ kennen[34], erkennt man den Kern des Arguments. Die Sprache ist in der Lage Exaktheit beliebig genau zu approximieren; mit der Anwendung dergleichen wird jedoch Vagheit eingeführt, die nicht von der Sprache herrührt, sondern vom Anwender selbst, der die Sprache vage verwendet.

[33] Vgl. Sainsbury 1993 S 69ff

[34] Vgl. Radtke

3.3 Andere Arten der Ungewissheit

Ungewissheit ist laut Duden[35] ein Zustand, in dem etwas nicht feststeht; das heißt, in dem nicht entscheidbar ist, ob etwas gilt oder nicht, ob etwas wahr ist oder nicht. Vages Wissen – oder unscharfes Wissen – ist – wie wir gesehen haben – ungewiss im Sinne der Definition. Das Phänomen der sprachlichen Vagheit oder Unschärfe lässt sich allerdings von anderen Arten der Ungewissheit unterscheiden: von Mehrdeutigkeit (Ambiguität), von Allgemeinheit, von Unspezifität (Ungenauigkeit) und auch von Relativität.

Zur Erklärung sollen hier Beispiele dienen: Der Satz „Ich gehe zur Bank." ist nicht vage, sondern mehrdeutig, falls „Bank" nicht näher bestimmt ist und nicht klar ist, ob etwa eine Sitzgelegenheit oder ein Geldinstitut gemeint ist. Diese Unbestimmtheit lässt sich aber durch eine genauere Bestimmung leicht aufklären. Mehrdeutigkeit ist also nur in dem Sinne ungewiss, da im Moment nicht klar entschieden werden kann, ob die Behauptung wahr oder falsch ist. Mit einer genaueren Beschreibung lässt sich diese Ungewissheit schnell beheben. Allgemeinheit ruft im folgenden Satz Ungewissheit hervor: „Dieser Mensch isst gerne Spaghetti" – hier kommt „Mensch" als Allgemeinbegriff vor und somit ist z.B. sein Geschlecht – männlich oder weiblich – nicht näher bestimmt; deshalb ist „Mensch" aber nicht vage, sondern eben nur allgemein, was ebenfalls zu einer momentanen Ungewissheit führt, die schnell und einfach zu beheben ist, etwa durch Ergänzungen und Spezifikationen.

Ungenauigkeit ist eine weitere Art von Ungewissheit, die deutlich von Vagheit differenziert werden muss. Der Satz „Karl wiegt zwischen fünf und fünfhundert Kilogramm." ist unbestimmt, nicht weil der Ausdruck vage, sondern weil seine Grenzen zu unspezifisch sind; sie sind nicht verschwommen, sondern einfach zu wenig präzise.[36] Diese Unspezifität, Impräzision oder Ungenauigkeit lässt sich durch die in der Praxis oftmals herrschende Situation der Unmöglichkeit der Feststellung eines beliebigen Grades an Genauigkeit erklären. Wenn ein Instrument nicht klein genug skaliert ist, gibt es Grenzen der Messbarkeit, die Impräzision ergeben. Im Falle der Darstellung irrationaler Zahlen ist Impräzision per se unvermeidbar und tritt somit im täglichen Leben wie auch im wissenschaftlichen Alltag unvermeidbar auf. Ungenauigkeit ist folglich nicht vermeidbar, weshalb sich die Frage nach der bestmöglichen Behandlung dergleichen stellt. In der Praxis beschreibt man derartige

[35] Duden Stichwort „Ungewissheit"
(http://www.duden.de/suche/index.php?begriff=Ungewissheit&bereich=mixed&pneu=#inhalte)
[36] Vgl. Buldt S 41

impräzise Informationen durch nichtstochastische Fehlerintervalle, die man auch als Spezialfall von unscharfen Mengen auffassen kann.[37]

Das Phänomen der Vagheit ist grundlegend vom Phänomen der Relativität verschieden, auch wenn diese beiden Phänomene oftmals gemeinsam auftreten. Im Satz „Der Mann ist groß." ist das Adjektiv „groß" sowohl vage als auch relativ. Im Unterschied zur Vagheit lässt sich die Relativität durch Zusatzinformation beheben. Der Ausdruck „Der Mann ist groß." ist kontext- und kulturbezogen: Fällt der Satz innerhalb eines Pygmäenstammes, hat er eine andere Bedeutung als beispielsweise in Mitteleuropa. Der Satz mag auch unterschiedliche Bedeutungen in Schweden und Italien haben, wo die Menschen gemeinhin als größer bzw. kleiner gelten. Unter Pygmäen wird eine Körpergröße als groß gelten, die mit Sicherheit in Europa als nicht groß gilt. Diese Relativität im Satz „Der Mann ist groß" lässt sich aber durch Zusätze wie zum Beispiel „Dieser Mann ist groß für einen Österreicher" beheben. Die Vagheit des Satzes lässt sich nicht beheben. Oftmals tritt Relativität auch ohne Vagheit auf: Das Prädikat „ist überdurchschnittlich groß" ist ein relatives, aber kein vages. Überdurchschnittlich groß ist jemand, der größer ist als der Durchschnitt – und die statistische Durchschnittsgröße ist ein Wert, der sich exakt bestimmen lässt. Dieses Prädikat ist also nicht vage, sondern nur relativ, da erst der Kontext bzw. die Menge, deren Durchschnittsgröße gemeint ist, spezifiziert werden muss.[38]

Eine weitere Art der Ungewissheit ist mit dem Begriff des Zufalls verbunden, der uns durch Würfelexperimente, Lottoziehungen und dergleichen vertraut ist. Diese spezielle mit Zufallsmechanismen verbundene Ungewissheit wird in einem späteren Kapitel noch eingehend diskutiert und deren Unterschied zur Vagheit herausgearbeitet werden.

3.4 Unterschiede zwischen Vagheit und auf Zufall basierender Ungewissheit

Das Phänomen der Ungewissheit, wie es in natürlichen Sprachen auftritt, wurde als solches in der Vergangenheit zu ungenau unterschieden. Vagheit als subjektive Ungewissheit, die nicht objektivierbar ist, wurde nicht als spezielle Form betrachtet und wird erst durch die Entwicklung der Fuzzy Theorie adäquat formalisiert.[39] Diese subjektive Ungewissheit, die Gegenstand der Fuzzy Theorie ist, ist eine Ungewissheit prinzipieller Natur, die im Gegensatz

[37] Vgl. Kruse/Gebhardt/Klawonn 1993 S 3
[38] Vgl. Sainsbury 1993 S 41f
[39] Vgl. Mukaidono 2004 S 95

zu ungenauem oder mehrdeutigen Wissen nicht behebbar ist. Weder Spezifizierung noch Beobachtungen oder theoretische Überlegungen reduzieren Vagheit. Die Ungewissheit, die Gegenstand der Wahrscheinlichkeitstheorie ist, rührt von einer Unkenntnis der Zukunft – die wird durch Kenntnis der Zukunft aufgehoben. Ist es heute ungewiss, ob es morgen schneit, ist dies morgen nicht weiter unklar, sondern gewiss. Die Frage, welche Augenzahl nach dem Wurf eines Würfels aufscheinen wird, ist lediglich vor dem Wurf interessant; vor dem Wurf beträgt die Wahrscheinlichkeit für eine bestimmte Augenzahl zwischen eins und sechs, bei einem fairen Würfel genau ein Sechstel. Nach dem Wurf ist der Sachverhalt geklärt und die Ungewissheit verschwunden; entweder es liegt die bestimmte Augenzahl obenauf oder nicht. Bei Vagheit ist eine breitere Kenntnis der Zukunft irrelevant. Bei dem Problem, ab wann eine Ansammlung Körner ein Haufen ist, ist es unwichtig, ob man die Körner tatsächlich angehäuft hat oder nicht. Die Frage, ab wann Körner einen Haufen bilden, ist per se interessant und zeitlos.

Ein Beispiel möge den Unterschied zwischen Ungewissheit der Fuzzy Theorie und der Wahrscheinlichkeitsrechnung veranschaulichen: Wenn ich mich frage, ob mein Nachbar – den ich nicht kenne – größer als 1.80 Meter groß ist, lässt sich dies mittels der Wahrscheinlichkeitsrechnung gut behandeln. Man nimmt Bevölkerungsstatistiken zu Rate und errechnet, wie groß die Wahrscheinlichkeit ist, dass ein Grazer Mann größer als 1.80 ist. Die Ungewissheit in der Behauptung „Mein Nachbar ist größer als 1.80.“ wird durch die Angabe einer Wahrscheinlichkeit modelliert. Lerne ich meinen Nachbarn kennen, kann ich ihn messen, erfahre ich seine tatsächliche Größe und kann die gestellte Frage nun eindeutig klären. Die Wahrscheinlichkeit, mit der mein Nachbar größer als 1.80 ist, spielt nunmehr keine Rolle mehr – entweder mein Nachbar ist nun größer oder nicht.

Von ganz anderer Art Ungewissheit ist die Behauptung „Mein Nachbar ist groß.“ Hier ist nicht die Frage ungeklärt, ob dies der Fall ist oder nicht, sondern vielmehr die Bedeutung des Wortes „groß“ an sich ist ungeklärt. Hier hilft es nicht, dass ich den Nachbarn kennen lernen und ihn messen kann; „groß“ ist ein vages Adjektiv und die Frage, wer groß ist oder nicht, ist subjektiv.

Wenn wir nun die Wahrscheinlichkeit eines Ereignisses angeben, tun wir dies, da wir nicht wissen, ob dieses Ereignis nun tatsächlich eintritt oder nicht. Diese Unkenntnis der Zukunft, dieses Unwissen, ob das Ereignis nun eintritt oder nicht, ist reduzierbar, falls man den Wissensstand erhöht; man kann dazu Statistiken durchführen oder Versuche machen.

Ebenfalls können theoretische Überlegungen die Ungewissheit mindern und die Wahrscheinlichkeitsprognosen genauer machen, z.B. mittels Theoreme der Wahrscheinlichkeitslehre.

Dies alles führt im Falle der Vagheit zu keiner Verminderung der Ungewissheit; weder können theoretische Überlegungen noch Versuche und Beobachtungen vage Begriffe spezifizieren.

3.5 Versuche Vagheit zu formalisieren

> *"All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life, but only to an imagined celestial existence."*[40]

Mit diesen Worten kritisiert Bertrand Russell die klassische Logik und deren Annahme präziser Begriffe. Es stellt sich nun die Frage, wie man vage Begriffe besser beschreiben kann.

Aufgrund der hohen Akzeptanz, Bewährtheit und Klarheit der klassischen binären Logik soll bei dem Versuch, eine Logik zu finden, die das Phänomen der Vagheit besser modellieren kann, die klassische Logik nicht aufgegeben werden, das heißt sie soll mitunter als Spezialfall einer neuen Logik ihre Gültigkeit bewahren.

Im Folgenden sollen Logiken präsentiert und diskutiert werden, die versuchen, das Konzept der Vagheit zu modellieren.

3.5.1 3-wertige Logik

In der binären Logik ist jede Behauptung entweder wahr oder falsch; ein drittes gibt es nicht (tertium non datur). Dies nennt man das Zweiwertigkeitsprinzip der klassischen binären Logik[41]. Dies impliziert jedoch nicht, dass von jeder Behauptung auch eindeutig festgestellt werden kann, ob sie wahr oder falsch ist – die Behauptung ist wahr oder falsch, nur kann dies eventuell nicht festgestellt werden. Es gibt Sätze empirischen Inhalts, von denen nicht festgestellt werden kann, welchen Wahrheitswert sie besitzen. Beispiel hierfür sei die Behauptung „Aristoteles hatte die Blutgruppe A positiv." Ferner können Sätze metaphysischen Inhalts wie zum Beispiel über die Existenz Gottes weder verifiziert noch falsifiziert werden; das Zweiwertigkeitsprinzip impliziert jedoch, dass auch diese Sätze wahr oder falsch sind, unabhängig von der Möglichkeit der Feststellung des Wahrheitswertes.

[40] Russell 1923
[41] Vgl. Gottwald 1989 S 1

Nun ist es der prädikative Halbschatten, der ein vages Prädikat nicht zweifelsfrei als wahr oder falsch erkennbar macht. Als Folge könnte man in naheliegender Weise die zweiwertige Logik um einen weiteren Wahrheitsgrad – der zwischen wahr und falsch liegt – erweitern. So wird in der 3-wertigen Logik des Jan Lukasiewicz, die als Beginn mehrwertiger Logiksysteme angesehen werden kann, neben den Wahrheitswerten wahr (1) und falsch (0) unbestimmt (0.5) als dritter Wahrheitswert eingefügt, um so den prädikativen Halbschatten vager Prädikate abzudecken. Doch reicht die Anwendbarkeit des dritten Wahrheitswertes noch weiter. Im klassischen Beispiel des unbestimmten Wahrheitswertes des Lukasiewicz „Ich werde Weihachten nächsten Jahres in Warschau sein“ ist kein vager Begriff, kein vages Prädikat vorhanden – dennoch ist die Behauptung ungewiss und eine Zuordnung der Wahrheitswerte wahr und falsch erscheint zum jetzigen Zeitpunkt als inadäquat. Die 3-wertige Logik erscheint also prima vista in der Lage, nicht nur Vagheit sondern, noch allgemeiner Ungewissheit formalisieren zu können.

Weiters ist die 3-wertige-Logik in der Lage die Russellsche Antinomie, der Menge aller Mengen, die sich nicht selbst enthalten, zu formulieren; jedoch ist die Antinomie in der 3-wertigen Logik nicht länger kontradiktorisch. Dies ist leicht einzusehen, wie folgende Argumentation zeigt. Wenn man die Menge aller Mengen, die sich selbst nicht enthalten, formalisiert, erhält man $R:=\{x \mid x \notin x\}$. Die Antinomie besteht nun in der Bedingung $R \in R \Leftrightarrow R \notin R$. Diese ist in der 3-wertigen-Logik jedoch erfüllt, wenn man als Wahrheitswert „unbestimmt“ wählt.[42]

Es zeigt sich allerdings, dass ein 3-wertiges wie auch n-wertiges Logiksystem ungeeignet ist, die bekannten Antinomien der naiven Mengenlehre zu umgehen. Konventioneller Weise wird die Inkonsistenz der Mengenlehre durch Einschränkungen – wie etwa im Axiomensystem von Zermelo-Fraenkel – behoben; eine grundsätzlich andere Möglichkeit wäre die Inkonsistenz nicht mehr in der klassischen Prädikatenlogik zu betrachten, sondern stattdessen in mehrwertigen Logiken. Für diesen Versuch wurde jedoch bereits gezeigt, dass zumindest endlich-wertige Logiken dafür nicht geeignet sind, da in ihnen Analoga zur Russellschen Antinomie ableitbar sind. [43]

Neben dem Zweiwertigkeitsprinzip basiert die binäre Logik noch auf einer zweiten Annahme, die – im Gegensatz zum Zweiwertigkeitsprinzip – mehrwertige Logiken ebenfalls treffen, nämlich das Extensionalitätsprinzip. Dieses Prinzip besagt in seiner aussagenlogischen Fassung, dass der Wahrheitswert einer zusammengesetzten Aussage nur

[42] Vgl. Malinkowski S 309ff
[43] Vgl. Gottwald 1989 S 291ff

von den Wahrheitswerten der atomaren Bestandteile der Aussage abhängt. Auf dieses Prinzip verzichtet auch die mehrwertige Logik nicht, insbesondere nimmt es auch die beschriebene 3-wertige Logik des Lukasiewicz an.[44]

Die Entstehung der mehrwertigen Logikkonzeption, die mit den Arbeiten von Lukasiewicz und der polnischen Logikerschule begann[45], ist eng mit der Entstehung des so genannten Intuitionismus verbunden. Zwar bedeutet die Kritik am Prinzip vom ausgeschlossenen Dritten und äquivalenter Formulierungen noch keineswegs den Übergang zu mehrwertigen Logiksystemen, doch erschüttert diese Kritik das Prinzip der Zweiwertigkeit von logischen Aussagen, welches das Fundament der binären Logik bildet.[46] Durch die Kritik am Zweiwertigkeitsprinzip sind mehrwertige Logiksysteme und der Intuitionismus also verbunden; doch reicht diese Verbundenheit nicht weiter, wie schon das Prinzip der doppelten Negation zeigt, welches in mehrwertigen Systemen gültig ist – wie noch zu sehen sein wird – und das die intuitionistische Schule verneint.

Die 3-wertige-Logik bietet also allerhand Möglichkeiten, die über die der binären Logik hinausreichen; nun soll untersucht werden, ob sie auch Vagheit adäquat modellieren kann.

Kehren wir zum Paradoxon des Sorites zurück, um diese Untersuchungen durchzuführen. Das Problem in der binären Logik war der abrupte Übergang von Nicht-Haufen zu Haufen – der Wechsel von Nicht-Haufen zu Unbestimmt-Haufen jedoch und von diesem zu Haufen muss ebenso abrupt erfolgen, da ein kontinuierlicher Übergang mittels drei Wahrheitswerten nicht möglich ist. Doch mit der Annahme einer Vagheit zweiter Stufe ist die Angabe einer solchen scharfen Grenze zwischen H(k), $k \in N$ ist wahr, unbestimmt oder falsch nicht sinnvoll. Mit der Annahme der Vagheit zweiter Stufe fällt also die Hoffnung, Vagheit mittels einer 3-wertigen-Logik adäquat modellieren zu können.[47]

3.5.2 Fuzzy Logik

In der fuzzy Logik stehen unendlich viele unterschiedliche Wahrheitswerte zur Verfügung, um eine Aussage zu valutieren. Zwischen den Extrema „Ein Korn ist kein Haufen“ ($\neg$H(1)) und „100.000 Körner sind ein Haufen“ (H(100.000)) können also beliebig viele verschiedene Abstufungen getroffen werden; die Wahrheitswertverteilung kann sogar stetig verlaufen. H(n) stehe – wie bereits eingeführt – für das Prädikat „n Körner sind ein Haufen“ und |H(n)| für

[44] Vgl. Gottwald 1989 S 1f
[45] Vgl. Gottwald 1989 S 5ff
[46] Vgl. Sinowjew 1986 S 34ff
[47] Vgl. Buldt 70f

dessen Wahrheitswert. Dann kann man eine Wahrheitswertverteilung bestimmen, die folgender Relation genügt:

$$0 = |H(1)|<|H(2)|<|H(3)|<\ldots<|H(n)| = 1 \text{ für sehr große } n\in N$$

Die Widersprüche im Paradoxon des Sorites ergeben sich durch den mehrfach angewandten Kettenschluss: Wenn eine Ansammlung kein Haufen ist und man gibt ein Korn dazu (und ein Korn ist auch kein Haufen), dann ist das Ganze auch kein Haufen, denn die Differenz von nur einem Korn kann nicht den Unterschied zwischen Haufen und Nicht-Haufen ausmachen. Dieser Schluss mag richtig sein, wenn man zu einem Korn ein zweites hinzufügt oder aber auch wenn man zu zwei Körnern ein drittes hinzufügt. Aber wie man sieht kann man durch das mehrmalige Hinzufügen nur eines einzelnen Kornes eine Ansammlung von 100.000 Körnern erzeugen, die sicherlich Haufen genannt werden kann. Versucht man dieses Problem mittels der fuzzy Theorie zu lösen, muss also ein unscharfer Folgerungsbegriff eingeführt werden, der die Widersprüchlichkeit des Sorites umgeht.

Wichtig für unscharfes Schließen ist die Formalisierung des Implikationsbegriffes, also das Definieren eines Operators, der die unscharfe wenn-dann Beziehung passend interpretiert und in eine formale Rechenvorschrift setzt, um den Kern des unscharfen Schlussfolgerns zu bilden.

Konjunktion und Disjunktion werden in unendlich-wertigen Logiken einheitlich mittels min bzw. max Operatoren definiert. Für die Subjunktion gibt es keine einheitliche Definition in den unendlich-wertigen Logiken: Verschiedene Logiker haben verschiedene Definitionen des Implikationsbegriffes gegeben. Hier soll zuerst auf die Definition Zadehs näher eingegangen werden, bevor andere Definitionen diverser Logiker kurz angedeutet werden.

Um die Subjunktionsdefinition Zadehs besser nachvollziehen und verstehen zu können, betrachten wir zuerst den scharfen Implikationsbegriff und äquivalente Formen[48]:

$$A \to B \Leftrightarrow \neg A \vee B$$

$$\Leftrightarrow (\neg A \vee B) \wedge (\neg A \vee A) \qquad \text{da } |\xi\wedge 1|=|\xi| \text{ und } |\neg\xi\vee\xi|=1$$

$$\Leftrightarrow \neg A \vee (B \wedge A)$$

$$\Leftrightarrow (A \wedge B) \vee (\neg A \wedge 1) \qquad \text{da } |\neg\xi\wedge 1|=|\neg\xi|$$

[48] Vgl. Jaanineh / Maijohann 1996 S 181

Mittels den bekannten Formalisierungen der Konjunktion und Disjunktion in der fuzzy Logik, der Festlegung der Negation durch $|\neg A| = 1-|A|$ und der Relation $\min(1-\xi, 1) = 1-\xi$ ergibt sich nun für die unscharfe Implikation in der fuzzy Logik:

$$|A \rightarrow B| := \max(\min(|A|,|B|), 1-|A|)$$ [49]

Wobei hier die Betragsstriche wieder den Wahrheitswert der Aussage symbolisieren.

Wie bereits erwähnt, wählen verschiedene Logiker diverse unterschiedliche Definitionen für den Implikationsoperator; im Folgenden werden der Vollständigkeit halber vier weitere Definitionen präsentiert[50]:

Mamdani Implikation: $|A \rightarrow B| := \min(|A|,|B|)$

Kleene Implikation: $|A \rightarrow B| := \max(1-|A|,|B|)$

Reichenbach Implikation $|A \rightarrow B| := 1-|A|+|A| \cdot |B|$

Lukasiewicz Implikation $|A \rightarrow B| := \min(1,1-|A|+|B|)$

Dies sind verschiedene Definitionen für denselben sprachlichen Ausdruck „Wenn A, dann B“ – was alle gemein haben ist, dass sie das wichtige Gebot der zweiwertigen Logik erfüllen, dass ein Schluss von etwas Wahrem zu etwas Falschem falsch ist: Gilt $|A|=1$ und $|B|=0$, hat eine Implikation $A \rightarrow B$ unter sämtlichen angeführten Implikationsoperatoren den Wahrheitswert 0, ist also falsch. Allgemeiner gesprochen gilt stets, dass der Schluss von etwas Wahrerem auf etwas weniger Wahres mit einem verringerten Wahrheitswert valutiert wird; umgekehrt wird der Schluss von etwas weniger Wahrem auf etwas Wahreres mit einem höheren Wahrheitswert bewertet. Je höher der Verlust an Wahrheitswert im Übergang von Antezedens zur Konklusion also ist, desto niedriger ist der Wahrheitswert, der dem gesamten Konditionalsatz zugesprochen wird.[51]

Betrachten wir nun den Schluss: Eine Ansammlung von einem Korn ist kein Haufen, das Hinzufügen eines Kornes macht daraus auch keinen Haufen, also ist eine Ansammlung von 2 Körnern kein Haufen; also formal: $\neg H(1) \wedge (\neg H(1) \rightarrow \neg H(2))$ ergo $\neg H(2)$. Dass eine 1-Korn-Ansammlung kein Haufen ist, wollen wir als völlig wahr bewerten, also $|\neg H(1)|=1$. Wie valutiert man nun den Wahrheitswert von $\neg H(2)$? Das ist tatsächlich ein schwerwiegendes Problem und eine Bewertung wird oftmals nicht eindeutig sein können; eine Bewertung ist oft subjektiv und in einer gewissen Form beliebig, vor allem bei der Bewertung von nicht

[49] Vgl. Jaanineh / Maijohann 1996 S 182
[50] Vgl. Jaanineh / Maijohann 1996 S 184
[51] Vgl. Sainsbury 1993 S 63ff

ganzzahligen Wahrheitswerten. Mathematische Sätze beispielsweise können hingegen oft eindeutig und objektiv als wahr oder falsch bewertet werden, etwa „5 ist eine Primzahl" ist wahr oder „2 teilt 3" ist falsch. Ebenso lassen sich empirische Sätze finden, denen eindeutig ein bestimmter Wahrheitswert zugeordnet werden kann; das Problem der Bewertung besteht hauptsächlich für ungewisse und vage Aussagen, oder für Aussagen, deren Wahrheitswert (noch) nicht exakt gekannt wird.

Was man nun bei der Bewertung von H(2) beachten muss, sind folgende zwei Bedingungen: $\neg H(2)$ soll weniger wahr sein als $\neg H(1)$ und es soll nicht viel weniger wahr sein. Dies sind äußerst vage Bedingungen, die vielerlei Wahrheitswertzuordnungen ermöglichen. Wir wollen uns hier nicht auf eine fixe Zuweisung eines exakten Wahrheitswertes beschränken; sei $|\neg H(2)|=1-\varepsilon$ und $\varepsilon>0$ möglichst klein. Der Unterschied im Wahrheitswert der beiden Aussagen $\neg H(1)$ und $\neg H(2)$ kann infinitesimal klein gewählt werden, sodass die beiden Aussagen grob geschätzt und oberflächlich betrachtet gleich wahr oder falsch sind. Der winzige Unterschied im Wahrheitswert lässt aber auch den Wahrheitswert des Konditionales $\neg H(1)\rightarrow\neg(2)$ schrumpfen, weshalb der ganze Schluss $\neg H(1) \wedge (\neg H(1)\rightarrow\neg H(2))$ ergo $\neg H(2)$ nicht folgerichtig im klassischen Sinn ist, und die Konklusion $\neg H(2)$ einen geringeren Wahrheitswert besitzen kann als seine Prämissen. Dies ist ein neuer Ansatz, das Problem des Sorites zu lösen: der modus ponens erhält nicht den Wahrheitsgrad der Prämissen, das heißt die Schlussfolgerung des Arguments $\neg H(n) \wedge (\neg H(n)\rightarrow\neg H(n+1))$ ergo$\neg H(n+1)$ kann einen geringeren Wahrheitswert haben als jede der Prämissen.[52] Anschaulich gesprochen kann man also den Schluss folgendermaßen interpretieren: „Eine n-Korn-Ansammlung ist kein Haufen, also ist eine (n+1)-Korn-Ansammlung auch kein Haufen" ist ein Schluss, der eine Konklusion mit verringertem Wahrheitswert impliziert. Zieht man also mehrmals Schlüsse der Form $\neg H(n)\wedge(\neg H(n)\rightarrow\neg H(n+1))$ ergo $\neg H(n+1)$ verliert jedes Mal die Konklusion an Wahrheitswert. Nach einer geeigneten Anzahl an Schlüssen ist der Wahrheitswert der Konklusion näher an 0 als an 1. Dieser Übergang von der Gültigkeit des Schlusses hin zu seiner Ungültigkeit wird mittels der Methoden der fuzzy Inferenz elastisch modelliert. Darin besteht der Vorteil des approximativen Schließens der fuzzy Theorie: Der stetige Übergang von wahr zu falsch und umgekehrt ermöglicht die elastische Modellierung von Problemstellungen.

[52] Vgl. Sainsbury 1993 S 63ff

Die fuzzy Theorie löst also nicht die Frage, ab wann nun eine Ansammlung an Körnern ein Haufen ist. Vage Termini bleiben vage und auch die fuzzy Theorie kann diese nicht präzisieren. Vagheit ist ein Phänomen, das sich nicht beheben lässt – zumindest nicht mit den hier vorgestellten Methoden. Was die fuzzy Theorie zur Lösung des Problems beitragen kann, ist die adäquate Modellierung des Phänomens Vagheit.

Ein abrupter Wechsel zwischen Haufen und Nicht-Haufen erschien unserer sprachlichen Auffassung zufolge als nicht sinnvoll; doch zwei- und dreiwertige Logiken implizieren abrupte Wechsel, weshalb wir sie als nicht adäquat zur Lösung des Sorites zurückwiesen. Mittels der fuzzy Theorie gelingt es nun, diesen Wechsel von Nicht-Haufen zu Haufen kontinuierlich verlaufen zu lassen, das Problem also stetig und elastisch zu modellieren, was unserem natürlichen Sprachempfinden von Vagheit am nächsten kommt. Die Modellierung des Problems ist nicht eindeutig; es lässt sich nicht präzise bestimmen, welchen Wahrheitswert welche Ansammlung bekommen sollte – dies ist eine Folgerung der Vagheit des Terminus „Haufen". Was eine Modellierung jedoch stets beinhaltet ist der kontinuierliche Übergang zwischen Haufen und nicht-Haufen; wie dieser Übergang numerisch genau verläuft ist nebensächlich.

3.5.3 Problemstellungen, die sich durch den Versuch, Vagheit zu formalisieren, ergeben

Bisher wurde versucht Paradoxien, die sich durch den Sorites ergeben, mittels der fuzzy Logik zu beseitigen. Es gelang uns, einen kontinuierlichen Übergang von Nicht-Haufen zu Haufen zu konstruieren, der unserem natürlichen Verständnis des Terminus Haufen sehr nahe kam. Wir diskutierten somit die Vorteile, die ein fuzzy theoretischer Zugang zu dem Problem gewährt; nun soll auf deren Nachteile verwiesen werden.

Bei der Analyse von Paradoxien wie dem Sorites ergeben sich meist drei alternative Lösungsansätze: Die erste mögliche Lösung ist die Verwerfung der Prämissen des paradoxen Arguments. Ein weiterer möglicher Lösungsansatz bietet sich durch das Akzeptieren der Schlussfolgerung der Paradoxie. Die dritte Möglichkeit, das Paradoxon zu umgehen, ergibt sich durch eine Ablehnung des logischen Folgerns.[53]

Alle drei Lösungsmöglichkeiten erscheinen als wenig wünschenswert. Die Alternative, die Prämissen zu verwerfen, gestaltet sich schwierig.[54] Die Schlussfolgerung zu akzeptieren, also zu akzeptieren, dass durch das Anhäufen vieler einzelner Körner kein Haufen entstehen

[53] Vgl. Sainsbury 1993 S 45
[54] Vgl. Sainsbury 1993 S 47ff

kann, erscheint ebenso wenig sinnvoll. Als dritte Alternative bleibt somit nur die Ablehnung des logischen Folgerns in seiner bisherigen gewohnten Form für Argumente vom Typus des Sorites. Dies ist der Lösungsansatz, der in dieser Arbeit verfolgt wurde und auch zu einer Vermeidung der Paradoxie führte.

Es muss jedoch hervorgehoben werden, welche Schwachstellen dieser Ansatz mit sich bringt. Dem Sorites liegt der modus ponens als Schlussform der klassischen Logik zugrunde; diesen abzulehnen, führt zu schwerwiegenden Problemen. Im in dieser Arbeit beschriebenen Lösungsansatz wurde die Gültigkeit des modus ponens – einem zentralen Prinzip des logischen Folgerns – angezweifelt. Der modus ponens ist der Schluss $p \wedge (p \rightarrow q)$ ergo q. Ihn anzuzweifeln bringt eine Reihe von neuen Problemstellungen mit sich und wird folglich auch wenn möglich vermieden – doch ist es genau diese Aufgabe des modus ponens, die die Widersprüchlichkeit des Sorites vermied und uns zu dem diskutierten Lösungsansatz führte. Untersucht man nun die Anforderungen an den modus ponens in der zweiwertigen Logik, erkennt man, dass die Aufgabe desselben in der mehrwertigen Logik der fuzzy Theorie durchaus mit der Beibehaltung der Schlussregel innerhalb der binären Logik vereinbar ist.[55] Sind Sätze entweder vollkommen wahr oder vollkommen falsch, bewahrt der modus ponens seine Gültigkeit.

Betrachten wir hierzu ein Beispiel: Der Wahrheitswert der Aussagen p und $p \rightarrow q$ sei gleich 1. Mittels dem modus ponens zieht man nun den Schluss $p \wedge (p \rightarrow q)$ ergo q – q hat also dank gültigem modus ponens den Wahrheitswert 1. Ohne den Schluss ergibt jedoch folgende Überlegung denselben Wahrheitswert für q: sei zunächst der Wahrheitswert von q unbekannt, also $|q| = x$ wobei $x \in [0,1]$. Wenn $|(p \rightarrow q)| = 1$ gelten soll und die Implikation der fuzzy Logik verwendet wird, ergibt sich (da $|p|=1$ gilt): $1 = |(p \rightarrow q)| = \max(\min(1, x), 0)$; also muss $\min(1,x)=1$ gelten. Dies impliziert nun $x = |q| = 1$. Das heißt, man kann innerhalb der fuzzy Theorie, falls sowohl die Aussage p als auch die Implikation $(p \rightarrow q)$ vollkommen wahr sind, folgern, dass auch q vollkommen wahr sein muss – ganz wie es der modus ponens implizieren würde. Es ist also unmöglich, dass die Prämissen des Arguments vollkommen wahr und die Konklusion vollkommen falsch ist – was die Definition des logischen Folgerns in der binären Logik erfüllt. Im Fall von nicht ganzzahligen Wahrheitswerten von p und $(p \rightarrow q)$ verliert der modus ponens an Gültigkeit, doch taucht dieser Fall in der klassischen Logik nicht auf.

Die fuzzy Theorie löst also die Paradoxie des Sorites mit dem Preis der Aufgabe des modus ponens als Schlussregel des logischen Folgerns; diese Aufgabe lässt sich jedoch

[55] Vgl. Sainsbury 1993 S 64

verteidigen mit der Begründung, dass der modus ponens nur aufgegeben werden muss, falls er auf Aussagen mit nicht ganzzahligen Wahrheitswerten angewandt werden soll. Für ganzzahlige Wahrheitswerte bleibt er weiterhin gültig, denn der modus ponens ist genau dann gültig, wenn [p∧(p→q)]→q eine Tautologie ist, also bei jeder beliebigen Wahrheitswertverteilung wahr ist, was folgende Wahrheitswerttabelle zeigt:

Aussage	p	q	p	p→q	p ∧ (p→q)	[p ∧ (p→q)] → q
Wahrheitswert	|p|	|q|	|p|	max(min(|p|,|q|),1-|p|)	min(|p|, |p→q|)	max(min(|p∧(p→q)|, |q|), 1-|p∧(p→q)|)
Verteilung	0	0	0	1	0	1
	0	1	0	1	0	1
	1	0	1	0	0	1
	1	1	1	1	1	1

Dieser Kompromiss, dass der modus ponens im allgemeinen Fall aufgehoben wird, jedoch als Spezialfall seine Gültigkeit bewahrt, lässt sich – meiner Ansicht nach – akzeptieren in Anbetracht der Möglichkeiten, die diese Aufgabe mit sich bringt.

3.6 Ein wahrscheinlichkeitstheoretischer Zugang zur Ungewissheit

„The only satisfactory description of uncertainty is probability."[56]

Mit diesen Worten beschreibt Dennis Lindley in seinem Aufsatz „The probability approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Systems" seine These von der Unumgänglichkeit der Wahrscheinlichkeitstheorie; er geht noch weiter und behauptet:

> *„that the calculus of probabilities is adequate to handle all situations involving uncertainty."*[57]

Der übliche Zugang zum Wahrscheinlichkeitskalkül ist ein axiomatischer; dieser hat den Vorteil, nicht darauf einzugehen, was Wahrscheinlichkeit und Zufall bedeuten; die Ergebnisse sind deshalb von der jeweiligen Interpretation unabhängig.

Der russische Mathematiker Andrei Nikolajewitsch Kolmogorow publizierte 1933 eine Schrift mit dem Titel „Grundbegriffe der Wahrscheinlichkeitsrechnung", in der er den Begriff der Wahrscheinlichkeit mit dem eines Maßes verband.[58] Die Maßtheorie bietet heute die Grundlage zum modernen Zugang zur Wahrscheinlichkeitstheorie.

Mathematisch werden Wahrscheinlichkeiten als eine Funktion P auf Ereignissystemen definiert, welche folgenden Axiomen genügt:

1.) $0 \le P(E) \le 1$ für alle Ereignisse E

2.) $P(S) = 1$ falls S ein sicheres Ereignis ist

3.) $P(A \cup B) = P(A)+P(B)$, falls $A \cap B = \{\}$

Mit diesen Eigenschaften ist die Wahrscheinlichkeitsfunktion P nun ein Maß.[59]

Es wurde bereits angedeutet, dass dieser axiomatische, maßtheoretische Zugang unabhängig von speziellen Interpretationen der Wahrscheinlichkeit ist; ein anderer Zugang ist der durch subjektive Wahrscheinlichkeit. Hier werden Wahrscheinlichkeiten als numerische Grade des Vertrauens in den Eintritt von Ereignissen E, die vom Informationsstand H

---

[56] Lindley 1987 S 17
[57] Lindley 1987 S 17
[58] Vgl. Hochkirchen S 129ff
[59] Vg. Viertel 1997 S 4ff

abhängen, interpretiert; die Notation hierfür ist $P(E|H)$ – dies entspricht also dem subjektiven Glauben, dass E eintritt, wenn man H weiß.

Wahrscheinlichkeit ist hier ein Maß des Vertrauens bzw. Glaubens in ein Ereignis, gegeben die Information H, das folgenden Eigenschaften genügen muss:

1.) $0 \leq P(E|H) \leq 1$

2.) $P(E_1 \vee E_2|H) = P(E_1|H)+P(E_2|H)$ für zwei sich gegenseitig ausschließende Ereignisse E_1, E_2

3.) $P(A \wedge B|H) = P(B|\ A \wedge H) \cdot P(A|H)$, falls das Ereignis A schon eingetreten ist und somit der Informationsstand $A \wedge H$ ist.[60]

Diese drei Eigenschaften könnten auch als Axiome eines Wahrscheinlichkeitskalküls dienen, da aus ihnen die gesamte Theorie ableitbar wäre. Man verzichtet jedoch gemeinhin darauf, da die Kolmogorowschen Axiome einfacher und intuitiver sind.[61]

3.7 Lindleys Kritik an der fuzzy Theorie

Wie eingangs erwähnt, ist Lindley der Auffassung, der Wahrscheinlichkeitskalkül sei der einzige Weg, Ungewissheit adäquat zu modellieren. Einen großen Vorteil der Wahrscheinlichkeitslehre sieht er in dem einfachen und intuitiven Fundament, welches die Axiome bilden. Anderen Wegen, Ungewissheit zu modellieren, unterstellt Lindley unsauber und ohne axiomatische Basis zu sein.[62] Wenigstens im Bezug auf die fuzzy Theorie sehe ich diesen Einwand nicht; die Theorie der unscharfen Mengen ist ebenfalls präzise formuliert, und trotz ihrer Beschäftigung mit unscharfen Objekten ist sie selbst scharf definiert. Eine axiomatische Basis mag fehlen, doch kann dies auch an dem jungen Alter der Theorie liegen.

Lofti Zadeh kritisierte die Behauptung, allein mittels Wahrscheinlichkeitsrechnung Ungewissheit modellieren zu können, und entgegnete, es sei eine gravierende Einschränkung der probabilistischen Methoden, dass diese nicht mit der durchdringenden Vagheit der Information zu Recht kommen.[63]

Lindley konterte diesen Einwänden mit der Aufforderung zu demonstrieren, dass ein Problem, gelöst mittels fuzzy Theorie, nicht mittels eines probabilistischen Zuganges besser gelöst werden könne. Diese Aufforderung ist äußerst pragmatischer Natur, denn sie richtet

[60] Vgl. Viertel 1997 S 6
[61] Vgl. Lindley 1987 S 18
[62] Vgl. Lindley 1987 S 19f
[63] Vgl. Zadeh 1983

sich auf die tatsächliche Lösbarkeit von Problemen.[64] Weiters impliziert sie nicht, dass alle Probleme, die Unsicherheit behandeln, durch die Wahrscheinlichkeitstheorie gelöst werden können; die Behauptung Lindleys besagt nur, dass sie es besser tun kann als die Alternativen.

Im Folgenden sollen Beispiele Lindleys erörtert werden, mit denen gezeigt werden soll, wie typische Aufgaben der fuzzy Theorie mittels der Wahrscheinlichkeitstheorie behandelt werden.

Als Beispiel einer unscharfen Aussage behauptete Zadeh „Berkeleys Bevölkerung ist über 100.000.“; diese Aussage ist unscharf aufgrund des impliziten Verständnisses von „über“ – es bedeutet etwas mehr als 100.000 aber nicht viel mehr. Lindleys probabilistischer Zugang wäre nun eine Wahrscheinlichkeitsaussage über eine Quantität, die bestimmt werden kann. (Im Allgemeinen sollten alle Quantitäten stets evaluierbar sein, da man diese auch verwenden will.) Eine mögliche bestimmbare Quantität wäre die Antwort des zuständigen Amtes für Statistik, Wahlen und Einwohnerwesen, welche X genannt sei. Die Wahrscheinlichkeitsaussage der Behauptung ist nun P(X|H), wobei H das Wissen desjenigen ist, der die Behauptung aufgestellt hat.

Alle vagen Aussagen dieser Gestalt können auf analoge Weise wahrscheinlichkeitstheoretisch interpretiert werden. Mehr Vorsicht ist bei Aussagen vom Typus „Heinrich ist jung.“ geboten; Heinrich ist hier eine wohldefinierte Person, sein Alter hingegen ein vages Prädikat X. Die Aussage hängt davon ab, wie und wo sie gemacht wird: Auf dem Universitätscampus bedeutet sie, Heinrich wäre um die 20, in einem Seniorenheim hingegen würde ein junger Heinrich vielleicht um die 60 sein. Folglich ist die Information vom Kontext äußerst relevant. Ohne Informationen über den Kontext der Aussage wird man P(X|H) abschätzen müssen.[65]

Als weiteren Einwand gegen die fuzzy Theorie brachte Lindley das Argument, diese sei ein komplizierteres Konzept als die Wahrscheinlichkeitslehre. Fuzzy Logik führt zu nichtlinearem Programmieren und beinhaltet große Komplexität in der Sprache. Im Gegensatz dazu ist der Wahrscheinlichkeitskalkül äußerst einfach, benötigt er doch nur drei einfache Axiome. Im Sinne der Methode des Willhelm von Ockham (Ockhams Rasiermesser) sollte nun die Wahrscheinlichkeitslehre gegenüber der fuzzy Theorie favorisiert werden, da sie nicht

[64] Vgl. Lindley 1987 S 19ff
[65] Vgl. Lindley 1987 S 20

nur alle Probleme, die sich mittels der fuzzy Logik behandeln lassen, gleich oder besser löst, sondern darüber hinaus auch noch einfacher ist.[66]

Ein letztes Argument gegen die fuzzy Theorie wird aus der Frage abgeleitet, warum man Ungewissheit untersucht. Abgesehen vom intellektuellem Vergnügen, das Lindley außer Acht lässt, ergibt sich ihm nur eine einzige mögliche Antwort: um Entscheidungen trotz Ungewissheit zu fällen. Eine axiomatische Aufbereitung der Theorie des Entscheidungstreffens zeigt nach Savage, dass nur eine Maximierung des erwarteten Nutzens eine befriedigende Prozedur darstellt. Diese benutzt allerdings Wahrscheinlichkeiten und nur diese sind die benötigten Quantitäten für den Prozess des Entscheidungstreffens.[67]

Gerade im Bereich der Anwendungen – auch in der Theorie der Entscheidungen – führte die fuzzy Theorie allerdings in den letzten Jahren einen großen Siegeszug, weshalb die Aktualität und Gültigkeit dieses letzten Arguments, warum der Wahrscheinlichkeitskalkül der fuzzy Theorie als Theorie der Ungewissheit vorzuziehen ist, verloren ging.

---

[66] Vgl. Lindley 1987 S 22
[67] Vgl. Lindley 1987 S 22

4 Ein geometrischer Zugang zur fuzzy Theorie

Die Einführung unscharfer Mengen in dieser Arbeit erfolgt analog zur Defnition Lofti Zadehs aus dem Jahr 1965 mittels Zugehörigkeitsfunktionen, also Abbildungen μ_A von der Grundmenge X in das Einheitsintervall [0,1]. Der Nachteil dieses Zuganges liegt in den Schwierigkeiten, die er bei der Veranschaulichung einer unscharfen Menge bereitet: Eine unscharfe Menge A ist die Menge aller Zahlenpaare (x; μ(x)), also

$$\mathrm{A} := \{\, (x;\mu(x) \big| x \in X, 0 \leq \mu(x) \leq 1 \in \Re\}$$

Betrachten wir zur Veranschaulichung die unscharfe Menge aller Jugendlichen. Als erstes muss eine Zugehörigkeitsfunktion μ definiert werden; diese Definition wird nicht eindeutig bestimmt sein, da „Jugendlicher" ein vager Begriff ist und man nicht zweifelsfrei feststellen kann, wer unter diesen Terminus fällt. Seien also beispielsweise alle 14- bis 18jährigen zweifelsfrei jugendlich; weiters seien Kinder ab zwölf und junge Erwachsene bis 21 zu einem gewissen Grad jugendlich – dann könnte die fuzzy Menge J aller Jugendlicher folgendermaßen definiert sein:

J = {(12jähriger,0.5), (13jähriger,0.75), (14jähriger,1), (15jähriger,1),
(16jähriger,1), (17jähriger,1), (18jähriger,1), (19jähriger,0.8),(20jähriger,0.4),
(21jähriger,0.2)}

Die Lesart dieser Notation gestaltet sich wie folgt: Das Element (x,y) bedeutet, dass ein Mensch mit dem Alter x einen Zugehörigkeitswert y zur Menge aller Jugendlichen besitzt – also hat beispielsweise ein 15jähriger einen Zugehörigkeitswert 1, hingegen ein 20jähriger nur eine Zugehörigkeit von 0.4. Elemente, die nicht in dieser Aufzählung aufscheinen, also unter 12jährige und über 21jährige, haben eine Zugehörigkeitswert von 0 und werden deshalb nicht aufgezählt.

Dieser Zugang ist wenig anschaulich und bietet somit Platz für Kritik, wie es unter anderem Dennis Lindley tut, wenn er einen der Vorzüge der Wahrscheinlichkeitslehre gegenüber der fuzzy Theorie in der Einfachheit des Wahrscheinlichkeitskalküls und seiner Axiome sieht.[68]

Dem gibt Bart Kosko in seinem Artikel „Fuzziness vs. Probability" Abhilfe, indem er eine neue Interpretation unscharfer Mengen einführt; diese Interpretation ist stark geometrisch und deshalb äußerst anschaulich – zumindest im Spezialfall einer nur zwei- bzw.

[68] Vgl. Lindley 1987 S 22

dreielementigen Grundmenge X. Die Geometrie unscharfer Mengen bezieht sich auf die Grundmenge $X=\{x_1,x_2,x_3,\ldots,x_n\}$ und den Bildbereich [0,1] der Zugehörigkeitsfunktionen. Ausgangpunkt der Interpretation ist die Menge aller unscharfen Teilmengen von X, der so genannten fuzzy Potenzmenge $F(2^X)$. Diese fuzzy Potenzmenge wird durch einen n-dimensionalen Einheitswürfel veranschaulicht, also einem Hyperwürfel mit n Kanten der Länge 1. Eine fuzzy Menge ist nun ein Punkt in diesem n-dimensionalen Hyperwürfel.[69] An dieser Stelle wird ersichtlich, dass die Ordnung der Grundmenge X wesentlich ist: Identifiziert man die Potenzmenge der n-elementigen Grundmenge X mit dem n-dimensionalen Einheitswürfel, ist die Ordnung von X insofern wesentlich, da sie sich auf die Geometrie des Einheitswürfels überträgt. Betrachtet man beispielsweise die Grundmenge {Hütte, Haus, Palast}, kann deren unscharfe Potenzmenge mit einem dreidimensionalen Einheitswürfel identifiziert werden, bei der Zugehörigkeitswerte von Hütte jeweils auf der X-Achse des Würfels, von Haus auf der Y-Achse und von Palast auf der Z-Achse angenommen werden. Ist die Reihenfolge der Elemente in der Grundmenge eine andere, etwa {Haus, Hütte, Palast} erkennt man, dass die entsprechenden Zugehörigkeitswerte auf anderen Achsen aufgetragen werden müssen. Die Ordnung der Grundmenge ist also wesentlich und sei von nun an stets festgelegt.

Die Stärke dieses Zuganges liegt – wie bereits erwähnt – in der Anschaulichkeit, die sie im Spezialfall einer zwei- oder dreielementigen Grundmenge X bietet; anders als Bart Kosko selbst sehe ich jedoch starke Veranschaulichungsschwierigkeiten bei beliebigen Grundmengen X, insbesondere unendlichdimensionalen: Der menschliche Geist ist nicht in der Lage sich einen n-dimensionalen Einheitswürfel bei $n\geq 4$ zu visualisieren; noch weniger ist er in der Lage, sich unendlichdimensionale Hyperwürfel vorzustellen.

Bart Kosko sieht in der Anschaulichkeit der Geometrie der fuzzy Mengen selbst eines der größten Argumente für die fuzzy Theorie.[70] Infolge der obigen Argumentation gegen eine allgemeine Vorstellbarkeit dieser geometrischen Interpretation sehe ich in ihr kein starkes Argument für die fuzzy Theorie. Im Spezialfall einer zweielementigen Grundmenge $X=\{x_1,x_2\}$ bietet die geometrische Interpretation jedoch eine große Hilfe im Verstehen der fuzzy Theorie und ihrer Theoreme; deshalb wollen wir uns im Folgenden auf diesen Fall konzentrieren:

Die fuzzy Potenzmenge $F(2^X)$ der Menge $X=\{x_1,x_2\}$ ist also ein Quadrat der Länge 1. Jede fuzzy Menge der Grundmenge X ist ein Punkt in diesem Quadrat; die Eckpunkte des

[69] Vgl. Kosko 1990 S 216ff
[70] Vgl. Kosko 1990 S 216

Quadrats entsprechen den klassischen scharfen Mengen {}, $\{x_1\}$, $\{x_2\}$ bzw. $\{x_1,x_2\}$. Die klassische Potenzmenge entspricht also der Menge sämtlicher Eckpunkte. Fasst man die Punkte des Quadrats als zweidimensionale Vektoren auf, kann man die vier Ecken folgendermaßen schreiben: {(0,0), (1,0), (0,1), (1,1)}. Diese vier Vektoren sind also die Elemente der klassischen Potenzmenge $2^X=\{\ \{\}, \{x_1\},\{x_2\},X\ \}$. Identifiziert man nun diese Menge mit der Menge der vier Vektoren, deutet die 1 im i-ten Argument des Vektors die Präsenz des Elementes x_i, i=1,2 an und die 0 die Absenz.

Analog gelingt die Identifizierung des Punktes $A=(\xi,\omega)$ mit der fuzzy Menge $A=\{(x_1,\xi), (x_2,\omega)\}$, das heißt der Wert an der i-ten Stelle des Punktvektors korrespondiert mit der Zugehörigkeit des Elements x_i zu A, die durch das Bild der Zugehörigkeitsfunktion $\mu_A(x_i)$ dargestellt wird.

Betrachten wir zur Veranschaulichung dieser Interpretation nochmals die in diesem Kapitel bereits eingeführte unscharfe Menge J aller Jugendlicher: Diese bestand aus zehn Zahlenpaaren. Mittels der neu eingeführten geometrischen Interpretation einer fuzzy Menge lässt sich diese Menge J nun als ein zehndimensionaler Vektor schreiben, der an der i-ten Koordinate jeweils den Zugehörigkeitswert des i-ten Zahlenpaares trägt.

$$J = (0.5, 0.75, 1, 1, 1, 1, 1, 0.8, 0.4, 0.2)$$

Ein Vorteil dieser Notation ist sicherlich die verkürzte Darstellung; was jedoch verloren geht, ist die Information, was denn den Zugehörigkeitswert trägt. In der alten Darstellung mittels Zahlenpaaren ließ sich sofort erkennen, dass (13jähriger, 0.75) dafür steht, dass ein Dreizehnjähriger zu 0.75 ein Jugendlicher ist. Mit der neuen Darstellung kann man nur erkennen, dass die zweite Koordinate des Vektors J die Zugehörigkeit 0.75 besitzt. Es muss also extern vermerkt werden, was jede einzelne Koordinate bedeutet, für wen oder was sie also steht.

Zum besseren Verständnis wollen wir nun noch klassische Mengen betrachten, etwa C={2,3} und D={1,2,3,4}; die Grundmenge für diese beiden Mengen sei X={1,2,3,4}. Dann lassen sich die beiden Mengen in der neuen Schreibweise als vierdimensionale Mengenvektoren auffassen; man erhält also: C = (0, 1, 1, 0) bzw. D = (1, 1, 1, 1)

Im Laufe des Kapitels wird ersichtlich werden, dass auf fuzzy Mengen verschiedene Operationen und Maße definiert werden können, die nur von den Zugehörigkeitswerten abhängen; es ist also durchaus zweckdienlich, nur diese bei der Definition einer Menge anzugeben. Weiters bietet die neue Darstellung die Möglichkeit einer geometrischen Darstellung, wie im Folgenden gezeigt werden soll.

Unten stehende Abbildung 1 visualisiert die neue geometrische Interpretation für fuzzy Mengen auf der zweielementigen Grundmenge $X=\{x_1,x_2\}$, was uns die Darstellung der fuzzy Potenzmenge als Quadrat erlaubt: Die Eckpunkte des Quadrats sind die scharfen Mengen $\{x_1\}$, $\{x_2\}$, X sowie die leere Menge {}. Der Punkt A=(1/3, 3/4) ist die fuzzy Menge $\{(x_1,1/3), (x_2,3/4)\}$.

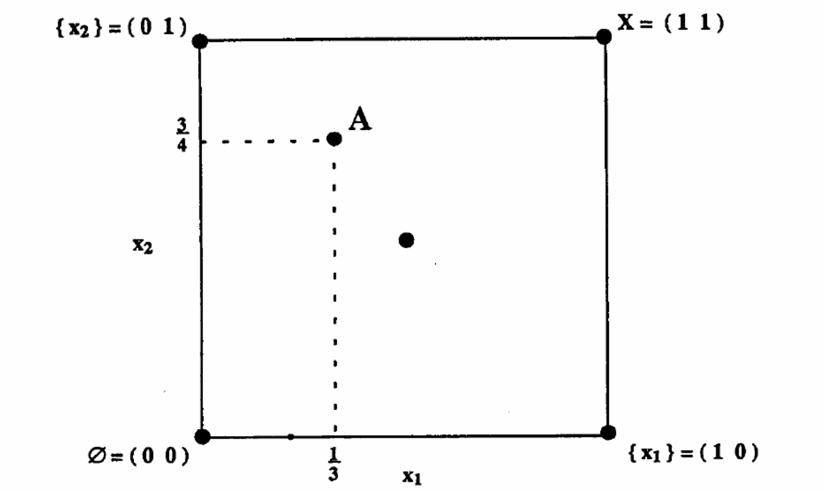

Abbildung 1

Mittels dieser Interpretation wird auch ein geometrischer Zusammenhang bei Komplementbildung, Durchschnitt und Vereinigung ersichtlich: Der Durchschnitt zweier Mengen wird durch paarweise Minimumbildung und die Vereinigung durch paarweise Maximumbildung berechnet. Folgendes Beispiel veranschauliche dies:

$$A=(\frac{1}{3},\frac{3}{4})$$

$$A^C=(\frac{2}{3},\frac{1}{4})$$

$$A\cap A^C=(\frac{1}{3},\frac{1}{4})$$

$$A\cup A^C=(\frac{2}{3},\frac{3}{4})$$

Dies wird durch folgende Abbildung dargestellt:

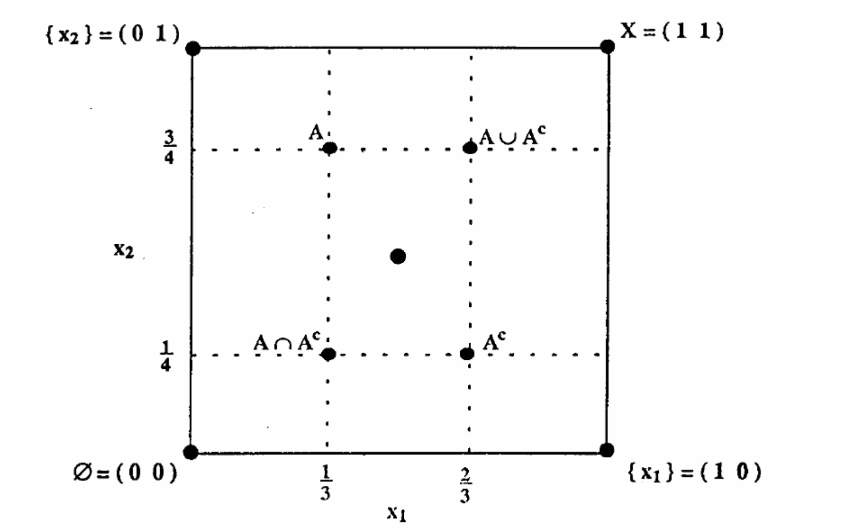

Abbildung 2

Als Spezialfall wird ersichtlich, dass der Mittelpunkt M=(0.5, 0.5) als einziger Punkt mit seinem Komplement übereinstimmt; ferner gilt:

$$M = M^C = M \cap M^C = M \cup M^C$$

Es mag eine praxisrelevante Heuristik sein, fuzzy Mengen mittels scharfer Mengen zu approximieren, indem man Punkte im Inneren des Quadrats durch den nächsten Eckpunkt ersetzt. Dies mag oft sinnvoll sein, da es bedeutend einfacher ist, von einem leeren Glas zu sprechen, als von einem fast leeren oder von einer schönen Frau als von einer ziemlich schönen Frau. Doch findet diese Abschätzungsheuristik ihre Limitationen, je näher man dem Mittelpunkt kommt; je näher man dem Mittelpunkt kommt, desto schwieriger wird es, die fuzzy Menge durch den nächstgelegenen Eckpunkt zu ersetzen, da für die Mitte alle Eckpunkte gleich weit entfernt sind, und somit keiner sinnvoll ausgewählt werden kann.[71]

4.1 Maße für fuzzy Mengen

Wie bereits festgehalten wurde, sind verschiedene Punkte des Quadrats verschieden unscharf – die Extrema sind die Eckpunkte, welche klassischen scharfen Mengen entsprechen und der Mittelpunkt, der als am unschärfsten bezeichnet wurde. Dies selbst ist eine äußerst vage Behauptung, da die Relation „ist schärfer als“ unscharf ist; dem soll jetzt Abhilfe geschaffen werden, indem ein Maß eingeführt wird, Unschärfe zu quantifizieren. Bevor dies getan wird,

[71] Vg. Kosko 1990 S 219f

soll noch ein natürlicher Abstandbegriff definiert werden: Man definiere das Maß M auf der Grundmenge $X=\{x_1,x_2,\ldots,x_n\}$ mittels

$$M(A) := \sum_{x\in X} \mu_A(x)$$

M entspricht dann der Kardinalität bzw. Mächtigkeit[72] einer Menge; im Spezialfall einer scharfen Menge gilt $\mu_A(x) = 1$ für alle $x\in A$, woraus $M(A) = |A|$ im Sinne der Definition von Mächtigkeit in der klassischen Mengenlehre gilt. Eine klassische dreielementige Menge hat also die Mächtigkeit drei. Wie berechnet sich nun die Mächtigkeit der hier eingeführten Menge aller Jugendlichen? Diese wurde mittels der geometrischen Interpretation folgendermaßen definiert:

$$J = (0.5, 0.75, 1, 1, 1, 1, 1, 0.8, 0.4, 0.2)$$

Die Mächtigkeit von J – M(J) – berechnet sich nach Definition als Summe sämtlicher Koordinaten, also

$$M(J) = 0.5+0.75+1+1+1+1+1+0.8+0.4+0.2 = 7.65.$$

Die Mächtigkeit M(A) einer Menge A entspricht darüber hinaus der so genannten l^1 Metrik, welche wir mit *d* bezeichnen wollen und folgendermaßen definiert ist:

$$d(A,B) := l^1(A,B) := \sum_{x\in X} |\mu_A(x) - \mu_B(x)|$$

Den Zusammenhang zwischen der Kardinalität M und der l^1 Metrik zeigt folgender kurzer Beweis[73]:

$$M(A) = \sum_{x\in X} \mu_A(x) = \sum_{x\in X} |\mu_A(x) - 0| = \sum_{x\in X} |\mu_A(x) - \mu_{\{\}}(x)| = l^1(A,\{\})$$

Wie man aus der Topologie, einer mathematischen Theorie, die sich mit Abständen beschäftigt, weiß, ist Metrik ein verallgemeinerter Abstandbegriff; die l^1 Metrik, die das oben eingeführte Maß M induziert, entspricht dem Abstand zweier Punkte, indem man stets entlang den Koordinatenachsen entlang von einem Punkt zum zweiten wandert und dabei die Entfernung misst. M(A) entspricht somit dem Abstand, den der Punkt A zu der linken unteren Ecke hat, wenn man zur Messung der Entfernung stets entlang der Koordinatenachsen wandert.

[72] Vgl. Jaanineh / Maijohann 1996 S 76f
[73] Vgl. Kosko 1990 S 220f

Durch das obige M gelingt nun die Definition des so genannten fuzzy Entropiemaßes E, das die Unschärfe einer Menge quantifiziert:[74]

$$E(A) := \frac{M(A \cap A^C)}{M(A \cup A^C)}$$

Aufgrund der besonderen geometrischen Eigenschaften der Durchschnitts- und Vereinigungsbildung von A und A^C, wie sie in Abbildung 2 visualisiert wurden, ergibt sich folgende geometrische Deutung des Entropiemaßes: $M(A \cap A^C)$ entspricht dem Abstand von A zum nächst gelegenen Eckpunkt; $M(A \cup A^C)$ entspricht dem Abstand von A zum weitest entfernt gelegenen Eckpunkt.

Das fuzzy Entropiemaß unterstreicht nochmals den Zusammenhang zwischen der Aufgabe des Prinzips vom ausgeschlossenen Dritten und der fuzzy Logik: Wird das Prinzip vom ausgeschlossenen Dritten angenommen, sind $M(A \cup A^C) = n$ und $M(A \cap A^C) = 0$ bei einer n-elementigen Grundmenge. Für das fuzzy Entropiemaß ergibt sich somit 0/n also E(A)=0 für eine klassische Menge. Für dem Mittelpunkt M ergab sich nach obigen Rechnungen $M \cup M^C = M \cap M^C$, also E(M)=1. Aufgrund der Monotonie des Maßes M und der Eigenschaft, dass $A \cap A^C$ stets in $A \cup A^C$ enthalten ist, also $A \cap A^C \subseteq A \cup A^C$ gilt, ergibt sich $M(A \cap A^C) \leq M(A \cup A^C)$ und folglich $0 \leq E(A) \leq 1$ für alle Mengen A. Folglich spricht man von einer unschärferen Menge, falls das fuzzy Entropiemaß E größer ist.

Zur Veranschaulichung betrachten wir zunächst die klassische Menge $C = \{x \in N | x<5\}$; das Komplement dieser Menge ergibt sich durch die Darstellung $C^C = [x \in N| x \geq 5\}$. Der Durchschnitt dieser beiden Mengen ist die leere Menge mit Mächtigkeit 0; die Vereinigung ist die Menge der natürlichen Zahlen – deren Mächtigkeit wird $\aleph_0$ (Aleph 0) genannt. Das Entropiemaß der klassischen Menge C ist also $E(C)=0/\aleph_0$ was definitionsgemäß 0 ist. Das Entropiemaß von C als klassischer Menge ist also gleich 0.

Betrachten wir nun die fuzzy Mengen A = (0.5, 0.4) und B = (0.7, 0.9). Die Komplemente ergeben sich durch $A^C = (0.5, 0.6)$ und $B^C = (0.3, 0.1)$. Die Entropiemaße der beiden Mengen sind folglich:

$$E(A) = \frac{M(A \cap A^C)}{M(A \cup A^C)} = \frac{M((0.5,0.4))}{M((0.5,0.6))} = \frac{0.9}{1.1} = 0.8181818$$

$$E(B) = \frac{M(B \cap B^C)}{M(B \cup B^C)} = \frac{M((0.3,0.1))}{M((0.7,0.9))} = \frac{0.4}{1.6} = 0.25$$

[74] Vgl. Kosko 1990 S 222f

Die Relation „schärfer als“ lässt sich also durch das Entropiemaß E formalisieren; Mengen mit hohem Entropiemaß (das maximal gleich 1 sein kann) werden als besonders unscharf eingestuft, Mengen mit niedrigem Entropiemaß als besonders scharf. Wie bereits an diesem einfachen Beispiel deutlich wurde, ist eine Menge unschärfer als eine andere, falls seine Zugehörigkeitsgrade näher bei 0.5 liegen.

4.2 Teilmengigkeit

Nach der Einführung der Inklusionsbeziehung für unscharfe Mengen durch Lofti Zadeh ist eine Menge A Teilmenge einer anderen Menge B oder nicht; die Relation $\subseteq$ selbst ist scharf und nicht fuzzy. Betrachtet man allerdings Mengen, wird man oft sehen, dass die eine zu einem großen Teil in der anderen enthalten ist – jedoch nicht ganz. Für solche Fälle scheint es wenig adäquat – wenn man einmal die Wege der klassischen scharfen Mengenlehre verlassen hat – die Teilmengeneigenschaft scharf aufzufassen und nur zwischen ist-Teilmenge-von und ist-es-nicht zu unterscheiden. Aus diesem Grund wird das Teilmengigkeitsmaß S eingeführt, dass den numerischen Grad der Teilmengigkeit von A in B entspricht; das heißt: S(A,B) = Grad($A \subseteq B$). Es ist unmittelbar evident, dass S folgende Eigenschaften haben soll: Ist A in B enthalten, soll S(A,B) = 1 gelten. Weiters soll S(A,B) = 0 sein, falls A kein gemeinsames Element mit B besitzt. Aus der Definition Zadehs[75] folgt, dass $\max(0, \mu_A(x)-\mu_B(x)) = 0$, falls x in B, jedoch nicht in A enthalten ist; $\max(0, \mu_A(x)-\mu_B(x)) > 0$, falls x in A und B enthalten ist. Dies motiviert folgende Definition:[76]

$$S(A,B) = 1 - \frac{\sum_{x \in X} \max(0, \mu_A(x) - \mu_B(x))}{M(A)}$$

Diese Definition ist wohldefiniert und stimmt für die Spezialfälle der alten Inklusionsrelation überein: Ist $A \subseteq B$ nach der Definition Zadehs, gilt also $\mu_A(x) \leq \mu_B(x)$ für alle $x \in X$, dann ist $\mu_A(x)-\mu_B(x)) \leq 0$ und folglich $\max(0, \mu_A(x)-\mu_B(x)) = 0$ für alle $x \in X$. Es gilt folglich S(A,B)=1-0=1, falls $A \subseteq B$. Hat umgekehrt A kein gemeinsames Element mit B, dann ist $\max(0, \mu_A(x)-\mu_B(x)) = 0$, falls $x \in B$ und $\max(0, \mu_A(x)-\mu_B(x)) = \mu_A(x)$ falls $x \in A$. Daraus ergibt sich für S(A,B)=1-M(A)/M(A)=1-1=0. Zur Illustration dieses Sachverhalts betrachte man die klassischen Mengen C = {2,3} und D = {1,2,3,4}: Es gilt $C \subseteq D$. Das Teilmengigkeitsmaß berechnet sich nun wie folgt:

[75] Vgl. Zadeh 1965 S 340
[76] Vgl. Kosko 1990 S 226f

$$S(C,D) =$$

$$= 1 - \frac{\sum_{i=1}^{4} \max(0, \mu_{\{2,3\}}(i) - \mu_{\{1,2,3,4\}}(i))}{M(\{2,3\})} =$$

$$= 1 - \frac{\max(0,-1) + \max(0,0) + \max(0,0) + \max(0,-1)}{2} =$$

$$= 1 - \frac{0}{2} =$$

$$= 1$$

Allerdings finden sich auch Elemente von D, die auch in C enthalten sind; es liegt also nahe eine gewisse Teilmengigkeit von D in C zu vermuten. Man ermittelt hierzu:

$$S(D,C) =$$

$$= 1 - \frac{\sum_{i=1}^{4} \max(0, \mu_{\{1,2,3,4\}}(i) - \mu_{\{2,3\}}(i))}{M(\{1,2,3,4\})} =$$

$$= 1 - \frac{\max(0,1) + \max(0,0) + \max(0,0) + \max(0,1)}{4} =$$

$$= 1 - \frac{2}{4} =$$

$$= \frac{1}{2}$$

Man sieht also, das Konzept der unscharfen Teilmengigkeit lässt sich auch auf klassische Mengen anwenden, denn auch scharfe Mengen können zu einem gewissen Grad in anderen enthalten sein, aber eben nicht ganz so wie beispielsweise D teilweise in C enthalten ist. Betrachtet man nun die fuzzy Mengen A = (2/3, 1/4) und B = (1/3, 2/3), dann erhält man:

$$S(A,B) = 1 - \frac{\max(0,1/3) + \max(0,-5/12)}{2/3 + 1/4} = 1 - \frac{1/3}{11/12} = \frac{7}{11}$$

Ebenso erhält man:

$$S(B,A) = 1 - \frac{\max(0,-1/3) + \max(0,5/12)}{1/3 + 2/3} = 1 - \frac{5/12}{1} = \frac{7}{12}$$

Die fuzzy Menge A ist also zum Grad 7/11 in der Menge B enthalten, während B in A zum Grad 7/12 enthalten ist; mit der Definition Zadehs einer scharfen Teilmengigkeit wäre weder A in B noch B in A enthalten gewesen, da für Teilmengigkeit jede Komponente des einen Mengenvektors kleiner oder gleich der Komponente des anderen Vektors hätte sein müssen, was bei den angeführten Mengen nicht der Fall ist.

Es zeigt sich nun ferner, dass 0 bzw. 1 die untere bzw. obere Schranke des Teilmengigkeitsmaßes von S ist; für sämtliche unscharfen Mengen A und B gilt somit: $0 \leq S(A,B) \leq 1$.

In der klassischen Mengenlehre korrespondiert die Inklusion $\subseteq$ mit der Subjunktion $\rightarrow$ vermöge der Festlegung: $A \subseteq B$ genau dann wenn $\forall x$: $x \in A \rightarrow x \in B$. Analoges gilt für das Teilmengigkeitsmaß S, das mit dem Lukasiewiczschen Subjunktor korrespondiert, falls M(A)=1.[77]

$$
\begin{aligned}
&S(A,B) = \\
&= 1 - \max(0, \mu_A(x) - \mu_B(x)) = \\
&= 1 - [1 - \min(1-0, 1 - (\mu_A(x) - \mu_B(x))] = \\
&= \min(1, 1 - \mu_A(x) - \mu_B(x)) = \\
&= Subjunktion_{\text{Lukasiewicz}}(A,B)
\end{aligned}
$$

$\text{Subjunktion}_{\text{Lukasiewicz}}(A,B)$ entspricht hierbei der Subjunktion $A \rightarrow B$ nach Definition von Lukasiewicz[78]

4.2.1 Das Teilmengigkeitsmaß in der geometrischen Deutung

Betrachtet man die geometrische Visualisierung der fuzzy Potenzmenge von B in Abbildung 3 erkennt man, dass A Teilmenge von B ist, falls sich A innerhalb des Vierecks $F(2^B)$ befindet; intuitiv vermutet man, dass je näher sich ein Punkt A zum Viereck befindet, desto größer seine Teilmengigkeit S(A,B) sein sollte.

[77] Vgl. Kosko 1990 S 227
[78] Vgl. Gottwald 1989 S 34

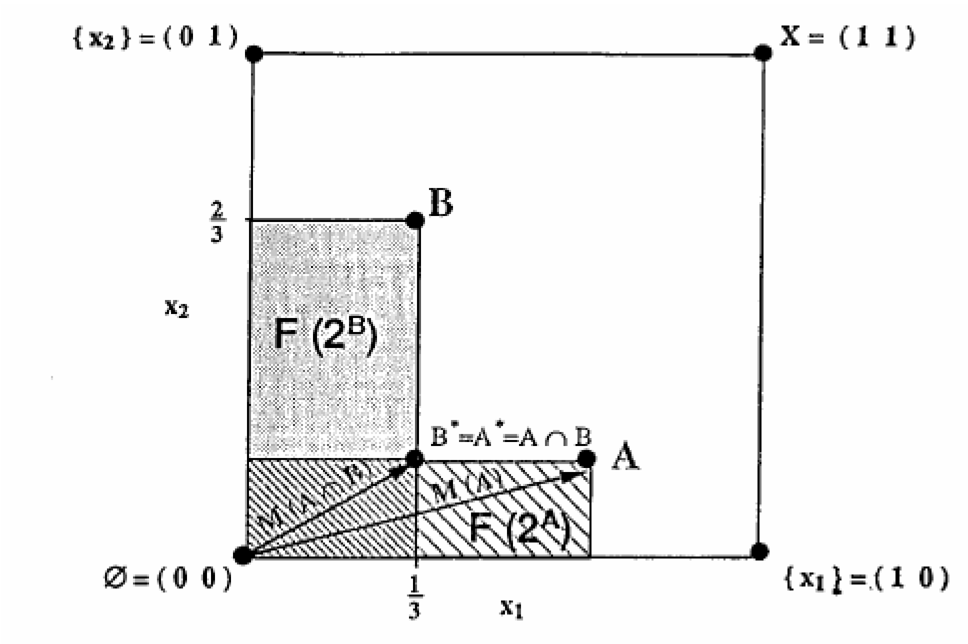

Abbildung 3

Die Vorgehensweise zur Findung einer geometrischen Charakterisierung des Teilmengigkeitsbegriffs wird also eine metrische sein: Es bezeichne d(A,B) also den Abstand von A zu B, der mittels der l^1 Metrik (sie sei von nun an wieder d genannt) berechnet wird. $d(A,F(2^B))$ bezeichne den minimalen Abstand zwischen A und dem Viereck $F(2^B)$, also $d(A,F(2^B)) = \min\{d(A,B') \mid B' \in F(2^B)\}$. Das Element aus $F(2^B)$, welches A am nächsten liegt, sei mit B* bezeichnet, folglich gilt $d(A,F(2^B)) = d(A,B^*)$. Mittels dieses Abstandes definieren wir nun:[79]

$$S(A,B) = 1 - \frac{d(A,B^*)}{M(A)}$$

Berechnen wir nun zur Gewöhnung an die Definition die Teilmengigkeit von A in B auf diese neu eingeführte Weise, wobei A und B wieder wie folgt definiert sind: A = (2/3, 1/4) und B = (1/3, 2/3). Wie bereits in Abbildung 3 ersichtlich, ist der minimale Abstand von A zu $F(2^B)$ die waagrechte Verbindung von A zum Rechteck der fuzzy Potenzmenge von B; es ist somit B* = (1/3, 1/4). Der minimale Abstand von A zu $F(2^B)$ berechnet sich folglich durch:

$$d(A,B^*) = \sum_{x \in X} |\mu_A(x) - \mu_{B^*}(x)| = |2/3 - 1/3| + |1/4 - 1/4| = 1/3$$

[79] Vgl. Kosko 1990 S 228ff

Womit für das Teilmengigkeitsmaß – analog zur Berechnung mittels der bisherigen Definition – folgt:

$$S(A,B) = 1 - \frac{d(A,B^*)}{M(A)} = 1 - \frac{1/3}{11/12} = 7/11$$

An dieser Stelle muss nun gezeigt werden, dass die obige Definition sinnvoll ist, das heißt, dass sie mit der bisherigen Definition des Teilmengigkeitsmaßes aus dem vorhergehenden Kapitel übereinstimmt.

Sei nun $B^* \in F(2^B)$ zu A am nächsten gemäß der l^1 Metrik. Dann gilt $\mu_A(x) \geq \mu_{B^*}(x)$ für alle $x \in X$ aufgrund der Orthogonalität im Hyperwürfel der fuzzy Potenzmenge $F(2^B)$. Wir unterscheiden nun zwei Fälle:

Fall 1: $\mu_A(x)=\mu_{B^*}(x)$

Dies gilt genau dann, wenn $\mu_A(x) \leq \mu_B(x)$ für alle x und A somit gleich B* ist. Dies impliziert $\max(0, \mu_A(x)-\mu_B(x)) = 0$ für alle $x \in X$.

Fall 2: $\mu_A(x)>\mu_{B^*}(x)$

Dies gilt genau dann, wenn $\mu_A(x)>\mu_B(x)$, da andernfalls B näher als B* wäre. Somit gilt: $\max(0, \mu_A(x)-\mu_B(x)) = \mu_A(x)-\mu_B(x)$.

Aus den beiden Fällen gemeinsam folgt, dass $\max(0, \mu_A(x)-\mu_B(x)) = |\mu_A(x)-\mu_{B^*}(x)|$. Summiert man diese Gleichung über alle $x \in X$, erhält man:

$$d(A,B^*) = \sum_{x \in X} |\,\mu_A(x) - \mu_{B^*}(x)\,| = \sum_{x \in X} \max(0, \mu_A(x) - \mu_B(x))$$

Aufgrund dieser Gleichung kann man die entsprechenden Ausdrücke in den beiden verschiednen Definitionen des Teilmengigkeitsmaßes S vertauschen, und ihre Äquivalenz wurde bewiesen.[80]

Für B*, der Teilmenge von B, die A am nächsten liegt, lässt sich noch eine weitere Charakterisierung finden: Da $\mu_B(x)=\mu_{B^*}(x)$ gilt, falls $\mu_A(x)>\mu_B(x)$, folgt $\min(\mu_A(x),\mu_B(x))=\mu_{B^*}(x)$ für alle x in X. Ferner gilt $\mu_A(x)=\mu_{B^*}(x)$, falls $\mu_A(x)<\mu_B(x)$, und somit $\min(\mu_A(x),\mu_B(x))=\mu_{B^*}(x)$ für alle x. Aufgrund der Definition des Durchschnitts mittels

[80] Vgl. Kosko 1990 S 231

der Minimumbildung, folgt B* = A $\cap$ B. Die Kommutativität des Durchschnitts impliziert ferner B* = A $\cap$ B = A*.

Aufgrund der Definition des Abstandes *d* gilt *d*(A,B) = M(A) – M(B), falls $\mu_A(x) \geq \mu_B(x)$ für alle x in X.[81] Da nun A $\cap$ B $\subseteq$ A (und daher $\mu_{A\cap B}(x) \leq \mu_A(x)$ für alle x) folgt, dass *d*(A,B*) = *d*(A, A$\cap$B) = M(A) – M(A$\cap$B). Setzt man dies in die obige Definition des Teilmengigkeitsmaßes ein, erhält man:

$$
\begin{aligned}
& S(A,B) = \\
& = 1 - \frac{d(A,B^*)}{M(A)} = \\
& = 1 - \frac{d(A,A\cap B)}{M(A)} = \\
& = 1 - \frac{M(A) - M(A\cap B)}{M(A)} = \\
& = 1 - 1 + \frac{M(A\cap B)}{M(A)} = \\
& = \frac{M(A\cap B)}{M(A)}
\end{aligned}
$$

Also insgesamt:

$$S(A,B) = \frac{M(A\cap B)}{M(A)}$$

Dieses wichtige Resultat soll nun in weiterer Folge Grundlage diverser Theoreme und Interpretationen sein, die zu wesentlichen Resultaten in dieser Arbeit führen werden.

4.3 Das Entropie-Teilmengigkeitstheorem

Aus der eben eingeführten Formel zur Berechnung der Teilmengigkeit S wird ersichtlich, dass diese rein als Funktion von Kardinalitäten M definiert werden kann; ebenso wurde das fuzzy Entropiemaß nur durch Kardinalitäten definiert. Dies motiviert die Suche nach einem Zusammenhang zwischen dem fuzzy Entropiemaß E und dem Teilmengigkeitsmaß S. Dieser Zusammenhang ergibt sich, wenn man den Ausdruck S(A$\cup$A^C,A$\cap$A^C) eingehender untersucht:

[81] Vgl. Kosko 1990 S 231

$$S(A \cup A^C, A \cap A^C) :=$$
$$:= \frac{M((A \cup A^C) \cap (A \cap A^C))}{M(A \cup A^C)} =$$
$$= \frac{M([(A \cup A^C) \cap A] \cap A^C)}{M(A \cup A^C)} =$$
$$\frac{M(A \cap A^C)}{M(A \cup A^C)} =:$$
$$=: E(A)$$

Dies ergibt somit das Entropie-Teilmengigkeitstheorem

$$S(A \cup A^C, A \cap A^C) = E(A),$$

welches den Zusammenhang beschreibt zwischen der fuzzy Entropie einer Menge, das den Grad der Unschärfe einer Menge quantifiziert, und der Teilmengigkeit von $A \cup A^C$ in $A \cap A^C$. Das Theorem besagt also, dass die Unschärfe einer Menge in direktem Verhältnis steht zu dem Grade, in welchem die Vereinigung $A \cup A^C$ in $A \cap A^C$ – einer ihrer eigenen Teilmengen – enthalten ist. Plakativ gesprochen gibt also der Grad, in welchem das Ganze in einem seiner Teile enthalten ist, die Unschärfe der Menge an.[82]

4.4 Teilmengigkeit und Wahrscheinlichkeit

Die im vorhergehenden Kapitel 4.2.1 hergeleitete Formel $S(A,B) = M(A \cap B)/M(A)$ für den Grad der Teilmengigkeit der Menge A in B stimmt mit der Definition der bedingten Wahrscheinlichkeit $P(B|A)$ überein[83]:

$$P(B \mid A) := \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} \cong \frac{M(A \cap B)}{M(A)} =: S(A,B)$$

Die bedingte Wahrscheinlichkeit $P(B|A)$ entspricht der Wahrscheinlichkeit P des Ereignisses B, falls A zutrifft – diese Interpretation ist jedem seit seiner Schulzeit an geläufig und gehört zu den Grundlagen der Wahrscheinlichkeitstheorie. Mittels der obigen Überlegungen gelingt nun eine neue Interpretation der bedingten Wahrscheinlichkeit. S(A,B) wurde eingeführt als der Grad des Enthaltenseins der Menge A in der Menge B. In der klassischen Mengenlehre gibt es nur die beiden Möglichkeiten $A \subseteq B$ oder $A \not\subset B$. Viele Mengen stimmen jedoch mit anderen nur teilweise überein, aber nicht ganz. Für Mengen, die fast gleich sind, aber nicht ganz, nutzen wir die fuzzy Theorie zur Modellierung dieses „fast gleich".

[82] Vgl. Kosko 1990 S 237f
[83] Vgl. Viertel 1997 S 12f

Bleiben wir für einen Moment in der klassischen Mengenlehre und betrachten die Mengen $A = \{2\}$ und $B = \{2,4,6\}$, dann gilt $A \subseteq B$ und somit $S(A,B) = 1$. Stellen wir uns nun den Wurf eines Würfels vor und betrachten zwei Ereignisse: Ereignis A entspreche der erscheinenden Augenzahl 2 nach dem Wurf, Ereignis B entspreche dem Ausgang, dass die gewürfelte Augenzahl gerade sei. Man schreibt hierfür $A = \{2\}$ und $B = \{2,4,6\}$. Wie groß ist nun die Wahrscheinlichkeit $P(B|A)$, also dass die gewürfelte Augenzahl gerade ist, falls sie 2 ist? Die Antwort ist selbstverständlich 1. Ist die Augenzahl 2, dann ist sie auch gerade. In diesem einfachen Fall stimmen der Grad der Teilmengigkeit und die bedingte Wahrscheinlichkeit überein. Die obige Argumentation impliziert ferner die Übereinstimmung der beiden Größen auch in allen anderen Fällen. Betrachten wir das umgekehrte Beispiel $P(A|B)$, also die Wahrscheinlichkeit, dass die Augenzahl 2 ist, falls sie zumindest gerade ist; sie berechnet sich mittels $P(A|B)=P(A\cap B)/P(B)=P(A)/P(B)=(1/6)/(1/2)=1/3$. Analog berechnet man die Teilmengigkeit von B in A, also $S(B,A)=M(A\cap B)/M(B)=M(A)/M(B)=1/3$. Wieder stimmen bedingte Wahrscheinlichkeit und Teilmengigkeit überein – wie es der obige Beweis auch impliziert.

Durch die obige Deduktion der Äquivalenz von Teilmengigkeit und bedingter Wahrscheinlichkeit gelingt also eine völlig neue Sichtweise und Interpretation der Wahrscheinlichkeit. Anstelle der abstrakten Interpretation „Wahrscheinlichkeit von B falls A eintritt“ gelingt die anschaulichere – weil geometrische – Interpretation von „Grad der Teilmengigkeit von B in A“. Da der Terminus „Wahrscheinlichkeit“ kein anschaulicher Begriff ist, wird intuitives Verständnis und Vorstellung in der Wahrscheinlichkeitstheorie sehr schwierig, da man nur mit abstrakten Größen rechnet. Dies mag sicherlich ein Vorteil der Interpretation durch Teilmengigkeit sein, da diese zumindest für einfache Beispiele eine gute Anschaulichkeit gewährleistet und sogar Skizzen ermöglicht:

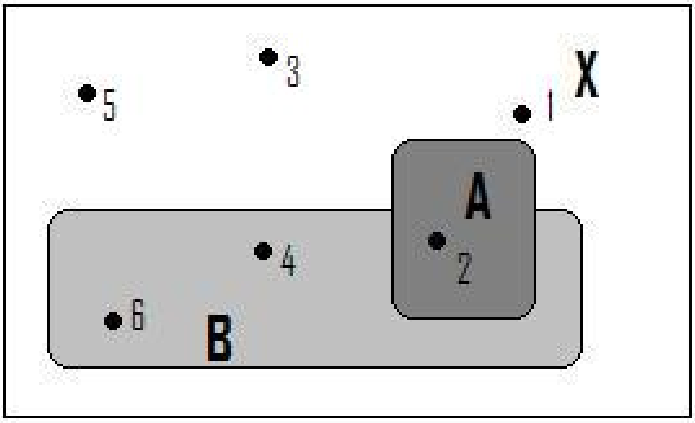

Abbildung 4

Nehmen wir nun an, A und B seien klassische, scharfe Mengen, A habe a Elemente und B sei gleich der Grundmenge $X = \{x_1, x_2, \dots, x_n\}$. Dann ergibt sich für S(A,B) die so genannte relative Häufigkeit:

$$S(A,B) = \frac{M(A \cap B)}{M(B)} = \frac{M(A)}{M(B)} = \frac{a}{n}$$

Mittels des Gesetzes der großen Zahl, das davon ausgeht, dass die relative Häufigkeit eines Ereignisses für große n gegen einen fixen Wert strebt, definiert man oft Wahrscheinlichkeit als idealisierte relative Häufigkeit[84].

4.4.1 Theoreme der fuzzy Theorie als Axiome der Wahrscheinlichkeitslehre

Es gelang uns nun wichtige Begriffe und Formeln der Wahrscheinlichkeitstheorie innerhalb der Theorie der unscharfen Mengen herzuleiten. Im Folgenden gehen wir noch weiter und deduzieren aus der fuzzy Theorie drei Theoreme, von denen Dennis Lindley behauptet, sie könnten als Axiome der Wahrscheinlichkeitslehre fungieren.

Wie bereits in Kapitel 4.2 ersichtlich wurde, gelten für alle fuzzy Mengen A,H folgende Relationen:

$$0 \le S(H,A) \le 1 \text{ und } S(H,A) = 1, \text{ falls } H \subseteq A$$

[84] Vgl. Viertel 1997 S 5

Aufgrund der Definition der Kardinalität M für unscharfe Menge ergibt sich auf evidente Weise folgende Additivitätseigenschaft: $M(A \cup B) = M(A)+M(B)-M(A \cap B)$. Mittels dieser Relation sowie der Definition des Teilmengigkeitsmaßes S und des Distributivgesetzes der Mengenlehre lassen sich nun zwei wichtige Gleichungen deduzieren:

$$
\begin{aligned}
&S(H, A_1 \cup A_2) = \\
&= \frac{M(H \cap (A_1 \cup A_2))}{M(H)} = \frac{M((H \cap A_1) \cup (H \cap A_2)}{M(H)} = \\
&= \frac{M(H \cap A_1) + M(H \cap A_2) - M(H \cap A_1 \cap H \cap A_2)}{M(H)} = \\
&= \frac{M(H \cap A_1)}{M(H)} + \frac{M(H \cap A_2)}{M(H)} - \frac{M(H \cap A_1 \cap A_2)}{M(H)} = \\
&= S(H, A_1) + S(H, A_2) - S(H, A_1 \cap A_2)
\end{aligned}
$$

Und weiters:

$$
\begin{aligned}
&S(H, A_1 \cap A_2) = \\
&= \frac{M(H \cap A_1 \cap A_2)}{M(H)} = \\
&= \frac{M(H \cap A_1 \cap A_2)}{M(H)} . \frac{M(H \cap A_1)}{M(H \cap A_1)} = \\
&= \frac{M(H \cap A_1)}{M(H)} . \frac{M(H \cap A_1 \cap A_2)}{M(H \cap A_1)} = \\
&= S(H, A_1) S(A_1 \cap H, A_2)
\end{aligned}
$$

Insgesamt konnten nun folgende 3 Beziehungen deduziert werden:

Konvexität: $0 \leq S(H,A) \leq 1$ und $S(H,A) = 1$, falls $H \subseteq A$

Addition: $S(H,A_1 \cup A_2) = S(H,A_1)+S(H,A_2)-S(H,A_1 \cap A_2)$

Multiplikation: $S(H,A_1 \cap A_2) = S(H,A_1)\ S(A_1 \cap H,A_2)$

Identifiziert man nun wieder S(H,A) mit P(A|H) ergeben sich folglich die drei Regeln (vgl. Kapitel 3.6), von denen Dennis Lindley postuliert:

> *"From these three rules, perhaps modiefied slightly, all of the many, rich and wonderful results of the probability calculus follow. They may be described as the axioms of probability."*[85]

[85] Lindley 1987 S 18

Dies ist ein beachtliches Resultat: Da innerhalb der fuzzy Theorie drei Theoreme abgeleitet werden konnten, mit denen man – wenn man sie als Axiome definiert – die Wahrscheinlichkeitslehre aufbauen kann, impliziert dies, dass die gesamte Wahrscheinlichkeitslehre als Folgerung der fuzzy Theorie betrieben werden könnte.

5 Zusammenhang zwischen Wahrscheinlichkeitslehre und fuzzy Theorie

Der Diskurs, ob die Wahrscheinlichkeitslehre ausreicht, sämtliche Arten der Ungewissheit adäquat zu modellieren, also in der Lage ist, als einzige Theorie der Ungewissheit zu fungieren, – wie es Befürworter der Wahrscheinlichkeitstheorie wie Dennis Lindley postulieren – wird oft mit Fragen erweitert, wie: Ist Ungewissheit dasselbe wie Zufall? Reichen unsere Begriffe der Wahrscheinlichkeit aus für unsere Vorstellungen von Ungewissheit?

Die Vermutung, Wahrscheinlichkeitslehre und fuzzy Theorie seien verschiedene Formulierungen desselben Problems, liegt nahe, wenn man die beiden Kalküle oberflächlich vergleicht und erkennt, dass beide im Wertebereich zwischen 0 und 1 operieren. Doch bereits konzeptionell erkennt man gewichtige Unterschiede in den beiden Theorien.

Die fuzzy Theorie beschreibt Vagheit; sie misst, zu welchem Grad ein Ereignis eintritt. Die Wahrscheinlichkeitslehre untersucht die Ungewissheit, ob ein Ereignis eintritt.[86] Dies sind zwei fundamental verschiedene Arten von Ungewissheit und es stellt sich bereits a priori die Frage, warum eine Theorie beide Arten von Ungewissheit adäquat modellieren können soll.

Wenn der Wetterbericht von einer 50%igen Chance spricht, dass es morgen regnet, ist es heute also noch ungewiss, wie das Wetter morgen sein wird. Vielleicht regnet es, vielleicht nicht – die Wahrscheinlichkeit für das eine wie das andere liegt bei 50%. Am nächsten Tag, wenn man aus dem Fenster sieht und das Wetter beobachtet, sieht man, ob es regnet oder nicht. Eine Wahrscheinlichkeitsaussage macht hier keinen Sinn mehr, weil es nicht mehr ungewiss ist, ob es regnet oder nicht. Die Wahrscheinlichkeit eines Regenschauers spiegelt also die Ungewissheit wider, ob dieses Ereignis in der Zukunft eintritt oder nicht. Anders verhält sich die Frage, ob die wenigen Tropfen, die zum Boden fallen, tatsächlich schon Regen genannt werden sollen oder nur Nieseln. Hierbei hilft es nicht, abzuwarten, bis das Nieseln bzw. der Regen vorbei ist, um die Frage zu klären, was das sei. Die Frage, ab wann Nieseln tatsächlich Regnen ist, ist eine Frage unabhängig von der tatsächlichen Realisierung desselben; es ist eine Frage, die auf Vagheit, einer speziellen Art von Ungewissheit, abzielt und nicht sinnvoll mittels Wahrscheinlichkeiten beantwortet werden kann.

[86] Vgl. Kosko 1990 S 213

5.1 Das Coxsche Theorem

Vertreter der Wahrscheinlichkeitstheorie wie der englische Statistiker und Bayesianer Dennis Lindley oder der Amerikaner Edwin Thompson Jaynes trennen den Wahrscheinlichkeitskalkül streng von jeglichem nicht klassischen Logiksystem wie insbesondere der fuzzy Logik. Jaynes war einer der ersten, der die Wahrscheinlichkeitslehre als Verallgemeinerung der aristotelischen Logik betrachtete.[87] Diese Vertreter sind es auch, die oftmals das Coxsche Theorem als Argument für die Wahrscheinlichkeitslehre verwenden; dieses wiederum wird oft falsch interpretiert bzw. aus diesem werden oft falsche Konsequenzen gezogen. Das Theorem von Cox besagt: Jedes Glaubensmaß ist isomorph zu einem Wahrscheinlichkeitsmaß.[88] Das heißt, dass jedes Glaubensmaß auf ein Wahrscheinlichkeitsmaß zurückgeführt werden kann. Dennis Lindley zieht daraus die Konsequenz:

> *"The message is essentially that only probabilistic descriptions of uncertainty are reasonable"*[89]

Doch die Annahmen der klassischen aristotelischen Logik sind teilweise problematisch, wenn man das Theorem von Cox als Argument gegen nicht klassische Systeme verwendet. Das Theorem zeigt nur, dass, falls Glauben derart definiert wird, wie es die Annahmen des Theorems vorgeben, die Wahrscheinlichkeitslehre ein legitimer Kalkül zur Behandlung von Glaubensgraden ist. Aber es wird nicht gezeigt, dass die Wahrscheinlichkeitslehre die einzig adäquate Methode ist, Ungewissheit zu formalisieren.[90]

5.2 Fuzzy Theorie als Extension der Wahrscheinlichkeitslehre?

Es gelang uns in Kapitel 4.4 innerhalb der fuzzy Theorie Beziehungen zu deduzieren, die als Axiomensystem eines Wahrscheinlichkeitskalküls dienen können. Die Wahrscheinlichkeitstheorie ist folglich innerhalb der fuzzy Theorie ableitbar. Von einem klassisch binären metatheoretischen Standpunkt betrachtet, kann dies zu zwei verschiedenen Konsequenzen führen: Erste mögliche Konsequenz wäre die logische Äquivalenz der fuzzy Theorie und des Wahrscheinlichkeitskalküls; dies ist der Fall, falls auch aus der Wahrscheinlichkeitslehre die fuzzy Theorie deduzierbar ist. Dann wären die beiden Theorien nur verschiedene Sichtweisen und Herangehensweisen, doch in ihrer Anwendbarkeit und

[87] Vgl. Bretthorst
[88] Vgl. Colyvan 2004 S 8
[89] Lindley 1982 S 1
[90] Vgl. Colyvan 2004 S 11

Aussagekraft gleichwertig. Die zweite Möglichkeit wäre – sollte die fuzzy Theorie nicht aus der Wahrscheinlichkeitslehre deduzierbar sowie die umgekehrte Deduktion jedoch möglich sein. – die fuzzy Theorie eine Extension der Wahrscheinlichkeitslehre wäre und diese als Anwendung besäße.

Wie bereits im Zuge dieser Arbeit ersichtlich gemacht wurde, gelten der Satz vom Widerspruch und das wohlbekannte Prinzip tertium non datur aus der klassischen Logik nicht in der fuzzy Theorie; das heißt: $A \wedge \neg A$ ist kein Widerspruch und $A \vee \neg A$ ist keine Tautologie in der fuzzy Logik – hingegen sind sie es in der klassischen. Diese Prinzipien fanden auch Einzug in die verschiedensten Theorien, die auf der klassischen Logik und Mengenlehre basieren – wie zum Beispiel in die Wahrscheinlichkeitslehre. Im Wahrscheinlichkeitskalkül gilt stets, dass ein Ereignis A eintritt oder nicht; eine dritte Möglichkeit ist ausgeschlossen, also $P(A \vee \neg A) = 1$. Ferner ist es unmöglich, dass ein Ereignis A eintritt und zugleich nicht eintritt, also formal $P(A \wedge \neg A) = 0$. Dies ist ein grundlegender Unterschied, der es prima vista unwahrscheinlich zu machen scheint, dass Wahrscheinlichkeitslehre und fuzzy Theorie insofern äquivalent sind, dass sich jedes Maß der fuzzy Theorie als Wahrscheinlichkeitsmaß schreiben lässt und umgekehrt. Vor Augen sollte man sich hierbei halten, dass die fuzzy Theorie als Extension der binären Logik das Gesetz vom ausgeschlossenen Dritten als Spezialfall beinhaltet – jedoch nur in Anwendung auf Aussagen mit Wahrheitswert 0 oder 1. Die Wahrscheinlichkeitslehre als auf der zweiwertigen Logik und Mengenlehre basierende Theorie ist hingegen nicht in der Lage das Prinzip vom ausgeschlossenen Dritten aufzugeben.

Dieser Gedankengang ist es auch, der folgenden Beweisversuch des fuzzy Theoretikers Bart Kosko zu Grunde liegen scheint.

Das Entropie-Teilmengigkeitstheorem aus Kapitel 4.3 impliziert, dass die Wahrscheinlichkeitstheorie nicht in der Lage ist, die fuzzy Theorie zu deduzieren. Um dies zu beweisen, zeigt Kosko, dass kein wahrscheinlichkeitstheoretisches Maß die Unschärfe einer Menge messen kann.

Angenommen das Gegenteil wäre der Fall und jedes Maß der fuzzy Theorie lässt sich durch ein wahrscheinlichkeitstheoretisches Maß ausdrücken. Wir betrachten nun das fuzzy Entropiemaß E, das für jede Menge seine Unschärfe angibt. Der obigen Annahme zufolge existiert nun ein Wahrscheinlichkeitsmaß P, das mit E übereinstimmt, also $P = E$. Da $P(X) = 1$ gilt, falls X die Grundmenge ist, kann P nicht identisch 0 sein. Somit existiert ein A, sodass $P(A) = E(A)>0$. Das Entropie-Teilmengigkeitstheorem impliziert nun

$$0<P(A) = E(A) = S(A\cup A^C, A\cap A^C),$$

wobei S das Teilmengigkeitsmaß ist, das den Grad der Teilmengigkeit von $A\cup A^C$ in $A\cap A^C$ angibt. Da hier versucht wird, aus dem Wahrscheinlichkeitskalkül die Theorie der unscharfen Mengen abzuleiten, muss die klassische Mengenlehre angewandt werden, weshalb die Vereinigung $A\cup A^C$ gleich ganz X ist und der Durchschnitt $A\cap A^C$ gleich der leeren Menge ist. Somit gilt:

$$0<P(A) = E(A) = S(A\cup A^C, A\cap A^C) = S(X,\{\})$$

Wenn $S(X,\{\})$ größer 0 sein soll, dann muss X in der leeren Menge enthalten sein. Dies ist aber nur der Fall, falls X selbst und somit auch $A \subseteq X$ leer ist. Somit gilt $1 = P(X) = P(\{\})$, also dass das unmögliche Ereignis sicher ist, was zu einem Widersprich führt.[91]

5.3 Wahrheitsfunktionalität

Eine weitere Beweisführung zur Demonstration der Nichtäquivalenz von fuzzy Theorie und Wahrscheinlichkeitslehre kann über den Begriff der Wahrheitsfunktionalität erfolgen. Ein logischer Junktor heißt wahrheitsfunktional (oder extensional), falls der Wahrheitswert eines durch ihn gebildeten Satzes nur vom Wahrheitswert seiner Teilsätze abhängt. Mit anderen Worten heißt dies, dass sich der Wahrheitswert einer durch einen Junktor zusammengesetzten Aussage zweier Teilsätze nur durch die Wahrheitswerte dieser beiden Teilsätze bestimmen lässt. Ein logischer Kalkül heißt wahrheitsfunktional, falls jeder Junktor wahrheitsfunktional ist.

Die fuzzy Logik ist wahrheitsfunktional,[92] wie sich leicht einsehen lässt, wenn man die Definition der fuzzy logischen Junktoren betrachtet:

Seien A,B fuzzy Aussagen und $|A|$ bzw. $|B|$ deren Wahrheitswerte, dann werden die Junktoren wie folgt berechnet:

$$\begin{aligned}
|\neg A| &= 1- |A| \\
|A\wedge B| &= \min(|A|,|B|) \\
|A\vee B| &= \max(|A|,|B|) \\
|A\rightarrow B| &= \max(\min(|A|,|B|), 1-|A|)
\end{aligned}$$

[91] Vgl. Kosko 1990 S 238
[92] Vgl. Luzzati S 123

Aus dieser Darstellung folgt auf evidente Weise, dass sich der Wahrheitswert jeder Komposition zweier fuzzy Aussagen mittels der Wahrheitswerte der Teilsätze bestimmen lässt.

Wahrheitsfunktionalität ist folglich eine Eigenschaft, die der fuzzy Theorie zukommt. Wären fuzzy Theorie und Wahrscheinlichkeitslehre äquivalent, müsste diese Eigenschaft auch für den Wahrscheinlichkeitskalkül gelten, das heißt die Wahrscheinlichkeitslehre müsste wahrheitsfunktional sein.

Doch dies ist nicht der Fall, falls die Ereignisse nicht unabhängig im Sinne der Wahrscheinlichkeitstheorie sind. Sei P ein Wahrscheinlichkeitsmaß, dann heißen zwei Ereignisse A und B unabhängig, falls $P(A \cap B) = P(A) \cdot P(B)$ gilt[93].

In diesem Falle lässt sich die Wahrscheinlichkeit des Durchschnittes als das Produkt der Wahrscheinlichkeiten der Teilereignisse bilden. Doch sind die Ereignisse nicht unabhängig, gilt dies nicht.

Betrachten wir nun die Axiome der Wahrscheinlichkeitsrechnung, wie sie im Kapitel 3.6 vorgestellt wurden. Das dritte Axiom lautet: $P(A \cup B) = P(A)+P(B)$, falls $A \cap B = \{\}$. Innerhalb des Wahrscheinlichkeitskalküls kann eine Additivitätseigenschaft auch für Ereignisse mit nichtleerem Durchschnitt deduziert werden; die Wahrscheinlichkeit der Vereinigung berechnet sich durch[94]:

$$P(A \cup B) = P(A)+P(A)-P(A \cap B)$$

Es wird also ersichtlich, dass die Vereinigung zweier Ereignisse nicht wahrheitsfunktional ist, falls die Ereignisse nicht unabhängig sind.

Ein Beispiel möge dies verdeutlichen: Sei A das Ereignis, dass bei einem Würfelwurf eine gerade Augenzahl erscheint und sei B das Ereignis, dass die zwei gewürfelt wird. Untersucht man nun die Wahrscheinlichkeit, dass eine gerade Augenzahl oder die zwei gewürfelt wird, berechnet sich diese wie folgt:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1/2 + 1/6 - 1/6 = 1/2,$$

wobei gilt: $P(A \cap B) = 1/6 \neq 1/12 = P(A) \cdot P(B)$.

[93] Vgl. Viertel 1997 S 15
[94] Vgl. Viertel 1997 S 12

Eine weitere Untersuchung verdeutlicht abermals das Nichtgelten der Wahrheitsfunktionalität des Wahrscheinlichkeitskalküls. Betrachten wir in Anlehnung an das berühmte Beispiel Jan Lukasiewicz' die Aussage G „Ich werde Weihnachten nächsten Jahres in Graz sein." Die Aussage habe aus nahe liegenden Gründen in der fuzzy Theorie den Wahrheitswert 0.5 und in der Wahrscheinlichkeitslehre die Wahrscheinlichkeit 0.5.

Aus der Wahrheitsfunktionalität eines Kalküls folgt insbesondere die Austauschbarkeit logisch äquivalenter Aussagen in zusammengesetzten Sätzen. Dies ist leicht einzusehen, wenn man bedenkt, dass sich der Wahrheitswert eines zusammengesetzten Satzes in extensionalen Systemen aus den Wahrheitswerten der Teilsätze ergibt; tauscht man also einen Teilsatz mit einem anderen aus, der jedoch denselben Wahrheitswert besitzt, kann dies an der Rechnung nichts ändern und der Wahrheitswert der gesamten Aussage bleibt konstant.

Beispielsweise kann in der fuzzy Logik – als wahrheitsfunktionales Kalkül – die Aussage G mit $\neg G$ vertauscht werden. Somit gilt:

$$|G \lor \neg G| = \max(0.5, 0.5) = 0.5$$

$$|G \lor G| = \max(0.5, 0.5) = 0.5$$

Die beiden Aussagen „Ich werde Weihnachten nächsten Jahres in Graz sein oder nicht." ist also innerhalb der fuzzy Logik äquivalent zur Aussage „Ich werde Weihnachten nächsten Jahres in Graz sein." Ob dies eine wünschenswerte Konsequenz ist, sei dahin gestellt – jedenfalls folgt diese aus der Wahrheitsfunktionalität und ist wesentliches Charakteristikum der fuzzy Logik.

Wäre die Wahrscheinlichkeitslehre auch wahrheitsfunktional, müsste dies also auch dort gelten. Betrachten wir also die Wahrscheinlichkeiten der Aussagen $G \lor \neg G$ bzw. $G \lor G$:

$$P(G \lor \neg G) = P(G)+P(\neg G)–P(G \land \neg G) = 1$$

$$P(G \lor G) = P(G)+P(G)-P(G) = P(G) = 0.5$$

Die Wahrscheinlichkeitstheorie ist also nicht wahrheitswertfunktional und kann daher nicht äquivalent sein zur fuzzy Logik oder anderen mehrwertigen wahrheitsfunktionalen Logiksystemen.[95]

[95] Vgl. Malinkowski 1993 S 327ff

6 Nachwort

Der philosophische Nutzen der fuzzy Theorie wurde an Hand der Analyse des Paradoxons des Sorites ersichtlich gemacht. Dennoch drängte sich der Verdacht auf, ob man nicht auf diese verzichten kann, indem man alles, was mittels der fuzzy Theorie ausgedrückt wird, mittels des Wahrscheinlichkeitskalküls modelliert. Vor allem Vertreter der Wahrscheinlichkeitslehre wie Dennis Lindley postulieren diese Möglichkeit fuzzy theoretische Ansätze mittels der Wahrscheinlichkeitsrechnung zu lösen. Dass dies – zumindest im allgemeinen Fall – nicht möglich ist, zeigte die abschließende Diskussion. Während wir innerhalb der fuzzy Theorie Theoreme deduzieren konnten, die als Axiome eines Wahrscheinlichkeitskalküls dienen können, konnten wir dies im umgekehrten Fall nicht tun. Darüber hinaus zeigten Überlegungen im Kapitel 5, dass es fuzzy theoretische Maße gibt, die durch kein Wahrscheinlichkeitsmaß ausgedrückt werden können. Ferner konnte gezeigt werden, dass zwar die fuzzy Theorie extensional, also wahrheitsfunktional ist – hingegen ist es nicht die Wahrscheinlichkeitslehre. Wahrscheinlichkeitslehre und fuzzy Theorie sind folglich nicht äquivalente Theorien, die nur verschiedene Interpretationen und Herangehensweisen an dieselbe Grundidee wären.

Dieser Schluss erscheint in Anbetracht unserer Analyse verschiedener Arten von Ungewissheiten als plausibel. Wir stellten fest, dass sich die Wahrscheinlichkeitslehre mit Ungewissheiten befasst, die von konzeptionell anderer Natur sind als die Ungewissheiten, mit der sich die fuzzy Theorie beschäftigt. Die Ungewissheit der Wahrscheinlichkeitsrechnung ist eine Ungewissheit, die sich auf die Unsicherheit zukünftiger Ereignisse bzw. die Unkenntnis gegenwärtiger oder vergangener Ereignisse bezieht. Hingegen rührt die Ungewissheit, die Gegenstand der fuzzy Theorie ist, von der Vagheit verwendeter Termini. In der fuzzy Theorie geht es also nicht darum, ob ein Ereignis eintritt oder nicht; in der fuzzy Theorie sind die Ereignisse, die sie behandelt, nicht scharf konzipiert sondern vage, was zu vollkommen anderen Problemstellungen führt als in der Wahrscheinlichkeitsrechnung.

7 Abbildungsverzeichnis

8 Literaturverzeichnis

ARISTOTELES: Metaphysik. In: Philosophie Schülerbibliothek. Directmedia Publishing GmbH. Berlin 1999 [Aristoteles Metaphysik]

AVENHAUS, Rudolf / SEISING Rudolf: Fuzzy Theorie – eine Alternative zur Stochastik? Eine Podiumsdiskussion. In: SEISING, Rudolf (Hrsg.): Fuzzy Theorie und Stochastik. Modelle und Anwendungen in der Diskussion. Vieweg. Wiesbaden 1999 [Avenhaus / Seising]

BOTHE, Hans-Heinrich: Fuzzy Logic. Einführung in Theorie und Anwendungen. Springer. Berlin/Heidelberg/New York 1995[2] [Bothe 1995]

BRETTHORST, Larry: Probability Theory As Extended Logic [Bretthorst]
unter: http://bayes.wustl.edu/ (Stand März 2007)

BULDT, Bernd: Supervaluvagefuzzysoritalhistorisch, oder: Ein kurzer Bericht der langen Geschichte, wie die Vagheit auf den Begriff und unter die Formel kam. In: SEISING, Rudolf (Hrsg.): Fuzzy Theorie und Stochastik. Modelle und Anwendungen in der Diskussion. Vieweg. Wiesbaden 1999 [Buldt]

COLYVAN, Mark: The Philosophical Significance of Cox's Theorem 2004 [Colyvan 2004]

DEISER, Oliver: Einführung in die Mengenlehre. Die Mengenlehre Georg Cantors und ihre Axiomatisierung durch Ernst Zermelo. Springer. Berlin 2002 [Dieser 2002]

FRAENKEL, Adolf: Zehn Vorlesungen über die Grundlegung der Mengenlehre. Teubner. Leipzig/Berlin 1927 [Fraenkel 1927]

FREGE, Gottlob: Grundgesetze der Arithmetik. Band II. Georg Olms Verlagsbuchhandlung. Hildesheim 1966 [Frege 1966]

GOTTWALD, Siegfried: Mehrwertige Logik. Eine Einführung in Theorie und Anwendungen. Akademie Verlag. Berlin 1989 [Gottwald 1989]

HAJEK, Alan: Probability, Logic, and Probability Logic In: GOBLE, Lou (Hrsg.): The Blackwell Guide to Philosophical Logic. Blackwell Publishing. Oxford 2001 [Hajek]

HASSE, Maria: Grundbegriffe der Mengenlehre und Logik. Teubner. Leipzig 1989[10] [Hasse 1989]

HOCHKIRCHEN, Thomas: Wahrscheinlichkeitsrechnung im frühen 20. Jahrhundert – Aspekte einer Erfolgsgeschichte. In: SEISING, Rudolf (Hrsg.): Fuzzy Theorie und Stochastik. Modelle und Anwendungen in der Diskussion. Vieweg. Wiesbaden 1999 [Hochkirchen]

JAANINEH, Georg / MAIJOHANN, Markus: Fuzzy-Logik und Fuzzy-Control. Vogel. Würzburg 1996 [Jaanineh / Maijohann 1996]

KOSKO, Bart: Fuzzy-logisch. Eine neue Art des Denkens. Econ. Düsseldorf 1995 [Kosko 1995]

KOSKO, Bart: Fuzziness vs. Probability. In: Int. J. General Systems. Vol. 17. Gordon and Breach Science Publishers S.A. 1990 S 211-240 [Kosko 1990]

KRUSE, Rudolf / GEBHARDT, Jörg / KLAWONN Frank: Fuzzy-Systeme. Teubner. Stuttgart 1993 [Kruse/Gebhardt/Klawonn 1993]

LINDLEY, Dennis: Scoring rules and the inevitability of probability. In: International Statistical Review 50 (1982) S 1–26 [Lindley 1982]

LINDLEY, Dennis: The Probability Approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Systems. In: Statistical Science, Vol.2, Nr.1 (Februar 1987) S 17-24 [Lindley 1987]

LUKASIEWICZ, Jan: Über den Satz des Widerspruchs bei Aristoteles. Georg Olms Verlag. Hildesheim 1993 [Lukasiewicz 1993]

LUZZATI, Claudio: Le metafore della vaghezza. In: Analisi e Diritto. Ricerche di giurisprudenza analitica. Giappichelli. Torino 1999 [Luzzati 1999]

MALINKOWSKI, Grzegorz: Many-Valued Logics (1993) In: GOBLE, Lou (Hrsg.): The Blackwell Guide to Philosophical Logic. Blackwell Publishing. Oxford 2001 [Malinkowski 1993]

MUKAIDONO, Masao: Fuzzy Logic. For Beginners. World Scientific Publishing. Singapore 2004[3] [Mukaidono 2004]

RADTKE, Peter: Sprache ist Denken [Radtke]
unter: www.peter-radtke.de/sprache.htm (Stand Jänner 2007)

RUSSELL, Bertrand: Vagueness. 1923 [Russell 1923]
unter: www.cscs.umich.edu/~crshalizi/Russell/vagueness/ (Stand: Jänner 2007)

SAINSBURY, Richard Mark: Paradoxien. Reclam. Stuttgart 1993 [Sainsbury 1993]

SINOWJEW, Alexander Alexandrowitsch: Über mehrwertige Logik. Ein Abriß. Friedrich Vierweg & Sohn. Basel 1968 [Sinowjew 1968]

VIERTEL, Reinhard: Einführung in die Stochastik. Mit Elementen der Bayes-Statistik und Ansätzen für die Analyse unscharfer Daten. Springer. Wien 1997 [Viertel 1997]

WANG, Pei: The Interpretation of Fuzziness. 1993 [Wang 1993]
unter: www.cogsci.indiana.edu/farg/peiwang/papers.html (Stand März 2007)

ZADEH, Lofti Asker: Fuzzy Sets. In: Information and Control 8. 1965 S 338-353 [Zadeh 1965]

ZADEH, Lofti Asker: The role of fuzzy logic in the management of uncertainty in expert systems. In: Fuzzy Sets and Systems, Vol. 11 1983 S 119-227 [Zadeh 1983]