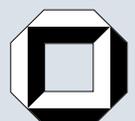
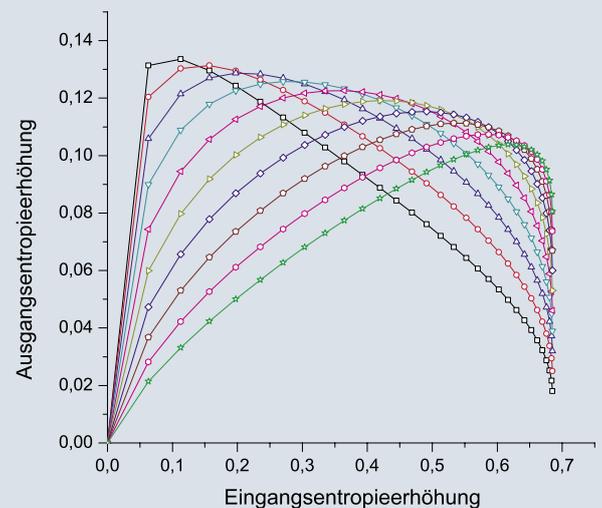
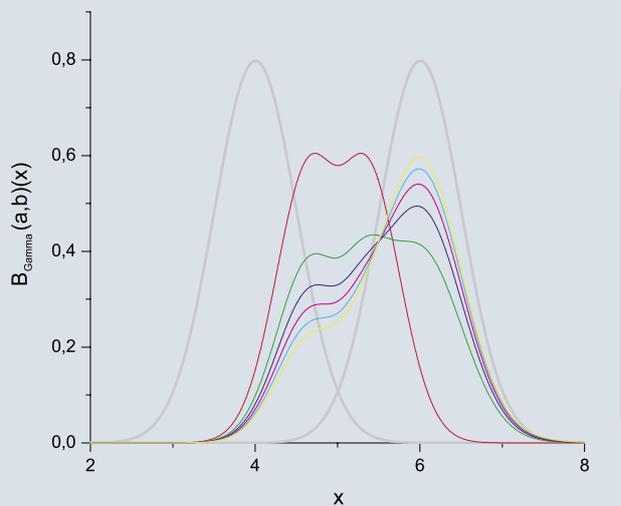


Florian Schaudel

Entropie- und Störungssensitivität als neues Kriterium zum Vergleich verschiedener Entscheidungskalküle



Florian Schaudel

**Entropie- und Störungssensitivität als neues Kriterium
zum Vergleich verschiedener Entscheidungskalküle**

Schriftenreihe des

Instituts für Angewandte Informatik / Automatisierungstechnik

an der Universität Karlsruhe (TH)

Band 11

Entropie- und Störungssensitivität als neues Kriterium zum Vergleich verschiedener Entscheidungskalküle

von
Florian Schaudel



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)
Fakultät für Maschinenbau, 2005

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de



Dieses Werk ist unter folgender Creative Commons-Lizenz
lizenziert: <http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Universitätsverlag Karlsruhe 2006
Print on Demand

ISSN 1614-5267
ISBN 3-86644-000-6

Entropie- und Störungssensitivität als neues Kriterium zum Vergleich verschiedener Entscheidungskalküle

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

der Fakultät für Maschinenbau der
Universität Karlsruhe (TH)
vorgelegte

Dissertation

von
Dipl. Math. Florian Schaudel
aus Langen

Hauptreferent: Prof. Dr. G. Bretthauer
Koreferent: Prof. Dr. H. Kiendl

Tag der Einreichung: 3. Mai 2005
Tag der mündlichen Prüfung: 15. Juli 2005

Vorwort

An dieser Stelle möchte ich mich bei all denen bedanken, ohne deren Unterstützung diese Arbeit nicht zustande gekommen wäre.

Mein ganz besonderer Dank gilt meinem wissenschaftlichen Betreuer und Hauptreferenten dieser Arbeit, Prof. Dr. Ing. habil. Georg Bretthauer, für seine fachliche Unterstützung sowie stetige konstruktive Kritik und wertvolle Ratschläge, mit denen er mich während der gesamten Entstehungszeit begleitete. Mit seiner ständigen Bereitschaft zur kritischen Diskussion hat er die Entwicklung dieser Arbeit maßgeblich vorangetrieben.

Genauso möchte ich Herrn Prof. Dr. rer. nat. Harro Kiendl für die Übernahme des Koreferats und sein Interesse an der Arbeit danken.

Wesentliche Teile dieser Arbeit entstanden während einer Freistellung durch meinen Arbeitgeber, McKinsey&Company, dem ich für die finanzielle und logistische Unterstützung während dieser Zeit dankbar bin.

Ein ganz besonderer Dank gilt schließlich auch meinen Eltern dafür, dass sie mir eine Ausbildung ermöglicht haben, die mir erlaubt, eine solche Arbeit zu schreiben und für ihre Unterstützung während der Arbeit. Nicht vergessen möchte ich natürlich auch, meiner Frau Christina für die stete moralische Unterstützung und Motivation sowie meiner Tochter Laura für ein immer strahlendes Lächeln zur Aufmunterung besonders zu danken.

Langen, im Mai 2005

Florian Schaudel

Inhaltsverzeichnis

1	Einleitung	1
1.1	Problemstellung	1
1.2	Entwicklungsstand	2
1.2.1	Entwicklung der Entscheidungstheorie	2
1.2.2	Modellierung von Information im Entscheidungskontext	8
1.2.3	Vergleich von Entscheidungskalkülen	10
1.3	Zielsetzung und Aufbau der Arbeit	13
1.3.1	Zielsetzung	13
1.3.2	Aufbau	14
1.3.3	Abgrenzung	15
2	Kalkülunabhängige Beschreibung von Entscheidungsmodellen	17
2.1	Klassische Matrixmodelle	17
2.1.1	Entscheidungen unter Sicherheit	17
2.1.2	Entscheidungen unter Unsicherheit	20
2.2	Neues kalkülunabhängiges Funktionalmodell	21
2.2.1	Allgemeines Entscheidungsproblem und dessen Eigenschaften	21
2.2.2	Entscheidungsmodell	24
2.2.3	Beispiel für die Anwendung des Funktionalmodells	28
2.3	Bestandteile des Funktionalmodells	33
2.3.1	Informationsfunktionen	34
2.3.2	Bewertungsoperatoren	39
2.3.3	Ordnungsfunktionale	46
2.4	Zusammenfassung	53
3	Neues Vergleichskriterium für Entscheidungsmodelle	55
3.1	Messung des Informationsgehaltes	56
3.1.1	Entropie	56
3.1.2	Entropie im Entscheidungskontext	58
3.1.3	Herausforderungen bei Verwendung des Entropiefunktional	59

3.2	Entropiesensitivität als neues Vergleichskriterium für Bewertungsoperatoren	61
3.3	Störungssensitivität als neues Vergleichskriterium für Ordnungsfunktionale	63
3.4	Zusammenfassung	65
4	Ergebnisse numerischer Vergleiche	67
4.1	Bewertungsoperatoren	67
4.1.1	Berechnungsmethode	68
4.1.2	Ergebnisse	70
4.1.3	„Anomalie“ des Gamma-Operators	77
4.1.4	Vergleich der Bewertungsoperatoren	80
4.2	Ordnungsfunktionale	81
4.2.1	Berechnungsmethode	81
4.2.2	Ergebnisse	84
4.2.3	Vergleich der Ordnungsfunktionale	90
5	Zusammenfassung und Ausblick	93
5.1	Allgemeine Erkenntnisse	93
5.2	Wichtigste Ergebnisse der Arbeit	96
5.3	Ausblick	98
	Literatur	99
6	Anhang	113
A	Bezeichnungen und Abkürzungen	113
B	Funktionsdefinitionen	115
C	Programme	117
D	Abbildungsverzeichnis	127
E	Verzeichnis der Definitionen	128

1. Einleitung

1.1. Problemstellung

Die Beschaffung von Information für Entscheidungsprozesse - sei es in Form von Sensorik für technische Anlagen oder für Experteneinschätzungen über wirtschaftliche Entwicklungen - kostet Geld. Darüber, ob die Information für den Nutzer auch ihr Geld wert ist, entscheidet nicht nur ihre intrinsische Qualität, sondern auch die Frage, wie effektiv sie in der Entscheidungsfindung überhaupt berücksichtigt wird.

Mit der klassischen Wahrscheinlichkeitstheorie (CDT¹) und der Fuzzy Decision Theory (FDT) sowie verschiedenen Varianten wie der Evidenztheorie, den subjektiven Wahrscheinlichkeiten etc. stehen zur Modellierung und Lösung von Entscheidungsproblemen leistungsfähige Kalküle zur Verfügung, die sich insbesondere darin unterscheiden, wie Informationen formalisiert und für die Bewertung unterschiedlicher Alternativen genutzt werden. Die Auswahl eines der Kalküle für ein gegebenes Problem erfolgt bei den meisten der in der Literatur dokumentierten Fälle eher intuitiv. Es scheint allerdings allgemein anerkannt zu sein, dass die klassische Wahrscheinlichkeitstheorie weniger für Entscheidungen unter großer Unsicherheit mit geringer Information geeignet ist, während es der FDT für einige kritische Anwendungen an der hinreichenden Präzision und "mathematischen Härte" fehlt (vgl. [162]).

In der vorliegenden Arbeit sollen diese Vorurteile quantitativ hinterfragt werden: Ausgehend von einer Sensitivitätsanalyse für Entscheidungsmodelle² wird untersucht, inwieweit sich die einzelnen Kalküle bezüglich der Nutzung vorhandener Information unterscheiden.

Dabei wird sich zeigen, dass es im Hinblick auf die Informationsnutzung zwar gravierende Unterschiede zwischen verschiedenen Entscheidungsmodellen gibt, sich diese aber nicht an den Grenzen der CDT und FDT fest machen lassen. Der subjektiv wahrgenommene Unterschied - insbesondere für Entscheidungen bei

¹nach dem englischen Begriff „Classical Decision Theory“, wobei sie wegen der wichtigen Rolle des Bayesschen Satzes gerade im angelsächsischen Sprachraum oft auch „bayesian decision theory“ genannt wird.

²wie sie z.B. Ott in [112] fordert

sehr geringer Information - kann dadurch erklärt werden, dass Anwender des Fuzzy-Kalküls gewohnheitsmäßig nicht so große Probleme damit haben, aus „Erfahrungswerten“ weit reichende Annahmen zu treffen, um das Problem zu vereinfachen. Während bei der Anwendung statistischer Methoden meist hunderte, wenn nicht tausende von Datensätzen erforderlich sind, um Aussagen über Wahrscheinlichkeitsverteilungen von Einflussgrößen machen zu können, wird die Fuzzy-Zugehörigkeitsfunktion, die bei unscharfen Methoden an die Stelle der Wahrscheinlichkeitsverteilung tritt, in vielen Fällen a priori als dreiecks- oder trapezförmig angenommen. Da aber auch in der klassischen Modellierung von Entscheidungsmodellen ein fehlerbehafteter Eingangswert meist als normalverteilt angenommen wird - auch wenn auf Grund physikalischer Beschränkungen nur wenige Messwerte wirklich normalverteilt sein können - und die beiden benötigten Parameter der Normalverteilung aus drei Datenpunkten geschätzt werden könnten, ist dies kein qualitativer Unterschied zwischen der CDT und der FDT.

Die Frage nach Vor- und Nachteilen der einen oder anderen Herangehensweise reduziert sich damit weitgehend auf die Vorlieben des Anwenders für Verwendung einer bestimmten Semantik bei der Beschreibung eines Entscheidungsmodells.

1.2. Entwicklungsstand

Um verschiedene Entscheidungskalküle miteinander vergleichen zu können, ist es zunächst notwendig, deren Entstehungsgeschichte zu betrachten, um ein Verständnis für deren grundsätzlich unterschiedliche Herangehensweisen an das Problem der optimalen Entscheidungsfindung zu entwickeln.

Anschließend wird ein kurzer Überblick über die bestehenden Arbeiten zum Vergleich verschiedener Entscheidungskalküle gegeben und dabei erläutert, wie sich die Methodik dieser Arbeit von den bestehenden Erkenntnissen abgrenzt. Schließlich werden mögliche Anwendungen der Ergebnisse diskutiert.

1.2.1. Entwicklung der Entscheidungstheorie

1.2.1.1. Klassische Entscheidungstheorie (CDT)

Die formale Beschäftigung mit der Entscheidungstheorie begann 1654 in einem Brief des Chevalier de Méré an Blaise Pascal[163]:

„Angenommen, mir werden zwei Spiele angeboten:

- 1. Ein Würfel wird 4-mal geworfen, wobei man darauf setzt, dass mindestens einmal eine 6 fällt.*
- 2. Zwei Würfel werden 24-mal geworfen, wobei man darauf setzt, dass mindestens ein Sechserpasch fällt.*

Gehe ich recht in der Annahme, dass ich bei Spiel 1 häufiger gewinne als bei Spiel 2?“

Aus dieser Frage entstand ein Briefwechsel zwischen Pascal und Fermat [116], in welchem schließlich nicht nur die Antwort gefunden wurde („ja“), sondern welcher auch als Ausgangspunkt für die noch heute vorherrschende klassische Entscheidungstheorie betrachtet werden kann: Der Erwartungswert eines mit Unsicherheit belegten Experimentes war als bester Indikator für dessen Ausgang identifiziert.

Nach grundlegenden Arbeiten von Moivre ([103]) und Laplace ([89]) war es schließlich Daniel Bernoulli [13], der im Rahmen seiner Beschäftigung mit dem Petersburger Paradoxon diejenige Frage formulierte, welche schließlich viel später für die Trennung von Wahrscheinlichkeits- und Entscheidungstheorie in zwei separate Zweige sorgen sollte:

Warum verhalten sich selbst vernünftige, mit den Grundlagen der Wahrscheinlichkeitsrechnung vertraute Menschen in einigen Situationen nicht so, dass sie ihren erwarteten Gewinn maximieren?

Da sich die mathematische Forschung aber zunächst dem theoretischen Fundament der Wahrscheinlichkeitsrechnung widmete, aus welchem dann die Maß- und Integrationstheorie entstand, dauerte es über 200 Jahre bis 1944 Neumann und Morgenstern in [107] Bernoullis Gedanken aufnahmen und darauf aufbauend die Theorie entwickelten, welche heute „klassische Entscheidungstheorie“ genannt wird. Ihr Kerngedanke ist weiterhin, dass eine „gute“ Entscheidung stets den Erwartungswert des Ergebnisses optimiert. Allerdings wird das Ergebnis nicht zwingend linear in Geld gemessen, sondern es wird berücksichtigt, dass bestimmte Ergebnisse unterschiedlich zu gewichten sind, wie einige triviale Beispiele verdeutlichen:

- Für die meisten Menschen ist es kaum erstrebenswert, bei einem Münzwurf 10 Mio. EUR auf Zahl zu setzen, selbst wenn sie im Erfolgsfall 25 Mio. EUR ausgezahlt bekommen, also einen erwarteten Gewinn von 2,5 Mio. EUR haben

- Für einen autonom navigierenden Roboter ist es meist wesentlich schädlicher wenn er versucht, 5 cm nach der Wand anzuhalten, als 10cm davor
- Bei den meisten medizinischen Diagnoseverfahren ist eine fälschlicherweise positive Diagnose einer falschen negativen vorzuziehen - zumindest wenn sie durch eine zweite Methode verifiziert werden kann.

Anwendungen dieses einfachen Prinzips finden sich heute in allen Bereichen der Ingenieur-, Natur-, Wirtschafts- und Sozialwissenschaften, daher existiert eine umfangreiche Literatur mit tausenden von Anwendungsbeispielen (z.B. [7, 90, 129, 136, 149]). Die klassische Entscheidungstheorie wird im angelsächsischen Sprachraum oft auch Bayessche Entscheidungstheorie genannt, was einige Autoren auf die Anwendung des Bayesschen Theorems zur Umrechnung der Messgeräteunsicherheit in die Messwertunsicherheit zurück führen.

1.2.1.2. Subjektive Wahrscheinlichkeiten

Das Konzept der subjektiven Wahrscheinlichkeiten (vgl. [124, 137]) ist ein erprobter Formalismus, um mit unterschiedlichem Vertrauen in verschiedene Informationsquellen umzugehen, welche sich im klassischen Sinne widersprechen. So kann z.B. die Aussage einer besorgten Mutter, über 50% der Autos fahren schneller als die erlaubten $50 \frac{\text{km}}{\text{h}}$, durchaus mit der Aussage eines Polizisten, in 10% aller Messungen sei eine überhöhte Geschwindigkeit festgestellt worden, in Einklang stehen.

Gerade in der psychologisch / soziologisch orientierten Literatur (vgl. [32, 80, 172]) werden derzeit sehr komplexe Modelle entwickelt, um von gegebener Information auf unterschiedliche Vertrauenswerte zu schließen. Für die Zwecke dieser Arbeit genügt es aber, anzunehmen, dass die Vertrauenswerte der verschiedenen Informationen bekannt sind. Die dann zu verwendenden Formalismen in den Entscheidungsmodellen unterscheiden sich bis auf Bewertungsfaktoren kaum von den Formalismen in der klassischen Entscheidungstheorie.

1.2.1.3. Fuzzy Decision Theory (FDT)

Die als Grundlage der Wahrscheinlichkeitstheorie dienende mathematische Logik ist, wie Zadeh in seinem die Fuzzy-Logik begründenden Aufsatz [175] bemerkt, nicht unbedingt die naheliegende Art und Weise, wie Menschen die Umwelt wahrnehmen. So sprechen wir von „großen Menschen“ ohne genau angeben zu können (und zu wollen), ob ein 1,85m langer Mann nun ein großer Mensch ist

oder nicht. Trotzdem hat jeder, der den Ausdruck „großer Mensch“ hört, intuitiv eine Vorstellung davon³. Zadeh schlug deswegen vor, die Zugehörigkeit eines Elementes zu einer Menge nicht binär (ein Element gehört entweder zu einer Menge oder nicht) zu beschreiben, sondern einen beliebigen Zugehörigkeitswert zwischen 0 und 1 zuzulassen. Dabei lassen sich nach einer möglichen Interpretation der Zugehörigkeitswerte ([42, 110]) scharfe und unscharfe Mengenzugehörigkeiten mit Mitteln der Wahrscheinlichkeitstheorie vereinen: Dieser Interpretation zufolge gibt der Wert der Zugehörigkeitsfunktion einer Fuzzy-Menge an einer bestimmten Stelle die Wahrscheinlichkeit an, mit der ein zufällig Befragter einen bestimmten Wert der entsprechenden scharfen Menge zuordnen wird. Gehört also ein Mann mit 1,85m Körperlänge mit einem Zugehörigkeitswert von 0.5 zur Fuzzy-Menge "große Menschen", so wird genau die Hälfte der Befragten einen solchen Mann als groß bezeichnen.

Zunächst 1970 von Bellman und Zadeh angeregt [11], dann von Bezdek [15] und Watson [160] Ende der Siebziger formalisiert und allgemeiner betrachtet, wurden die Methoden der Fuzzy-Logic bald auch gezielt auf entscheidungstheoretische Fragestellungen angewandt.

Gerade wegen ihrer Anschaulichkeit und der Verwendung natürlichsprachiger Ausdrücke ist die FDT heute in allen Anwendungsgebieten der Entscheidungstheorie sehr weit verbreitet (vgl. [18]):

Prozessmodellierung: Insbesondere bei komplexen Prozessen, die nicht analytisch durch einzelne Gleichungen zu beschreiben sind, findet die Fuzzy-Modellierung mit Sätzen einfacher Regeln häufig Anwendung (z.B. [68, 75, 79, 86]). Relativ modern sind dabei Methoden, bei denen die verwendeten Regeln nicht abstrakt aus einem tiefen theoretischen Verständnis des Systems sondern automatisch aus Beobachtungsdaten generiert werden (vgl. [25, 101, 144]), wobei das entstehende System nicht in jedem Fall das eigentlich zu regelnde System modelliert, sondern vielmehr auch den normalerweise menschlichen Bediener des Systems (vgl. [145]). In diesem Zusammenhang sei auf das von Kiendl, Krabs und Krone entwickelte Fuzzy-ROSA-Verfahren (vgl. [78, 84, 85]) hingewiesen, aber auch moderne Methoden zur automatischen Generierung der Regelsätze mittels neuronaler Netzwerke oder genetischer Algorithmen (vgl. [67, 72, 105, 114, 177]) kommen zur Anwendung. Als ein Vorteil der Anwendung automatisch generierter Fuzzy-Regelsätze erweist sich dabei, dass die derart generierten Regeln oft eine inhaltliche Interpretation erlauben und somit zum Verständnis des zu modellierenden Systems beitragen (vgl. [23, 100, 106]).

³wenn auch nicht notwendig dieselbe

Steuerungs- und Regelaufgaben (Fuzzy-Control)⁴: Die Regelungstechnik gehört nicht nur theoretisch zu den am besten beschriebenen Anwendungen des Fuzzy-Kalküls, sondern hier hat sich auch ein weites Spektrum von Anwendungen herausgebildet - angefangen mit der reinen Modellierung komplexer Prozesse (z.B. von Magnetlagern [161]) über die Regelung großtechnischer Anlagen (z.B. von Kraftwerksblöcken [3], Stahlwerken [45]) bis hin zur Steuerung von Hausgeräten (z.B. von Klimaanlage [12] oder Heißwasserboilern [173]). Der Grund der Verwendung von Fuzzy-Logik liegt häufig in der Nutzung vorhandener linguistischer Information, manchmal auch in der Nichtverfügbarkeit aussagekräftiger numerischer Zielgrößen (z.B. „angenehmes Raumklima“).

Mustererkennung: Ähnlich wie in der Prozessmodellierung treten auch in der Mustererkennung oft analytisch nur sehr komplex geschlossen zu beschreibende Strukturen auf. Deshalb werden erfolgreich Fuzzy-Ansätze eingesetzt (z.B. [4, 16, 30]). Einige neuere Arbeiten versuchen darüber hinaus, das in diesem Bereich vorherrschende klassische Paradigma der Support-Vector-Maschinen mit Fuzzy-Methoden zu kombinieren (z.B. [63]).

Optimierungsprobleme: Hierunter fallen die Lösung klassischer Optimierungsaufgaben vom Typ „Travelling-Salesman“, wie die Verteilung von Fabriken [153] oder Logistikzentren [48], und die Produktionsablaufplanung in der Automobilmontage [113]). Dieser Gruppe von NP-vollständigen Problemen ist gemeinsam, dass sie analytisch nicht exakt oder nur mit sehr großem Aufwand lösbar sind. Fuzzy-Logik bietet hier einen natürlichen Rahmen zur Suche nach „sehr guten“ Näherungslösungen

Entscheidungsunterstützungssysteme: Diese werden sowohl im technischen Bereich (z.B. bei der Auswahl einer geeigneten Stahlsorte für ein Werkzeug [98]) als auch im nicht technischen Bereich (z.B. bei der Prüfung der Kreditwürdigkeit eines Bankkunden [127] oder der Entscheidung über Investitionen [152]) eingesetzt. Hier liegt die Stärke der Fuzzy-Logik darin, dass oft unstrukturiert bzw. natürlichsprachig vorhandenes Expertenwissen unmittelbar genutzt werden kann (vgl. [17, 128, 131]).

Das theoretische Fundament der auf Fuzzy-Logik basierenden Entscheidungstheorie ist dagegen noch relativ schwach. Ansätze wie der von Ott in [112, S.100ff] scheitern - was von Ott selbst auch unmittelbar bemängelt wird - sowohl an der mangelnden Kanonisierung der FDT als auch an einigen Lücken im

⁴eine sehr umfassende, wenn auch nicht mehr ganz aktuelle Übersicht findet sich z.B. in [21]

theoretischen Unterbau im Bereich der Fuzzy-Maß- und Integrationstheorie. Die mangelnde Kanonisierung der FDT ist es auch, die einen exakten Vergleich zu klassischen Methoden erschwert: Während bei solchen klassischen Methoden das „Gesetz der großen Zahlen“ weitgehend die Bedeutung der auftretenden Wahrscheinlichkeiten und Wahrscheinlichkeitsverteilungen festlegt, ist eine Fuzzy-Zugehörigkeitsfunktion für sich genommen zunächst bedeutungsfrei. Schon in den Lehrbüchern zur Fuzzy-Theorie gibt es mindestens drei mögliche Interpretationen des Bedeutungsinhalts einer solchen Zugehörigkeitsfunktion (vgl. [42]), welche streng genommen jede für sich eine andere Modellierung erforderlich machen.

1.2.1.4. Evidenztheorie (DST)

Etwa gleichzeitig mit den Arbeiten von Zadeh und Bellman entwickelten Dempster und Shafer ([37], [141]) aufeinander aufbauend eine andere Alternative zur klassischen Entscheidungstheorie, die nach ihnen benannte „Dempster-Shafer-Theorie“ (DST). Ihr Hauptaugenmerk lag dabei auf der Behebung zweier ihrer wesentlichen Schwächen:

- Aus *theoretischer* Sicht sind die Anforderungen der Wahrscheinlichkeitstheorie sehr hoch. So ist es z.B. unmöglich, *jeder* Teilmenge von \mathbb{R} einen Inhalt und damit eine Wahrscheinlichkeit zuzuweisen [8, S. 53ff] ⁵.
- Aus *praktischer* Sicht fällt es oft schwer, in konsistenter Weise Wahrscheinlichkeiten und Nutzenfunktionen für Ereignisse anzugeben. Auch lassen sich bekannte Rahmeninformationen nicht immer leicht in Wahrscheinlichkeitsaussagen umformulieren.

Die der Theorie zugrunde liegende neue Idee ist, nun nicht mehr aus der gesammelten Information für jeden möglichen Ausgang eine Wahrscheinlichkeit und einen Nutzenwert zu berechnen, sondern für jedes Stück Information (Kohlas und Monney sprechen in diesem Zusammenhang von einem „Hinweis“ (vgl. [81])) einzeln zu überprüfen, was es darüber aussagt, ob der Ausgang gut oder schlecht ist. Aus der Kombination der einzelnen Hinweise ergeben sich für jede Aktion eine „obere“ und „untere“ Schranke dafür, dass sie tatsächlich die beste zur Verfügung stehende Aktion ist. In „mathematisch schönen“ Fällen konvergieren diese beiden Werte mit zunehmender Information gegeneinander, was die Interpretation als Wahrscheinlichkeiten gestattet.

⁵Originalarbeit von Hausdorff [59, S.401f]

Das grundsätzliche Konzept der DST wurde von unterschiedlichen Autoren zur Anwendungsreife weiter entwickelt. So lieferte Smets mit seinem "Transferable Belief Model" [146] einen axiomatischen Aufbau der DST, der ohne Rückgriffe auf Elemente der Wahrscheinlichkeitstheorie auskommt, während Dubois und Prade [44, 41] sowie Kohlas und Monney [81] von vornherein darauf achteten, die DST in solche einzubetten. Alle Formulierungen, die Smets in [147] anschaulich gegenübergestellt hat, fallen jedoch bei endlichen Ereignisräumen zusammen. Da sich die vorliegende Arbeit aber ausschließlich mit solchen Räumen beschäftigt, kann im Folgenden ohne Einschränkung der Allgemeinheit die Formulierung von Kohlas und Monney gewählt werden, die unter dem Namen „Mathematical Theory of Hints“ bekannt ist. Als Überbegriff für alle verschiedenen Formulierungen hat sich in der Zwischenzeit der Begriff „Evidenztheorie“ etabliert, der im Folgenden verwendet wird. Eine aktuelle Übersicht zur Verwendung der Evidenztheorie in Entscheidungssystemen findet sich in [117]-[122].

Die praktische Anwendung der Evidenztheorie ist derzeit noch recht selten vertreten (aber z.B. [108]) - nicht selten werden allerdings evidenztheoretische Überlegungen dem Design von klassischen oder FDT-Systemen vorangestellt (vgl. [24, 26, 55]). In der Beschreibung struktureller Zuverlässigkeit von Materialien und Bauteilen scheint sich aber in letzter Zeit ein fruchtbares Anwendungsfeld zu öffnen, wenngleich auch die veröffentlichten Arbeiten stets die nahe Verwandtschaft zur Fuzzy-Logik betonen (vgl. [36, 56, 102]). Allgemein sind die Grenzen zwischen Methoden der Evidenztheorie und Fuzzy-Logik in der praktischen Anwendung nur sehr schwer scharf zu ziehen (vgl. [43]).

1.2.2. Modellierung von Information im Entscheidungskontext

Alle vorgenannten Entscheidungskalküle dienen dem Zweck, Information über die Umwelt in einer Weise zu verknüpfen, die eine Bewertung möglicher Aktionen in einem gegebenen Kontext erlaubt. Damit das möglich ist, müssen die Informationen in strukturiert, mathematischer Form vorliegen. Die Information wird dabei in den verschiedenen Entscheidungskalkülen unterschiedlich dargestellt, sei es in Form von Wahrscheinlichkeitsdichten, Fuzzy-Zugehörigkeitsfunktionen, Möglichkeitsmaßen, oder anderen. Es ist sogar einer der wesentlichen Unterschiede der verschiedenen Kalküle, wie die Information formalisiert wird (vgl. [83]).

Um so erstaunlicher erscheint es zunächst, dass es keine eindeutige wissenschaftliche Definition davon gibt, was Information eigentlich ist (vgl. [47, 95,

132]). So finden sich (vgl. [14, 82]) alleine in der klassischen Literatur der Informationstheorie drei sich deutlich voneinander unterscheidende Ansatzpunkte:

Fisher Für Fisher [51] drückt sich Information in ihrer Wirkung auf die bedingte Wahrscheinlichkeit für ein Ereignis E aus: Ein Messwert M hat dann einen um so höheren Informationsgehalt, je stärker $P(E|M)$ von M abhängt.

Shannon und Hartley Shannon [142] vertritt die Ansicht, dass Information zunächst ein Fehlen von Ungewissheit ist: Werden zwei Situationen miteinander verglichen, so liegt in der Situation mehr Information vor, in der die Ungewissheit über den wahren Zustand am geringsten ist.

Auch wenn diese Definition zunächst tautologisch erscheint, erweist sie sich in vielen Situationen als äußerst hilfreich. Ist nämlich der Raum aller möglichen Umweltzustände beschränkt und wird gleichzeitig angenommen, dass eine Gleichverteilung über alle möglichen Zustände maximale Ungewissheit ausdrückt, so liefert die Shannon-Information ein absolutes Maß für den Informationsgehalt einer bestimmten Situation. Da beide Annahmen nicht zuletzt in der Quantenmechanik getroffen werden, hat dieser Informationsbegriff in der Physik weite Verbreitung gefunden (vgl. [47, 97]). Da darüber hinaus mit der Shannon-Entropie ein sehr praktisches Informationsmaß zur Verfügung steht, ist es auch in vielen anderen Bereichen wie beispielsweise der Messtechnik und der Signalverarbeitung das vorherrschende Informationsparadigma.

Kolmogorov Ein gänzlich anderer Ansatz findet sich bei Kolmogorov [82]: Der Informationsgehalt eines Objektes ist seine algorithmische Komplexität, also die Länge des kürzesten Programmes, das dieses Objekt auszugeben vermag.

Neben der Tatsache, dass auch diese Definition kein absolutes Maß für Information erzeugen kann, da die Komplexität stets von der dem Algorithmus zu Grunde liegenden Semantik abhängig ist, bedingen starke Beschränkungen bei der praktischen Umsetzung, dass diese Definition in realen Anwendungen kaum Verwendung findet (vgl. [54])

Für den Vergleich der Informationsnutzung verschiedener Entscheidungskalküle eignet sich der Shannon-Informationsbegriff von den drei vorgestellten am besten, da die Fisher-Information nur im klassischen Wahrscheinlichkeitskontext sinnvoll anwendbar bzw. der Aufwand, sie auf Fuzzy-Mengen zu übertragen, relativ groß (vgl. [52]) ist, und Kolmogorovs Informationsbegriff sehr theo-

retisch ist. Daher ist der Begriff „Information“ im Folgenden stets als Shannon-Information zu verstehen.

1.2.3. Vergleich von Entscheidungskalkülen

Die drei wichtigen Entscheidungskalküle CDT, FDT und DST wurden in der Literatur vielfach unter den verschiedensten Gesichtspunkten verglichen. Dabei werden insbesondere drei Kriterien⁶ immer wieder zum Vergleich herangezogen:

Komplexität: Wie hoch ist der Berechnungsaufwand und Speicherbedarf für nach den verschiedenen Kalkülen entworfene Entscheidungssysteme?

Expressivität: Inwiefern sind die einzelnen Theorien geeignet, verschiedene Probleme zu lösen bzw. unterschiedliche Informationen abzubilden?

Praxistauglichkeit: Wie lösen unterschiedliche Implementierungen eines Entscheidungssystems für einen vorgegebenen Zweck die gestellte Aufgabe?

1.2.3.1. Vergleich der Komplexität

Entscheidend für den praktischen Einsatz der Entscheidungskalküle sind neben der Qualität der Ergebnisse natürlich stets der notwendige Rechen- und Speicheraufwand und damit auch die Kosten der technischen Realisierung. Das gilt insbesondere, da automatische Entscheidungssysteme häufig bei sehr nichtlinearen Optimierungsaufgaben mit vielen Parametern und möglichen Ausgängen verwendet werden. Selbst moderne Computer kommen bei solchen Berechnungen schnell an die Grenzen ihrer Leistungsfähigkeit. Daher hat sich mit dem „Operations Research“ auch ein eigener Forschungszweig etabliert, welcher sich ausschließlich mit der Entwicklung von effizienten Algorithmen zur Berechnung solcher Optimierungsprobleme beschäftigt.

Aber auch bei kleineren Anwendungsbeispielen spielt die Berechnungskomplexität eine wichtige praktische Rolle: Da die Marktpreise für VLSI-Schaltungen in dem für Regelanwendungen relevanten Leistungsbereich annähernd linear mit der Rechenleistung⁷ skalieren, ist für den Hersteller jedes Prozent weniger Rechenaufwand bares Geld wert. In letzter Zeit kommt als weiterer Faktor der Stromverbrauch in mobilen Geräten hinzu, der sogar exponentiell mit der benötigten Rechenleistung steigt.

⁶auch ein Vergleich bezüglich „Stabilität“, also der Anfälligkeit gegen einzelne falsche Eingabeparameter, ist denkbar, aber bisher in der Literatur (noch) nicht zu finden

⁷bzw. Datendurchsatz, Speicherkapazität,...

Die Arbeit wird sich aus zwei Gründen nicht mit Fragen der Berechnungseffizienz befassen:

1. Die Komplexität ist sehr stark von Details des gewählten Modells abhängig. Die Unterschiede in der Komplexität zwischen einem sehr effektiven und einem weniger effektiven klassischen Modell für ein konkretes Entscheidungsproblem wird dadurch um Größenordnungen größer sein als der Unterschied zwischen einem sehr effektiven klassischen und einem sehr effektiven Fuzzy-Modell. Da im Rahmen der Arbeit die unterschiedlichen Kalküle anhand eines jeweils möglichst typischen Beispielmmodells verglichen werden, welches insbesondere nicht nach Berechnungseffizienz optimiert ausgewählt wird, kann ein Vergleich der Komplexität im gewählten Rahmen keine aussagekräftigen Ergebnisse liefern.
2. Für Modelle der CDT stehen heute bereits in viel größerem Umfang sehr leistungsfähige Standard-Algorithmen zur Verfügung als es für FDT- oder gar DST-Modelle der Fall ist. Ein fairer Vergleich der Kalküle kann also nicht auf Standard-Algorithmen aufbauen, sondern muss zumindest für die FDT- und DST-Modelle zunächst eigene, optimierte Algorithmen entwerfen, was den Aufwand für einen solchen Vergleich erheblich steigert. Die Diskussion um die Berechnungseffizienz der einzelnen Kalküle wird auch deswegen in der Literatur teilweise sehr emotional geführt (vgl [35]).

1.2.3.2. Vergleich der Expressivität

Thomas Whalen schlägt in seinem Übersichtsartikel „Decisionmaking under Uncertainty with Various Assumptions about Available Information“ [162] vor, die Auswahl eines Entscheidungskalküles für ein gegebenes Entscheidungsproblem von der Struktur der vorhandenen Information abhängig zu machen: Hat der Entscheider gesicherte Wahrscheinlichkeitsverteilungen für die relevanten Entscheidungsparameter zur Verfügung, so wird die CDT gewählt, liegen dagegen nur unzusammenhängende Informationen vor, so wird die DST gewählt, usw. Ein solcher Ansatz verkennt allerdings, dass die Informationen für den Entscheidungsprozess meist entweder unstrukturiert vorliegen oder extra für den Entscheidungsprozess erhoben werden. In beiden Fällen muss der Anwender ohnehin bestimmen, in welcher Form er die Information sammeln will.

J. Drakopoulos nimmt in [39] einen (mengen-)theoretischen Vergleich vor. Er zeigt, dass über endlichen Ereignisräumen - und nur solche sind letztendlich für eine technische Anwendung relevant - die CDT „ausdrucksstärker“ ist als die anderen beiden Kalküle. Hierbei ist allerdings zu beachten, dass die Aussage

nur über dem jeweils gleichen Ereignisraum gilt. Über einem entsprechend größer gewählten Raum lässt sich durchaus ein gleich ausdrucksstarkes System in jeweils der anderen Semantik formulieren. Da andererseits die Komplexität eines CDT-Systems bei gleichem Ereignisraum höher ist als bei FDT/DST-Systemen, kommt Drakopoulos ohne Beweis zu dem Schluss, dass die Wahl eines Entscheidungssystems letztendlich in Abwägung von Ausdrucksstärke und Komplexität erfolgen muss, wobei in der Abwägung keines der Systeme einen signifikanten Vor- oder Nachteil aufweist.

Noch abstrakter behandeln Zhen und YunJuan die Gemeinsamkeiten und Unterschiede von CDT und FDT, indem sie in [178] feststellen, dass beide Theorien komplementäre Ableitungen des Zermelo-Fraenkel-Systems seien, wobei die CDT ein „von-Neuman-Modell“ repräsentiere, während die FDT ein „Henkin-Modell“ darstelle.

Während die vorgenannten Autoren den Vergleich auf abstrakter Ebene führen, beschäftigt sich eine andere Gruppe damit, die unterschiedlichen Konzepte, die hier als verschiedene Semantiken zur Beschreibung des übergreifenden Konzeptes der Unsicherheit betrachtet werden, konstruktiv ineinander überzuleiten. (z.B. Wu et al für DST und CDT [174]). Jumarie schließlich [73] entwickelt sowohl die FDT als auch die DST von Grund auf neu, indem er sein Konzept der „Informations-Invarianz“ auf klassische informationstheoretische Transformationen anwendet.

1.2.3.3. Praktische Vergleiche

Praktische Vergleiche der verschiedenen Entscheidungskalküle sind vor allem aus der Regelungstechnik bekannt. Das ist insbesondere damit zu begründen, dass es in Regelungssystemen klare Kriterien und leicht nachprüfbare Zielwerte bezüglich der Qualität des Ergebnisses gibt (z.B. die maximale Abweichung der Ist- von der Soll-Temperatur). In anderen Anwendungsgebieten, z.B. Expertensystemen, ist es oft selbst a posteriori schwer, die Qualität einer Antwort zu bewerten.

Noch auf eher theoretischer Ebene behandeln einige Arbeiten die Äquivalenz von klassischen (PI)-Reglern mit Fuzzy-Reglern. So zeigt Byung Soo Moon in [104] zumindest für den Spezialfall des Controllers mit einem einzigen Eingang, dass und wie sich PI- und Fuzzy-Controller ineinander überführen lassen.

Interessanter für die Zwecke der vorliegenden Arbeit, da methodisch näher, sind die Vergleiche der unterschiedlichen Methoden in realen Anwendungen. Ribo und Pinz berichten in [125] über die Ergebnisse eines Versuches, bei welchem

sie unterschiedliche Ansätze zur Kartenerstellung in autonomen Robotern mittels Daten von Ultraschallsensoren verglichen. Aufgabe war es, aus mehreren tausend Distanzsignalen eines Ultraschallsensors, welche von mehreren Punkten des Raumes in unterschiedliche Blickrichtung ermittelt worden waren, eine Karte der von Hindernissen bedeckten bzw. freien Flächen zu generieren⁸. Dabei stellten sie empirisch fest, dass die Verwendung der CDT oder der DST bei relativ schwach gestörten Eingangsdaten (wenige Wandreflexionen) bessere Ergebnisse liefert, während die Implementierung der FDT den freien Raum zu konservativ einschätzt, also Hindernisse als dicker identifiziert als sie tatsächlich sind. Bei stark gestörten Daten mit vielen Reflexionen an den Wänden ist wiederum die Implementierung der FDT in dem Versuch die einzige der drei, welche zuverlässig alle Hindernisse erkennt. Ähnliche Vergleiche finden sich in [29, 53, 150].

1.3. Zielsetzung und Aufbau der Arbeit

1.3.1. Zielsetzung

Wie in Abschnitt 1.2.3 deutlich gemacht wurde, mangelt es derzeit in der Literatur an quantitativen und damit nachvollzieh- und übertragbaren Vergleichen der unterschiedlichen Entscheidungskalküle. Insbesondere der Aspekt der Informationsnutzung wurde bisher nicht systematisch untersucht, was auch daran liegt, dass die bestehenden Konzepte zur Sensitivitätsanalyse von einzelnen Bestandteilen eines Entscheidungsmodells kaum von einem in ein anderes Kalkül zu übertragen sind.

Ziel der vorliegenden Arbeit ist es daher, eine neue Vergleichsmethodik zu entwickeln und anzuwenden, welche

- es erlaubt, verschiedene Entscheidungskalküle in einem einheitlichen mathematischen Rahmen quantitativ gegenüber zu stellen,
- die Effizienz der Informationsausnutzung durch verschiedene konkrete Entscheidungsmodelle bewerten kann und damit Aussagen über die Anwendbarkeit und Wirtschaftlichkeit unterschiedlicher Modelle erlaubt,
- letztendlich dem praktisch tätigen Entwicklungsingenieur dabei hilft, das für seinen Anwendungsfall optimale Entscheidungskalkül und -modell zu

⁸zu diesem Zweck wurden zwei verschiedene Räume mit einer Fläche von 13 bzw. 25 qm in Planquadrate mit 10 cm Kantenlänge aufgerastert

wählen und hierbei insbesondere zur Klärung der Frage beizutragen, ob es in einer gegebenen Anwendung sinnvoll erscheint, weitere Ressourcen in die Verbesserung der Datenqualität oder der Verfeinerung des mathematischen Modells zu investieren.

Dabei wird Wert darauf gelegt, so weit wie möglich bekannte Konzepte aus der Informationstheorie zu übernehmen und den mathematischen Aufwand möglichst gering zu halten, um auch eine über die Arbeit hinausgehende praktische Anwendung zu ermöglichen.

Fragen der Berechnungseffizienz und des Ressourcenverbrauchs werden dagegen bewusst ausgelassen, da sie zum einen sehr stark von der konkreten Implementierung des Systems abhängen, welche nicht Gegenstand der Betrachtung sein soll, und zum anderen komplexe Optimierungsprobleme, bei denen die Berechnungseffizienz eine vorrangige Bedeutung spielt, nicht im Fokus der Arbeit stehen.

1.3.2. Aufbau

Zur Durchführung eines Vergleiches, welcher die oben genannten Anforderungen erfüllt, wird in der vorliegenden Arbeit in drei Schritten vorgegangen:

- Kapitel 2: *Entwurf eines neuen Meta-Modells (Funktionalmodell) zur kalkülübergreifenden Beschreibung von Entscheidungsmodellen.*
Zunächst wird aufbauend auf bestehenden Konzepten (vgl. z.B. [7, 90, 112]) ein Meta-Modell für Entscheidungssysteme entworfen, welches sowohl klassische als auch unscharfe Ausgestaltungen zulässt (Abschnitt 2.2). Hierzu ist es notwendig, die auch in der Literatur meist mit leicht unterschiedlicher Bedeutung verwandten Begriffe exakt zu definieren und das Entscheidungsmodell sauber vom zu lösenden Entscheidungsproblem zu trennen. Zur Abstraktion von einem konkreten Kalkül wird das Konzept der Informationsfunktion eingeführt, welches sowohl eine Fuzzy-Zugehörigkeitsfunktion als auch eine Wahrscheinlichkeitsdichte sein kann. Schließlich wird in Abschnitt 2.3 aufgezeigt, wie sich die zu untersuchenden Kalküle in dem neu entwickelten Meta-Modell darstellen.
- Kapitel 3: *Entwicklung einer neuen, auf der Informationsnutzung des Modells basierenden Vergleichsmethodik.*

Aufbauend auf dem Meta-Modell werden für dessen einzelne Bestandteile Methoden entwickelt, mit denen deren Informationsnutzung unabhängig vom gewählten Entscheidungskalkül quantitativ überprüft werden kann. Hierzu ist zunächst ein kurzer Exkurs in die Informationstheorie notwendig, um geeignete (Entropie-)Maße für die quantitative Erfassung der in einer Informationsfunktion codierten Information zu definieren (Abschnitt 3.1).

Anschließend werden in den Abschnitten 3.2 und 3.3 Störungsmodelle dargestellt, mit Hilfe derer die Sensitivität einzelner Komponenten der Entscheidungsmodelle auf Änderungen im Informationsgehalt der Eingangsfunktionen überprüft werden kann.

- Kapitel 4: *Anwendung der neuen Vergleichsmethodik auf typische Vertreter der verschiedenen Entscheidungskalküle.*

Schließlich wird die neue Methodik auf typische Vertreter der einzelnen Entscheidungskalküle angewandt, um sie miteinander zu vergleichen. Hierzu werden die einzelnen Modellbestandteile im CAS-System *Mathematica 5.0* implementiert⁹ und die unterschiedlichen Störungsmodelle darauf angewandt. Die numerischen Berechnungen werden durch Diskretisierung der auftretenden Funktionen vereinfacht.

Die Ergebnisse der Simulationsrechnungen finden sich in den Abschnitten 4.1.2 und 4.2.2.

In Kapitel 5 werden schließlich die Ergebnisse zusammen gefasst und ein Ausblick auf die Übertragung der entwickelten Konzepte auf die praktische Modellentwicklung gegeben.

1.3.3. Abgrenzung

Kernziel der Arbeit ist ein grundsätzlicher Vergleich von Entscheidungskalkülen anhand eines neu entwickelten, auf Entropie basierendem Qualitätskriterium, sowie die Etablierung eines neuen Meta-Modells zur Beschreibung von Entscheidungsmodellen, das solch einen kalkül-übergreifenden Vergleich ermöglicht. Ziel ist es nicht, einen umfassenden Überblick über alle in den jeweiligen Kalkülen möglichen Varianten zu geben, oder diese Varianten gegeneinander zu bewerten. Deswegen werden bei der Durchführung des Vergleiches nur für die

⁹Die verwandten Quelltexte finden sich im Abschnitt C.

jeweiligen Kalküle typische Vertreter berücksichtigt.

Darüber hinaus werden einige Anforderungen an die Struktur der betrachteten Probleme gestellt, welche „pathologische“ Spezialfälle ausschließen, sowie die Darstellung durch Vermeidung vieler Fallunterscheidungen übersichtlich halten. Zu den wichtigsten dieser Anforderungen gehören: beschränkter und kontinuierlicher Handlungsraum, stetige Bewertungsoperatoren, nicht mehr als zwei entscheidungsrelevante Umweltparameter.

Die hier entwickelte Methode kann aber auch auf Entscheidungsmodelle angewandt werden, bei denen eine der oben genannten Anforderungen nicht erfüllt ist. Allerdings muss in diesem Fall eine Reihe von Spezialfälle separat behandelt, bzw. das verwendete (integrale) Entropiemaß durch ein geeignetes diskretes ersetzt werden.

2. Kalkülunabhängige Beschreibung von Entscheidungsmodellen

Bei dem Versuch, mit bestehenden Methoden Entscheidungsmodelle aus verschiedenen Kalkülen miteinander zu vergleichen, gibt es zwei wesentliche Hindernisse. Zum Einen wird in der Literatur oft nicht sauber zwischen zu lösendem Entscheidungsproblem und dem eigentlichen Entscheidungsmodell unterschieden, zum Anderen verwenden CDT und FDT unterschiedliche Formalismen, um Informationen und Modelle zu beschreiben.

Um den mit dieser Arbeit bezweckten Vergleich verschiedener Kalküle vornehmen zu können, wird daher zunächst in Abschnitt 2.1 am Beispiel klassischer Matrixmodelle eine begrifflich saubere Trennung von Problem und Entscheidungsmodell vorgenommen, um dann in den Abschnitten 2.2 und 2.3 ein verallgemeinertes Funktionalmodell zu entwerfen, mit dem sowohl klassische, als auch unscharfe Entscheidungsmodelle formuliert werden können.

2.1. Klassische Matrixmodelle

2.1.1. Entscheidungen unter Sicherheit

Bei einem Entscheidungsproblem geht es darum, aus einer gegebenen Menge von Handlungsalternativen die „beste“ auszuwählen. Es wird bestimmt durch:

- Handlungsraum \mathbb{A} : Eine Menge sich gegenseitig ausschließender Handlungsmöglichkeiten.
- Umweltparameter U : Eine Menge von Parametern, deren Wert die Qualität der einzelnen Entscheidungen beeinflusst, ohne ihrerseits von der Entscheidung beeinflusst zu werden.
- Optimalitätskriterium O : Eine Regel, nach der ein Ereignis „besser“ ist als ein anderes. Ohne Einschränkung sei das Optimalitätskriterium numerisch messbar und somit der Grad der Optimalität einer Entscheidung mit \mathbb{R} parametrisierbar.

Die Forderung der Ausschließlichkeit der Handlungsalternativen stellt keine Einschränkung der Allgemeinheit von Entscheidungsproblemen dar. Können mehrere Handlungen durchgeführt werden, so wird als Raum der Handlungsalternativen entweder der Produktraum dieser Handlungen (also Alternativen der Form „Handlung A und Handlung B“) betrachtet oder zunächst eine Handlung ausgewählt und danach aus den verbleibenden Alternativen eine weitere.

Grundsätzlich wird zwischen Entscheidungsproblemen unter Sicherheit und solchen unter Unsicherheit unterschieden. Bei einem Entscheidungsproblem unter Sicherheit sind dabei die Werte der Umweltparameter zum Entscheidungszeitpunkt exakt bekannt oder werden zumindest als exakt bekannt angenommen.

Zur Lösung eines solchen Entscheidungsproblem es wird ein Entscheidungsmodell benötigt:

Definition 1 (Entscheidungsmodell unter Sicherheit)

Ein Entscheidungsmodell für eine Entscheidungssituation unter Sicherheit $(\mathbb{A}, \mathbb{U}, \mathbb{O})$ ist eine Funktion $f : \mathbb{A} \times \mathbb{U} \rightarrow \mathbb{O}$, so dass gilt:

$$f(\alpha_k, u_1, u_2, \dots) = f(\alpha_k) = \max \Leftrightarrow \alpha_k \text{ ist optimale Handlungsalternative aus } \mathbb{A}. \quad (2.1)$$

Oder in Worten: Ein Entscheidungsmodell zu einer gegebenen Entscheidungssituation (unter Sicherheit) ist eine numerische Funktion, die für alle Alternativen \mathbb{A} definiert ist, und die genau dann für ein α_k den maximalen Wert annimmt, wenn α_k eine optimale Handlungsalternative darstellt.

Da das Optimalitätskriterium oft nicht durch einen einzelnen, unmittelbar messbaren Parameter gegeben ist, sondern sich aus verschiedenen Größen zusammensetzt, wird die Optimalität einer Entscheidung nicht direkt betrachtet, sondern zunächst die messbare Folge der Entscheidung in einem sog. Ergebnisraum. Die Lösung eines Entscheidungsproblem es bei Sicherheit läuft dann in drei Stufen ab (vgl. Abbildung 2.1):

1. Ermittlung des Ergebnisses jeder einzelnen Aktion

Zunächst wird für jede mögliche Handlungsalternative das Ergebnis als Vektor in einem Ergebnisraum \mathbb{E} ermittelt.

2. Bewertung jedes Ergebnisses und damit jeder Aktion

Dann werden die Ergebnisse mittels einer Bewertungsfunktion v bewertet, insbesondere die verschiedenen Komponenten des Ergebnisvektors gegeneinander abgewogen. Bei eindimensionalen Ergebnisvektoren kann Schritt

Aktionen	Ergebnisse		Bewertung der Aktionen		Ordnung der Aktionen
α_1	$\vec{\kappa}_1 := \hat{f}(\alpha_1)$	→	$\nu(\alpha_1) := \hat{\nu}(\vec{\kappa}_1)$	→	4.
α_2	$\vec{\kappa}_2 := \hat{f}(\alpha_2)$	→	$\nu(\alpha_2) := \hat{\nu}(\vec{\kappa}_2)$	→	3.
α_3	$\vec{\kappa}_3 := \hat{f}(\alpha_3)$	→	$\nu(\alpha_3) := \hat{\nu}(\vec{\kappa}_3)$	→	1.
α_4	$\vec{\kappa}_4 := \hat{f}(\alpha_4)$	→	$\nu(\alpha_4) := \hat{\nu}(\vec{\kappa}_4)$	→	2.
...	...	→	...	→	...

ABBILDUNG 2.1: Klassisches dreistufiges Vorgehen bei Entscheidungsproblemen unter Sicherheit

2 meist entfallen und das Ergebnis unmittelbar zur Ordnung der Alternativen herangezogen werden.

3. Ordnung der Aktionen bzw. Auswählen der besten Aktion

Schließlich werden die Alternativen entsprechend ihrer Bewertung sortiert bzw. wird die beste Alternative ausgewählt.

Die Bewertung ν einer Aktion α ist also die Bewertung $\hat{\nu}$ des von dieser Aktion erzeugten Ergebnisses $\hat{f}(\alpha)$, welches mit Sicherheit vorhersagbar ist.

$$\nu(\alpha_k) =: \hat{\nu}(\hat{f}(\alpha_k)) \tag{2.2}$$

Vom theoretischen Blickwinkel aus betrachtet sind Entscheidungsprobleme unter Sicherheit - zumindest dann, wenn wie in allen praktisch relevanten Fällen nur endlich viele Handlungsalternativen in Frage kommen - einfach zu lösen, da „nur“ die Ergebnisse für alle Alternativen bewertet werden und die beste Alternative auszuwählen ist. Die einzige nennenswerte Schwierigkeit besteht in der Abwägung zwischen verschiedenen Dimensionen des Ergebnisraumes (also z.B. zwischen Preis und Haltbarkeit eines zu kaufenden Werkzeuges).

In der Praxis gibt es allerdings oft sehr viele - wenn auch nur endlich viele - mögliche Alternativen. So kommt es regelmäßig vor, dass die Berechnung der optimalen Lösung dennoch zu so großem Aufwand führen kann, dass sie nach derzeitigem Stand der Technik nur näherungsweise möglich ist. Als Beispiel hierfür kann das „Travelling Salesman“ Problem gelten (vgl. [22, 27, 155]), bei welchem

es darum geht, aus den $n!$ möglichen Rundreisen durch n verschiedene Städte diejenige mit der kürzesten Weglänge auszuwählen, und welches nicht nur in der Planung von Logistiknetzen, sondern auch z.B. bei der optimalen Verteilung von Sendemasten für Mobilfunknetze zu lösen ist. Schon bei nur 170 Städten/Standorten gibt es $170! > 7,2 * 10^{306}$ verschiedene Möglichkeiten.

2.1.2. Entscheidungen unter Unsicherheit

Bei der Entscheidung unter Unsicherheit sind nun die Werte der Umweltparameter zum Entscheidungszeitpunkt nicht mehr exakt bekannt, und die Unsicherheit wird bei der Modellierung explizit berücksichtigt. Hierzu müssen bei der Charakterisierung eines solchen Entscheidungsproblem es zwei zusätzliche Größen eingeführt werden:

Zustandsraum \mathbb{S} : Eine diskrete Menge endlich vieler sich gegenseitig ausschließender Umweltzustände

$$s_i := (u_1(\sigma_1), u_2(\sigma_i), \dots). \quad (2.3)$$

Informationen \mathbb{I} : Eine Menge von Informationen p_i , welche Auskunft über das Eintreten der Umweltzustände geben.

Solche Umweltzustände können bei der Entscheidung eines Bauers, wann er sein Kornfeld mähen soll, z.B. sein „Es regnet morgen“, „Es regnet in der nächsten Woche nicht, aber übernächste Woche“ oder „Es regnet erst nächstes Jahr wieder“. Da das Modell praktisch ausschließlich im Kontext der klassischen Wahrscheinlichkeitstheorie seine Anwendung fand, sind die Informationen p_i meist als Eintrittswahrscheinlichkeiten der einzelnen Zustände zu interpretieren.

Entscheidungsprobleme unter Unsicherheit werden in der Literatur [62, 94] analog zu Entscheidungsproblemen bei Sicherheit oft als Matrix dargestellt (vgl. Abbildung 2.2). Dabei wird ebenfalls prinzipiell dreistufig vorgegangen. Im Unterschied zu Entscheidungsproblemem bei Sicherheit (Abbildung 2.1) müssen allerdings mehrere (oder ein ganzes Spektrum) möglicher Ergebnisse der Handlung berücksichtigt werden. Auch wenn das Matrixmodell aufgrund seiner Anschaulichkeit zur Demonstration der Funktionsweise eines Entscheidungsmodelles sehr gut geeignet ist, so ist es für eine mathematische Analyse wegen seiner mehrschrittigen Struktur nachteilig.

Aktionen	Umweltzustände					→	Bewertung der	→	Ordnung der
	σ_1	σ_2	σ_3	σ_4	...		Aktionen		Aktionen
α_1	$\vec{\kappa}_{1,1}$	$\vec{\kappa}_{1,2}$...			→	$v(\alpha_1)$	→	2.
α_2	$\vec{\kappa}_{2,1}$	$\vec{\kappa}_{2,2}$				→	$v(\alpha_2)$	→	3.
α_3	\vdots		\ddots			→	$v(\alpha_3)$	→	4.
α_4						→	$v(\alpha_1)$	→	1.
...						→	...	→	...

ABBILDUNG 2.2: Dreistufiges Vorgehen bei Entscheidungsproblemen unter Unsicherheit.

Daher wird in den nächsten beiden Abschnitten ein Funktionalmodell eingeführt, das der Behandlung mit funktionalanalytischen Methoden zugänglich ist. In Abschnitt 2.2 wird das Modell formal definiert, wobei im Unterschied zu bestehenden ähnlichen Modellen das entwickelte Modell sowohl mathematisch formalisiert ist (im Gegensatz zu z.B. [34]) also auch zur Beschreibung von CDT und FDT Modellen gleichermaßen geeignet (im Gegensatz zu z.B. [112]).

2.2. Neues kalkülunabhängiges Funktionalmodell

2.2.1. Allgemeines Entscheidungsproblem und dessen Eigenschaften

Um Entscheidungsprobleme sinnvoll mit funktionalanalytischen Methoden behandeln zu können, ist der im Matrixmodell angenommene diskrete Zustandsraum zunächst hinderlich. Deswegen wird von einer allgemeineren Formulierung des Entscheidungsproblem auszugehen:

Definition 2 (Allgemeines Entscheidungsproblem)

Ein Tupel $(\mathbb{A}, \mathcal{U}, I, O)$ heißt „Entscheidungsproblem“, wenn

- \mathbb{A} eine Menge von sich gegenseitig ausschließenden Handlungsalternativen,

- $U \equiv \mathbb{R}^n$ ein endlich-dimensionaler \mathbb{R} -Vektorraum der Umweltparameter,
- $I = \{i_1, \dots, i_n\}$ eine Menge von Informationen über den Wert dieser Umweltparameter und
- O ein Optimalitätskriterium ist.

Zur Verbesserung der Lösbarkeit solcher Entscheidungsprobleme und um in allgemeinen Untersuchungen die Betrachtung einer Reihe pathologischer Fälle, welche in der Praxis selten auftreten, zu vermeiden, werden an die Struktur des Entscheidungsproblem eine Reihe weiterer Anforderungen gestellt, die in den nächsten drei Abschnitten kurz diskutiert werden.

2.2.1.1. Handlungsalternativen \mathbb{A}

Die Menge der Handlungsalternativen sollte zunächst nicht mächtiger als \mathbb{R} sein, so dass gleichzeitig eine Identifizierung der einzelnen Alternativen mit reellen Zahlen ermöglicht wird:

Definition 3 (Parametrisierbarer Handlungsraum)

Ein Handlungsraum \mathbb{A} heißt parametrisierbar genau dann, wenn es eine Variable α gibt, so dass alle möglichen Handlungsalternativen α_i eindeutig und umfassend durch $\alpha = x_i$ mit $x_i \in \mathcal{X} \subseteq \mathbb{R}$ beschrieben werden.
 α heißt dann Handlungsparameter.

Die Parametrisierbarkeit stellt dabei keine Einschränkung der Allgemeinheit für praktisch relevante Fälle dar, da jeder endliche Handlungsraum über eine beliebige Abzählung der Alternativen parametrisierbar ist. Erst zusammen mit bestimmten Regularitätsbedingungen für ein Entscheidungsmodell, wie z.B. Stetigkeit, können sich Einschränkungen ergeben.

Insbesondere aus Gründen der Berechenbarkeit ist es darüber hinaus vorteilhaft, wenn auch tatsächlich alle Handlungsalternativen aus einem bestimmten Intervall möglich sind.

Definition 4 (Vollständiger Handlungsraum)

Ein parametrisierbarer Handlungsraum mit Handlungsparameter α heißt vollständig, wenn es ein Intervall $[a; b] \in \mathcal{P}(\mathbb{R})$ gibt, so dass $\alpha = x$ für jedes maschinen-darstellbare $x \in [a; b]$ eine zulässige Handlungsalternative ist.

Insbesondere ist bei einem vollständigen Handlungsraum der Wertebereich des Handlungsparameters α beschränkt.

Im Folgenden werden nur vollständige Handlungsräume betrachtet. Die Forderung nach Vollständigkeit von Handlungsräumen erscheint zunächst als starke Einschränkung - im nicht technischen Bereich wird sie streng genommen so gut wie nie erfüllt sein. So können z.B. bei der Planung der Produktionsmenge immer nur ganze Werkstücke berücksichtigt werden wie man auch bei Kapitalanlageentscheidungen nur in bestimmten Stückelungen agieren kann. Es hat sich aber gezeigt, dass in den meisten Fällen eine Vernachlässigung dieser diskreten Struktur des Handlungsraumes zunächst gerechtfertigt ist, wenn dann in einem zweiten Schritt unter den tatsächlichen Alternativen diejenige ausgewählt wird, die am nächsten an der Lösung des vervollständigten Problems liegt. Für die spezifischen Probleme, welche bei der Lösung von Entscheidungsproblemen mit diskreten Handlungsräumen auftreten, sei auf die reichhaltige Literatur(z.B. [19, 50, 58]) verwiesen.

2.2.1.2. Umweltparameter

Der beim Matrix-Entscheidungsmodell unter Unsicherheit eingeführte Zustandsraum lässt sich auf das Funktionalmodell übertragen:

Definition 5 (Zustandsraum im Funktionalmodell)

Der von den Umweltparametern $U = \{u_1, u_2, \dots, u_n\}$ aufgespannte n -dimensionale \mathbb{R} -Vektorraum heißt „Zustandsraum“. Seine Elemente heißen „Umweltvektoren“. Die Informationen über die Werte dieser Umweltparameter werden mit $I = \{i_1, \dots, i_n\}$ bezeichnet.

Analog zum Handlungsraum kann auch vom Zustandsraum eine bestimmte Vollständigkeit gefordert werden:

Definition 6 (Vollständiger Zustandsraum)

Ein parametrisierbarer Zustandsraum mit Umweltvektor \vec{s} heißt vollständig, wenn es einen (Hyper-)Quader $Q \subset \mathbb{R}^n$ gibt , so dass $\vec{s} = x$ für jedes maschinen-darstellbare $x \in Q$ ein möglicher Umweltzustand ist.

Da die Forderung nach Vollständigkeit nur bedeutet, dass die Annahme eines solchen Umweltzustandes denkbar ist, auch wenn er in der Realität nicht

notwendig auftritt, bedeutet die ausschließliche Betrachtung vollständiger Zustandsräume keine Einschränkung der Allgemeinheit. In der Praxis werden unmögliche Zustände aus einem solchen vollständigen Zustandsraum durch eine entsprechende Information (also z.B. Eintrittswahrscheinlichkeit = 0) ausgeschlossen.

2.2.1.3. Information

Die Informationen I können prinzipiell auf vielfältige Art und Weise vorliegen: unstrukturiert, als linguistische Aussagen, in Form von Messwertreihen etc. Für deren mathematische Behandlung ist es allerdings sinnvoll zu fordern, dass sie die Form von Funktionen haben. Um das Modell offen für die verschiedensten Arten von Entscheidungskalkülen zu halten, wird im Folgenden der Begriff der "Informationsfunktion" verwendet:

Definition 7 (Informationsfunktion (IF))

Sei $B \subset \mathbb{R}$ ein abgeschlossenes Intervall. Eine Funktion

$$i : \mathbb{R} \rightarrow \mathbb{R} \tag{2.4}$$

heißt Informationsfunktion (IF) genau dann, wenn

1. $\forall x \notin B : i(x) = 0$,
2. $i(x) \geq 0$ für alle x ,
3. $i(x) < i$ für ein universelles i .

Der Raum aller IF über $B = [a; b]$, die durch i beschränkt sind, heiße $\mathcal{J}_{[a;b]}^i$.

Endliche Fuzzy-Zugehörigkeitsfunktionen (ZGF) und Wahrscheinlichkeitsdichten sind demnach spezielle Informationsfunktionen.

Von der Frage, wie aus unstrukturierten oder natürlichsprachigen Informationen solche Informationsfunktionen erzeugt werden können, handelt Abschnitt 2.3.1.

2.2.2. Entscheidungsmodell

Sind die Informationen in Form von Informationsfunktionen gegeben, so kann das Entscheidungsmodell für ein allgemeines Entscheidungsproblem als kanonische Erweiterung des Entscheidungsmodells unter Sicherheit (Definition 1)

definiert werden:

Definition 8 (Entscheidungsmodell als Funktional)

Sei $(\mathbb{A}, \mathbb{U}, \mathbb{I}, \mathbb{O})$ ein allgemeines Entscheidungsproblem mit $\mathbb{I} \in (\mathcal{J}_{[a;b]}^i)^n$ für geeignete a, b, i und $n = \dim(\mathbb{U})$, so ist ein Entscheidungsmodell \mathcal{E} für diese Entscheidungssituation ein Funktional

$$\mathcal{E} : (\mathbb{A}; (\mathcal{J}_{[a;b]}^i)^n) \rightarrow \mathbb{R} \tag{2.5}$$

mit

$$\mathcal{E}(\alpha, \mathbb{I}) = \max \Leftrightarrow \alpha \text{ ist optimale Handlungsalternative} \tag{2.6}$$

Mit anderen Worten: Ein Entscheidungsmodell ist ein Funktional, das - abhängig von der vorhandenen Information - jeder Handlungsalternative eine reelle Zahl in der Weise zuordnet, dass jede optimale Handlungsalternative den maximalen Wert erhält.

Zur mathematische Analyse eines solchen Entscheidungsmodells ist es zweckmäßig, das Funktional in zwei Teile zu zerlegen, die auch in der praktischen Modellierung meist getrennt voneinander betrachtet werden: einen *Bewertungsoperator* und ein *Ordnungsfunktional*. Sei zur Beschreibung $\mathcal{J} := (\mathcal{J}_{[a;b]}^i)$ für geeignete a, b und i , dann gilt:

Bewertungsoperator $B : (\mathbb{R}; \mathcal{J}^n) \rightarrow \mathbb{R} \times \mathcal{J}$: Der Bewertungsoperator ordnet jeder Alternative (also jedem möglichen Wert des Handlungsparameters α) eine „Nutzeninformationsfunktion“ zu, welche die Qualität der Aktion beschreibt. Diese Funktion kann zum Beispiel als Wahrscheinlichkeitsverteilung für den aus der Handlung entstehenden Nutzen oder als Fuzzy-Zugehörigkeitsfunktion für den Nutzen interpretiert werden.

Ordnungsfunktional $O : \mathcal{J} \rightarrow \mathbb{R}$: Das Ordnungsfunktional bewertet nun die verschiedenen Nutzenfunktionen, um sie miteinander vergleichen zu können, da es im Gegensatz zu reellwertigen Nutzen bei Nutzeninformationsfunktionen nicht offensichtlich ist, welche Handlungsmöglichkeit vorzuziehen ist. So zeigt zum Beispiel Abbildung 2.3 drei Fuzzy-Zugehörigkeitsfunktionen, eine dreiecksförmige, eine trapezförmige sowie eine mit zwei Spitzen (hier aus $I_{[800;1200]}^1$), bei denen nicht ohne weiteres zu entscheiden ist, welche die „bessere“ Alternative beschreibt, da alle denselben Schwerpunkt haben.

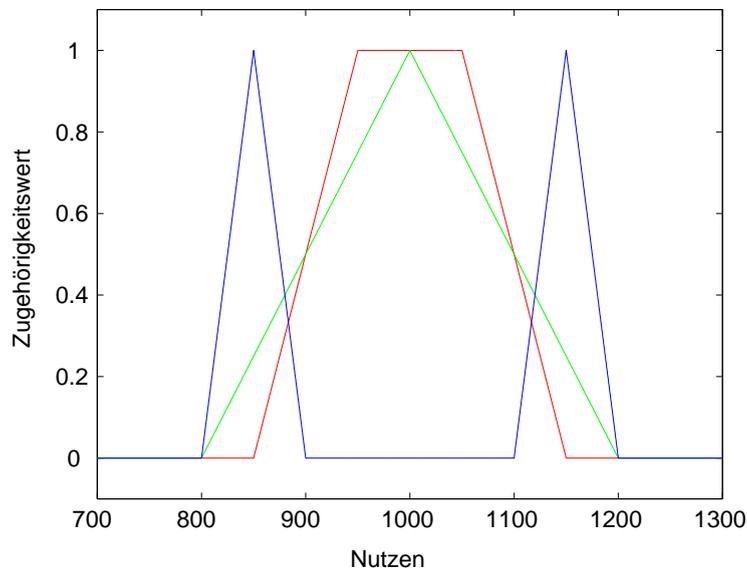


ABBILDUNG 2.3: Graph von drei verschiedenen Nutzenzugehörigkeitsfunktionen

Offensichtlich ist die Aufteilung eines Entscheidungsmodells in Ordnungsfunktional und Bewertungsoperator formal gesehen immer möglich, da zu einem gegebenen Entscheidungsmodell

$$\mathcal{E} : (\mathbb{A}; (\mathcal{J}_{[a;b]}^i)^n) \rightarrow \mathbb{R}; \quad \mathcal{E}(\alpha; i_1, \dots, i_n) = x_\alpha \quad (2.7)$$

lediglich

$$O : (\mathbb{A}; (\mathcal{J}_{[a;b]}^i)^n) \rightarrow I; \quad O(\alpha; i_1, \dots, i_n)(x) := \begin{cases} 1 & \text{für } x = x_\alpha \\ 0 & \text{sonst} \end{cases} \quad (2.8)$$

$$B : I \rightarrow \mathbb{R}; \quad B(i(x)) := \max_{x \in \mathbb{R}} x \mid i(x) = 1 \quad (2.9)$$

gesetzt werden muss.

Ein so konstruierter Bewertungsoperator verfehlt natürlich seinen eigentlichen Zweck, eine wirkliche Informationsfunktion über den Nutzen einer Aktion zu liefern.

Oft ergibt sich aber der Bewertungsoperator unmittelbar aus der Modellierung des Systems. In der Semantik des Funktionalmodells gesprochen stellen sich Methoden zur automatischen Regelgenerierung in Fuzzy-Systemen im beschriebenen Kalkül genau als Verfahren zur Ermittlung eines geeigneten Bewertungsoperators dar. Allerdings wird deutlich, dass die Formulierung eines komplexen Systems mehrerer Dutzend Fuzzy-Regeln in einem einzigen Bewertungsoperator zwar theoretisch möglich ist, für die praktische Arbeit aber eher weniger geeignet. Dies gilt insbesondere auch für die Zusammenfassung eines hierarchischen

Regelsystems, wie z.B. bei der Hyperdefuzzifizierung (vgl. [74]). Für die Analyse der Störungssensitivität allerdings, ist eine solche Zusammenfassung allerdings sehr hilfreich und da sie in jedem Fall möglich ist, ist das im Folgenden beschriebene Verfahren auch auf komplexe Modelle anwendbar.

Viele der in der Praxis vorkommenden Bewertungsoperatoren haben die günstige Eigenschaft, stetig bezüglich der Information zu sein. Das bedeutet, dass kleine Änderungen in der Information über die Umweltzustände auch die Bewertung jeder einzelnen Aktion nur wenig verändern. Ein bezüglich der Information stetiger Bewertungsoperator muss dagegen nicht notwendigerweise stetig bezüglich dem Ergebnis der Aktion sein. Daher können auch Bewertungsoperatoren in nichtlinearen Regelungssystemen in diesem Sinne stetig sein, da in diesen zwar eine kleine Änderung des Eingangsparameters zu großen Änderungen des Ergebnisses führen kann, nicht aber eine kleine Änderung der *Wahrscheinlichkeitsverteilung* dieses Eingangsparameters.

Lässt sich ein Entscheidungsmodell mit einem solcherart stetigem Bewertungsoperator darstellen, wird es quasistetig genannt.

Definition 9 (Quasistetiges Entscheidungsmodell)

Ein Entscheidungsmodell

$$\mathcal{E} : (\mathbb{R}; I^n) \rightarrow \mathbb{R}$$

heißt „*quasistetig*“, wenn es einen in der zweiten Variablen ($\in I^n$), also in jeder Komponente des Informationsvektors, stetigen Operator

$$B : (\mathbb{R}; I^n) \rightarrow \mathbb{R} \times I \tag{2.10}$$

und ein Funktional

$$O : I \rightarrow \mathbb{R} \tag{2.11}$$

gibt, so dass gilt

$$\mathcal{E}(\alpha; i_1, \dots, i_n) = O(i_\alpha) = O(B(\alpha; i_1, \dots, i_n)) \tag{2.12}$$

Analog heißt ein Bewertungsoperator „*quasistetig*“, wenn er Teil eines quasistetigen Entscheidungsmodells ist.

Alle in der vorliegenden Arbeit verwendeten Entscheidungsmodelle sind quasistetig. Daher können bei der numerischen Berechnung in Kapitel 4 konventionelle (Monte-Carlo-) Integrationsmethoden verwendet werden.

Zur Veranschaulichung des Funktionalmodells ist im folgenden Abschnitt seine Anwendung an einem Beispiel demonstriert. Da das Beispiel nur der Veranschaulichung der Vorgehensweise dient, entspricht es keiner realen Anwendung

sondern ist für die Zwecke dieser Demonstration konstruiert.

Dabei zeigt sich, dass die Trennung des Funktionalmodells in Bewertungsoperator und Ordnungsfunktional dem natürlichen Vorgehen beim Entwurf eines Entscheidungsmodells entgegen kommt.

2.2.3. Beispiel für die Anwendung des Funktionalmodells

Beispiel 1 (Farbabfüllung)

Ein Produzent für 1000g-Eimer Dispersionsfarbe automatisiert seine Abfüllanlage. Wo bisher noch ein Angestellter mit der Hand zum Befüllen jedes Eimers einen Hebel auf- und wieder zudrehen musste, soll das in Zukunft ein automatisches Ventil erledigen. Bei der Modellierung des Reglers mit Hilfe des Expertenwissens des Mitarbeiters stellt sich heraus, dass der Mitarbeiter ein hervorragendes Gefühl für die notwendige Öffnungszeit des Ventils entwickelt hatte. Auch im automatischen System ist eine Zeitsteuerung vorgesehen, da eine kontinuierliche Kontrolle des Eimerfüllstandes technisch nicht möglich und eine Massedurchflussmessung im Rohr zu teuer ist. Der Experte berichtet, dass zwei Parameter die notwendige Öffnungszeit des Ventils entscheidend beeinflussen:

- Die aktuelle Temperatur der Farbe als bestimmende Variable für die Viskosität und
- der von Charge zu Charge variierende Anteil Luft, der in kleinen Blasen eingeschlossen ist.

Während die Temperatur leicht messbar ist, ist bei der Abschätzung des Luftanteils weiterhin die Beurteilung durch den Mitarbeiter notwendig, welche er für jede Charge in das Regelungssystem eingibt.

Das für die Anwendung zu modellierende Entscheidungsmodell $\mathcal{E}(x, i_T, i_{NL})$ muss also unter Berücksichtigung der vorhandenen Informationen über die Umgebungstemperatur i_T und den Nicht-Luftanteil der Farbe i_{NL} jeder zulässigen Öffnungszeit x einen Qualitätswert in der Weise zuordnen, dass

$$\mathcal{E}(x; i_T, i_{NL}) = \max \Leftrightarrow (\text{Füllmenge nach } x \text{ Sekunden} = 1000\text{g}). \quad (2.13)$$

Die Aufteilung des Entscheidungsmodells in einen Bewertungsoperator und ein Ordnungsfunktional ist dabei inhaltlich naheliegend. Der Bewertungsoperator ordnet jeder Öffnungszeit eine Informationsfunktion über die Abweichung der Füllmenge von der Idealmenge 1000 g zu.

Wird der Einfachheit halber angenommen, die Massenflussgeschwindigkeit hängt linear von der Umgebungstemperatur ab, so kann sich die Abweichung z.B. berechnen als¹

$$\begin{aligned} A(x) &:= - \left| 1000\text{g} - \left(x * \left(100 \frac{\text{g}}{\text{s}} * i_{\text{NL}} * \frac{i_{\text{T}}}{20^\circ\text{C}} \right) \right) \right| \\ &= - \left| 1000\text{g} - \left(5 \frac{\text{g}}{\text{s}^\circ\text{C}} x * (i_{\text{NL}} * i_{\text{T}}) \right) \right|, \end{aligned} \quad (2.14)$$

wobei $100 \frac{\text{g}}{\text{s}}$ die Massenflussgeschwindigkeit der Farbe durch die Abfüllanlage bei 20°C ohne Luftblasen ist, $0 < i_{\text{NL}} \leq 1$ der Nichtluftanteil und $\frac{i_{\text{T}}}{20^\circ\text{C}}$ der Korrekturfaktor für die Berücksichtigung der geänderten Viskosität bei einer Temperatur $i_{\text{T}} \neq 20^\circ\text{C}$.

Sind i_{T} und i_{NL} keine Funktionen, sondern - wie bei einem Entscheidungsproblem unter Sicherheit - feste Zahlen, so lässt sich durch diese Gleichung zu jeder Öffnungszeit x die Abweichung unmittelbar angeben. Ist zum Beispiel $i_{\text{T}} = 24^\circ\text{C}$ und $i_{\text{NL}} = 0.9$, so ergibt sich die Abweichung in Abhängigkeit der Einfüllzeit x als

$$A_{0.9;24^\circ\text{C}}(x) = - \left| 1000\text{g} - \left(x * 100 \frac{\text{g}}{\text{s}} * 0.9 * 1.2 \right) \right| \quad (2.15)$$

$$= - \left| 1000\text{g} - \left(108 \frac{\text{g}}{\text{s}} * x \right) \right| \quad (2.16)$$

Die minimale Abweichung $A_{0.9;24^\circ\text{C}}(x) = 0$ wird bei $x = 9.26\text{s}$ erreicht.

Seien nun $i_{\text{T}} = i_{\text{T}}(y)$ und $i_{\text{NL}} = i_{\text{NL}}(y)$ Informationsfunktionen, in diesem Beispiel Wahrscheinlichkeitsdichten, über den Wert von i_{T} und i_{NL} , also $i_{\text{NL}}(y) = P(i_{\text{NL}} = y)$. Dann ist eine Möglichkeit den Bewertungsoperator zu bilden, die Wahrscheinlichkeitsverteilung für die Abweichung der Füllmenge $A(x)$ bei Öffnungszeit x zu berechnen.

Bei gegebener Füllzeit x_0 berechnet sich die Wahrscheinlichkeit $P(A(x_0) = A_0)$ für eine gegebene negative Abweichung A_0 als:

$$P(A(x_0) = A_0) = - \left| 1000\text{g} - \left(5 \frac{\text{g}}{\text{s}^\circ\text{C}} x * (i_{\text{NL}} * i_{\text{T}}) \right) \right| = A_0 \quad (2.17)$$

$$\begin{aligned} &= P(1000\text{g} - (5x * (i_{\text{NL}} * i_{\text{T}})) = A_0) + \\ &\quad + P(1000\text{g} - (5x * (i_{\text{NL}} * i_{\text{T}})) = -A_0) \end{aligned} \quad (2.18)$$

$$= P\left(i_{\text{NL}} * i_{\text{T}} = \frac{1000 - A_0}{5x}\right) + P\left(i_{\text{NL}} * i_{\text{T}} = \frac{1000 + A_0}{5x}\right) \quad (2.19)$$

¹Das Modell wurde lediglich zur Veranschaulichung entworfen und entspricht keiner realen Anwendung

Sind i_{NL} und i_T kontinuierlich verteilte Zufallsgrößen so errechnet sich die Wahrscheinlichkeitsverteilung für die Abweichung bei gegebener Öffnungszeit x daher als Faltung der Verteilungen von i_{NL} und i_T :

$$\begin{aligned} \mu_x(y) = & \int \frac{1}{5x} \left| \frac{1}{z} \right| i_T(z) * i_{NL} \left(\frac{1000-y}{z} \right) dz + \\ & + \int \frac{1}{5x} \left| \frac{1}{z} \right| i_T(z) * i_{NL} \left(\frac{1000+y}{z} \right) dz \end{aligned} \quad (2.20)$$

$$=: B(x, y). \quad (2.21)$$

$B(x_0, y_0)$ ist somit die Wahrscheinlichkeit, dass bei einer Öffnungszeit von x_0 Sekunden die Füllmengenabweichung y_0 erreicht wird.

Angenommen, an einem bestimmten Tag ist die Temperatur nun gleichverteilt zwischen 22°C und 26°C und der Nicht-Luftanteil in der Farbe normalverteilt um 0.9 mit $\sigma = 0.02$, also

$$i_T(z) := \begin{cases} 0.25 & \text{für } 22 \leq z \leq 26 \\ 0 & \text{sonst} \end{cases} \quad (2.22)$$

$$i_{NL}(z) := \frac{1}{0.02 * \sqrt{2} * \pi} * e^{-\frac{1}{2} * \left(\frac{z-0.9}{0.02} \right)^2}. \quad (2.23)$$

$$(2.24)$$

Damit vereinfacht sich (2.20) zu:

$$\begin{aligned} B(x, y) = & \int_{22}^{26} \frac{1}{5x} \left| \frac{1}{z} \right| * 0.25 * i_{NL} \left(\frac{1000-y}{z} \right) dz + \\ & + \int_{22}^{26} \frac{1}{5x} \left| \frac{1}{z} \right| * 0.25 * i_{NL} \left(\frac{1000+y}{z} \right) dz. \end{aligned} \quad (2.25)$$

Aus dieser Gleichung kann $B(x, y)$ numerisch berechnet werden (vgl. Abbildung 2.4; die eingezeichneten Linien in dieser Abbildung sind die Funktionsgraphen der Informationsfunktion über die Qualität der Füllung, also die Wahrscheinlichkeitsdichten der Qualität, bei gegebener Öffnungszeit und werden in Abbildung 2.5 detaillierter dargestellt). Zur Ermittlung der optimalen Einfüllzeit müssen nun die Funktionen $B|_x(y)$ betrachtet, also die Schar der eindimensionalen Funktionen, die sich beim Festhalten von jeweils einem x ergeben (vgl. Abbildung 2.5), die für jede Öffnungszeit x die Wahrscheinlichkeitsverteilung für den Qualitätswert y darstellt. Um nun das beste x auswählen zu können, müssen die Qualitätsfunktionen auf eine Zahl reduziert werden, wozu das Ordnungsfunktional angewendet wird. Da hier sehr viele Eimer zu füllen sind und das

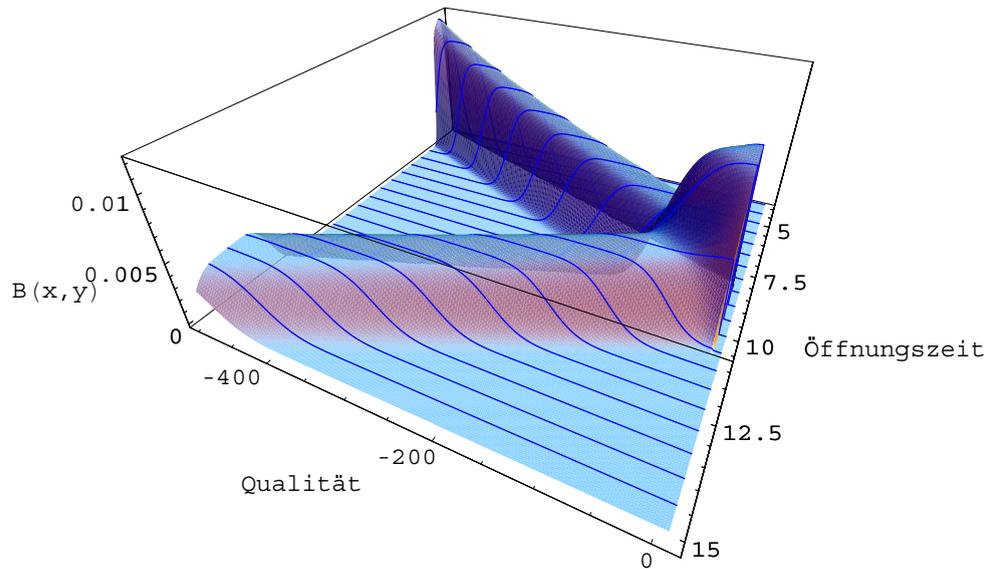


ABBILDUNG 2.4: Funktionsgraph des Bewertungsoperators $B(x, y)$.

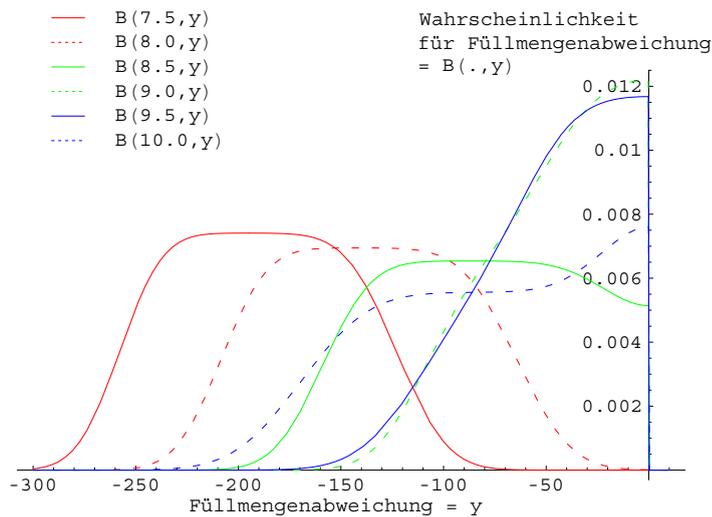


ABBILDUNG 2.5: Funktionsgraph der Informationsfunktion $B(x, y)$ über die Qualität der Füllung bei verschiedenen Einfüllzeiten x . Jede Kurve gibt die Wahrscheinlichkeitsverteilung für die Füllmengenabweichung bei der entsprechenden Öffnungszeit an.

Unternehmen eine im Mittel richtige Entscheidung benötigt, bietet sich als Ordnungsfunktional der Erwartungswert bzw. der Schwerpunkt der Informationsfunktion an:

$$O(B) = \int_{-\infty}^{\infty} y * B(x, y) dy. \quad (2.26)$$

Die Anwendung des Ordnungsfunktionales auf $B(x, y)$ ergibt nun für jede Öffnungszeit x des Ventils einen Qualitätswert. Dieser Qualitätswert lässt sich interpretieren als der Erwartungswert der stets negativ gerechneten Abweichung der Füllmenge von 1000g:

$$\text{Qual}(x) = O(B|x(y)). \quad (2.27)$$

Abbildung 2.6 zeigt den Graphen der Funktion im Bereich $7 \leq x \leq 12$. Da im

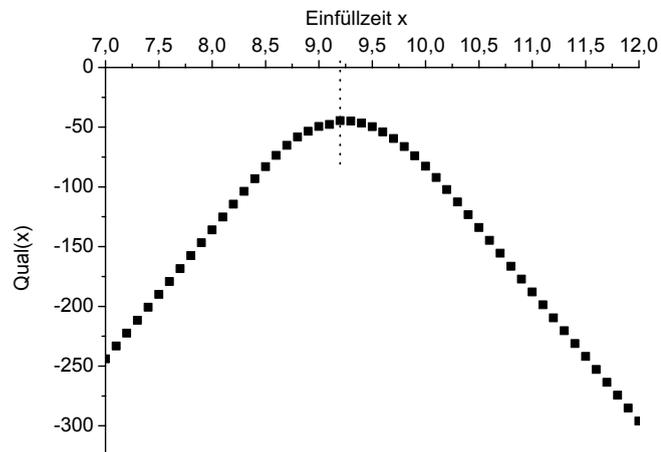


ABBILDUNG 2.6: Der Qualitätswert für verschiedene Einfüllzeiten. Die maximale Qualität wird bei $x_{\text{opt}} = 9,23$ erzielt.

verwendeten Modell ein mögliches Überlaufen des Eimers nicht berücksichtigt wurde, um das Beispiel übersichtlich zu halten², ist die Qualitätsfunktion um das Maximum beidseitig abfallend.

Die Stelle des Maximums dieser Qualitätsfunktion ist die gesuchte optimale Entscheidung. Numerisch ergibt sich für dieses Beispiel der Wert

$$x_{\text{opt}} = 9,23 \text{ Sekunden.} \quad (2.28)$$

Dieser Wert ist etwas niedriger als der im sicheren Fall für $i_T = 24^\circ\text{C}$ und $i_{NL} = 0.9$ berechnete, da im verwendeten Modell eine beliebige Überfüllung des Eimers möglich ist, während die schlechteste durch Unterfüllung erreichbare Bewertung (keine Füllung) -1000 darstellt.

²Eine Berücksichtigung des Überlaufens hätte eine Reihe von Fallunterscheidungen notwendig gemacht.

2.3. Bestandteile des Funktionalmodells

Die bisherige Definition des Funktionalmodells ist bewusst unabhängig vom gewählten Kalkül gehalten, um als Basis für einen Vergleich verschiedener Kalküle untereinander dienen zu können. Im folgenden Abschnitt wird nun gezeigt, welche spezifischen Ausprägungen die verschiedenen Entscheidungskalküle aufweisen.

Die Unterschiede zwischen verschiedenen Entscheidungskalkülen liegen in drei Eigenschaften:

- *Informationsfunktionen:* Im klassischen Fall werden Wahrscheinlichkeitsdichten für die Werte der Umweltparameter herangezogen, im unscharfen Fall Fuzzy-Zugehörigkeitsfunktionen.
- *Bewertungsoperatoren:* Im klassischen Fall werden hauptsächlich punktweise Additionen und Multiplikationen sowie Faltungen der Umweltinformationsfunktionen untereinander oder mit stetigen Funktionen verwendet, während im unscharfen Fall ein breites Spektrum an Fuzzy-Operatoren angewendet wird.
- *Ordnungsfunktional:* Das Ordnungsfunktional O ist im klassischen Fall meist ein Erwartungswert

$$O(\alpha, \mu_\alpha(r)) = \int_{\mathbb{R}} r d\mu_\alpha(r), \quad (2.29)$$

seltener der wahrscheinlichste Ausgang

$$O(\alpha, \mu_\alpha(r)) = x_0 \Leftrightarrow \forall x \in \mathbb{R} : \mu_\alpha(x) < \mu_\alpha(x_0). \quad (2.30)$$

Im unscharfen Fall gibt es eine breite Auswahl an möglichen Ordnungsfunktionalen, auf die in Abschnitt 2.3.3.2 separat eingegangen wird.

Eine scharfe Trennung zwischen unscharfen und klassischen Methoden ist in solch einer Betrachtungsweise allerdings schwierig, da sich die verwendeten Methoden oft nur in der Bezeichnung, nicht aber in der Definition unterscheiden: So ist das in der klassischen Theorie als „Mittelwert der wahrscheinlichsten Ausgänge“ bekannte Ordnungsfunktional in der FDT unter der Bezeichnung „Mittlere maximale Zugehörigkeit“ oder „AVG-MAX-Prinzip“ sehr gebräuchlich. Dadurch reduziert sich der Unterschied auf verschiedene Normierungen der Informationsfunktionen.

In den folgenden Abschnitten werden die im Weiteren verwendeten Funktionen/Funktionale im Einzelnen definiert.

2.3.1. Informationsfunktionen

Informationen über Umweltparameter werden im Folgenden stets in Form von Informationsfunktionen (IF) beschrieben. Deren Definition ist allerdings interpretationsneutral, so dass es Aufgabe der einzelnen Entscheidungskalküle ist, eine solche Interpretation zu liefern. Die Herausforderung in der praktischen Anwendung ist hierbei stets, die unstrukturiert vorliegende Information aus der realen Welt mathematisch in einer IF zu formalisieren.

2.3.1.1. Definitionen

Bei den hier betrachteten Entscheidungskalkülen treten grundsätzlich zwei verschiedene Formen von IF auf:

1. Wahrscheinlichkeitsdichten in der klassischen Entscheidungstheorie und bei den subjektiven Wahrscheinlichkeiten.

Definition 10 (Wahrscheinlichkeitsdichte)

Eine Funktion $\mu \in \mathcal{B}(\mathbb{R})$ heißt „Wahrscheinlichkeitsdichte“ oder „Wahrscheinlichkeitsmaß“, wenn sie \mathbb{R} -integrierbar ist und

$$\int_{\mathbb{R}} d\mu = \int_{\mathbb{R}} \mu(x) dx = 1. \quad (2.31)$$

Anmerkung: Mit der Forderung $\mu \in \mathcal{B}(\mathbb{R})$ ist es offensichtlich nicht möglich, sicheres Wissen als Wahrscheinlichkeitsmaß zu definieren. Um dies zu ermöglichen, sind an Stelle integrierbarer Funktionen Distributionen zu verwenden (vgl. sehr ausführlich [60]). Da für die Zwecke dieser Arbeit die Unterscheidung nicht notwendig ist, wird im Folgenden stets von Funktionen gesprochen, auch wenn es um die Modellierung von Mehr-Punkt Verteilungen geht.

2. Fuzzy-Zugehörigkeitsfunktionen (ZGF) in der Fuzzy-Entscheidungstheorie und der Evidenztheorie.

Definition 11 (Fuzzy-Zugehörigkeitsfunktion (ZGF))

Eine Funktion $\mu \in \mathcal{B}(\mathbb{R})$ heißt „Fuzzy- Zugehörigkeitsfunktion“, wenn

$$\sup_{x \in \mathbb{R}} \mu(x) \leq 1 \quad (2.32)$$

gilt.

Die in der Evidenztheorie vorkommenden Möglichkeits- und Glaubwürdigkeitsfunktionen sind dabei aus formalen Gründen den ZGF zugeordnet. Sie erfüllen die Definition 11 (vgl. [55, 169]), lediglich ihre inhaltliche Interpretation unterscheidet sie von typischen ZGF. Der Bedeutungsgehalt ist aber lediglich für die Modellierung eines Entscheidungssystems, nicht für die Analyse der mathematischen Eigenschaften eines gegebenen Modells relevant. Daher ist eine spezielle Behandlung von Möglichkeits- und Glaubwürdigkeitsfunktionen im Folgenden nicht notwendig.

Es fällt unmittelbar auf, dass eine enge Verwandtschaft zwischen Wahrscheinlichkeitsverteilungen und Zugehörigkeitsfunktionen besteht, wenn einige kleine Einschränkungen für die betrachteten Funktionen gemacht werden. Hierzu werden einige weitere Begriffe eingeführt:

Definition 12 (Träger)

Zu einer gegebenen Informationsfunktion $i \in \mathcal{B}(\mathbb{R})$ wird die Menge

$$S(i) = \{x \in \mathbb{R} | i(x) > 0\} \quad (2.33)$$

als „Träger von i “ oder auch als „Support von i “ bezeichnet.

Definition 13 (Beschränkte Zugehörigkeitsfunktion)

Eine Zugehörigkeitsfunktion μ heißt „beschränkt“, wenn ihr Träger beschränkt ist, also wenn es zwei Zahlen $a, b \in \mathbb{R}$ gibt, so dass gilt

$$\forall x \in \mathbb{R} : (x < a \vee x > b) \Rightarrow \mu(x) = 0. \quad (2.34)$$

Definition 14 (Beschränkte Wahrscheinlichkeitsdichte)

Eine Wahrscheinlichkeitsdichte μ heißt „beschränkt“, wenn ihr Träger beschränkt ist, also wenn es zwei Zahlen $a, b \in \mathbb{R}$ gibt, so dass gilt

$$\forall x \in \mathbb{R} : (x < a \vee x > b) \Rightarrow \mu(x) = 0. \quad (2.35)$$

Aus diesen Definitionen ist unmittelbar ersichtlich, dass sich beschränkte Wahrscheinlichkeitsdichten und beschränkte ZGF nur um einen Normierungsfaktor unterscheiden.

Lemma 1 (Zusammenhang zwischen Wahrscheinlichkeitsdichten und ZGF)

Ist die überall definierte Funktion $i_1 : \mathbb{R} \rightarrow \mathbb{R}$ eine beschränkte Wahrscheinlichkeitsdichte, so ist die Funktion

$$i_2 := \frac{i_1}{\max_{x \in \mathbb{R}} i_1(x)} \tag{2.36}$$

eine ZGF.

Ist umgekehrt i_2 eine beschränkte ZGF, so ist die Funktion

$$i_1 := \frac{i_2}{\int_{\mathbb{R}} i_2} \tag{2.37}$$

eine Wahrscheinlichkeitsdichte, wenn mit $\int_{\mathbb{R}} i_2$ der notwendigerweise beschränkte Flächeninhalt von i_2 bezeichnet wird. Für den Fall $\int_{\mathbb{R}} i_2 = 0$ ist im Sinne der Bemerkung zu Definition 10 die entsprechende (Delta-)Distribution zu verwenden

Ebenso unmittelbar ist ersichtlich, dass jede (auch unbeschränkte) Wahrscheinlichkeitsdichte durch Division mit ihrem Supremum in eine ZGF überführt werden kann. Lediglich im umgekehrten Fall bereiten ZGF i_2 Probleme, für die $\int_{\mathbb{R}} i_2$ nicht existiert, wobei die Verwendung solcher ZGF für die Beschreibung einer unsicheren Größe ohnehin zweifelhaft erscheint, da ihre Interpretation alles andere als offensichtlich ist.

Da der Definitionsbereich aller von Computern verarbeiteten numerischen Funktionen ohnehin auf den (endlichen) Zahlenbereich des Computers beschränkt ist, stellt eine Forderung von beschränkten Zugehörigkeitsfunktionen bei technischen Anwendungen keine Einschränkung der Allgemeinheit dar.

2.3.1.2. Modellierung der Information

Aus unstrukturierten, linguistischen Informationen oder einzelnen Messwerten Informationsfunktionen zu modellieren, welche die Unsicherheit der Information widerspiegeln, ist derzeit kaum systematisch möglich(vgl. [109]).

Im Fall der klassischen Entscheidungstheorie gibt es zumindest im Fall einer Vielzahl von Messwerten mit dem zentralen Grenzwertsatz (vgl. [9][S. 226ff]) einen Anhaltspunkt für den Anwender:

Satz 1 (Zentraler Grenzwertsatz)

Sei X_n eine Familie beliebig verteilter Zufallsvariablen, $\sigma_n := \sigma(x_n)$, $s_n := \sigma(X_1, X_2, \dots, X_n)$ und $\mu_n := E(X_n)$, so konvergiert die Funktion

$$S_n := \frac{1}{s_n} \sum_{j=1}^n (X_j - \mu_j) \tag{2.38}$$

in Verteilung gegen $N_{0,1}$ genau dann, wenn für jedes $\epsilon > 0$ gilt

$$\lim_{n \rightarrow \infty} L_n(\epsilon) = 0 \quad (2.39)$$

mit

$$L_n(\epsilon) := s_n^{-2} \sum_{j=1}^n E((X_j - \mu_j)^2; |X_j - \mu_j| \geq \epsilon s_n). \quad (2.40)$$

Insbesondere konvergiert die Verteilung des Mittelwertes identisch verteilter Zufallsvariablen stets gegen eine Normalverteilung.

Ist also eine Vielzahl von Messwerten eines Parameters vorhanden und deren Mittelwert für die Entscheidung relevant, so kann dieser Mittelwert mit gewisser Berechtigung als normalverteilt angenommen werden.

Für den Fall einzelner Messwerte eines Messgerätes mit bekanntem - und normalverteilten - statistischen Messfehler liefert das Bayessche Theorem

$$P(b|a) = \frac{P(b) * P(a|b)}{P(a)} \quad (2.41)$$

einen Anhaltspunkt für die Modellierung des realen Wertes (vgl. [115]).

Beispiel 2 (Anwendung des Bayesschen Satzes)

Als Eingangsinformation eines Entscheidungsmodells wird eine physikalische Messgröße verwendet, wobei deren Unsicherheit auf Grund stochastischer Messfehler explizit im Modell berücksichtigt werden soll.

Bei bekanntem Messwert a wird die Wahrscheinlichkeitsverteilung für folgendes Ereignis benötigt: Der wahre Wert ist b unter der Bedingung, dass der Messwert a beträgt. Aus den Datenblättern des Messgerätes ist allerdings nur die Wahrscheinlichkeitsverteilung für das folgende Ereignis ablesbar: Das Messgerät zeigt den Wert a unter der Bedingung an, dass der wahre Wert b ist.

Unter der Annahme, dass weder das Messgerät noch der wahre Wert eine Präferenz für bestimmte Werte hat, also $P(a) = \text{const}$ und $P(b) = \text{const}$, besagt das Bayessche Theorem, dass $P(b|a)$ bei einem Messgerät mit normalverteiltem Messfehler ebenfalls normalverteilt ist.

Bei der Modellierung von einzelnen, linguistischen Aussagen wie „Morgen wird es sehr heiß“ bietet aber die Wahrscheinlichkeitstheorie ebenso wenig Hilfestellung wie die FDT: In beiden Fällen ist der Anwender auf grobe Annahmen angewiesen, was dazu führt, dass verschiedene Anwender in derselben Situation die gleiche

linguistische Information sehr unterschiedlich formalisieren (siehe z.B. [69]). Es sind meist Gründe wie z.B. die Einfachheit der Darstellung, die dazu führen, dass solche Informationen im klassischen Kontext ebenfalls als normalverteilte Werte und im Fuzzy-Kontext als dreiecksförmige ZGF modelliert werden.

Im Rahmen der Arbeit wird die eigentliche Übertragung linguistischer Information in Informationsfunktionen nicht untersucht, da eine derartige Fragestellung für sich genommen ein eigenes Forschungsgebiet darstellt (Linguistische Semasiologie) und damit den Rahmen dieser Arbeit sprengt. Allerdings wird bei der Analyse verschiedener Ordnungsfunktionale in Abschnitt 4.2 die Auswirkung der verschiedenen Funktionsformen auf die Störungsanfälligkeit der Ergebnisse diskutiert.

2.3.1.3. Spezielle Eigenschaften von Informationsfunktionen

Bei der Auswertung von Informationsfunktionen durch Bewertungsoperatoren und Ordnungsfunktionale werden insbesondere in der Fuzzy-Theorie einige spezifische Eigenschaften verwendet, die hier kurz eingeführt werden sollen (vgl. z.B. [112, 179]).

Definition 15 (α -Schnitt)

Zu einer gegebenen Informationsfunktion μ und einer gegebenen reellen Zahl $\alpha \in [0; 1]$ wird die Menge

$$\alpha(\mu) = \{x \in \mathbb{R} \mid \mu(x) \geq \alpha\} \quad (2.42)$$

als „ α -Schnitt von μ “ bezeichnet (siehe Abbildung 2.7).

Definition 16 (Kern)

Der 1-Schnitt einer ZGF μ , also die Menge

$$K(\mu) = \{x \in \mathbb{R} \mid \mu(x) = 1\}, \quad (2.43)$$

wird auch als „Kern von μ “ genannt (siehe Abbildung 2.8).

Mit Hilfe dieser Parameter lassen sich nun einige Sonderfälle von ZGF definieren:

Definition 17 (Normale Zugehörigkeitsfunktion)

Eine Zugehörigkeitsfunktion μ heißt „normal“, wenn

$$\sup_{x \in \mathbb{R}} (\mu(x)) = 1. \quad (2.44)$$

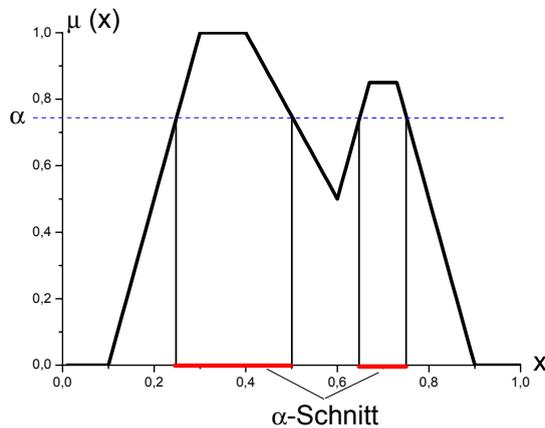


ABBILDUNG 2.7: Für eine gegebene ZGF μ ist der α -Schnitt die (klassische) Menge aller Punkte x , für die $\mu(x) = \alpha$ ist.

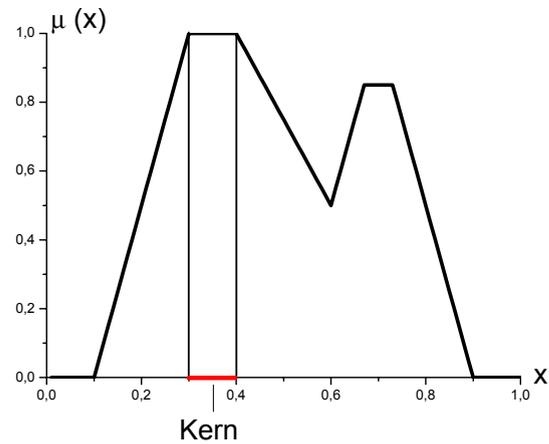


ABBILDUNG 2.8: Der Kern einer ZGF ist ihr 1-Schnitt.

Definition 18 (Einzelpunkt-Verteilung)

Eine Funktion $\mu \in \mathcal{B}(\mathbb{R})$ heißt „Einzelpunkt-Verteilung“, wenn es eine endliche Menge $M \subset \mathcal{P}(\mathbb{R})$ gibt, so dass gilt:

$$\forall x \in M^C : \mu(x) = 0. \tag{2.45}$$

Definition 19 (Fuzzy-Zahlen und -Intervalle)

Eine stückweise stetige ZGF μ wird

- Fuzzy-Zahl genannt, wenn ihr Kern genau ein Element enthält,
- Fuzzy-Intervall, wenn ihr Kern ein Intervall in \mathbb{R} ist, und
- Echtes Fuzzy-Intervall, wenn alle α -Schnitte Intervalle sind.

(Siehe Abbildung 2.9).

2.3.2. Bewertungsoperatoren

Aufgabe des Bewertungsoperators ist es, die Informationen über den Umweltzustand sowie die Gewichtungen des Entscheiders zu einer Bewertungsfunktion für jede Handlungsalternative zusammenzufassen. Diese Bewertungsfunktion

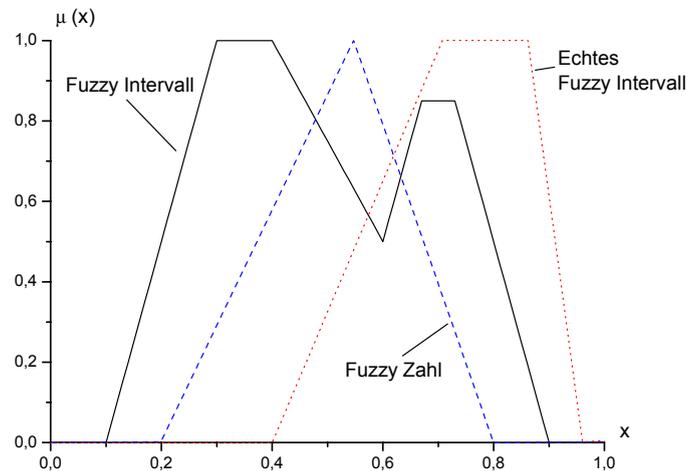


ABBILDUNG 2.9: Einige spezifische Formen von ZGF: Fuzzy-Zahl, Fuzzy-Intervall und echtes Fuzzy-Intervall

trägt die Information über den Nutzen der möglichen Ergebnisse der Handlung und hat deswegen wiederum die Form einer Informationsfunktion.

Da der Bewertungsoperator die eigentliche Information des Entscheidungsmodells über die zu entscheidende Fragestellung enthält, gibt es natürlich letztendlich genauso viele verschiedene Bewertungsoperatoren wie Entscheidungsmodelle. Trotzdem treten in jedem Entscheidungskalkül typische Formen auf, welche in vielen Anwendungen Verwendung finden.

Im Folgenden werden nur Bewertungsoperatoren mit zwei Eingangsfunktionen betrachtet. Die Hinzunahme weiterer Informationen kann durch Kaskadierung mehrerer solcher Operatoren erreicht werden ($B(i_1, i_2, i_3) := B_1(B_2(i_1, i_2), i_3)$).

2.3.2.1. Klassische Entscheidungstheorie

In der klassischen Entscheidungstheorie kommen - je nachdem, ob die Informationen sich auf den gleichen Parameter beziehen oder unterschiedliche Parameter beschreiben - insbesondere zwei Typen von Bewertungsoperatoren vor:

Mittelwertoperator Sind i_1 und i_2 zwei unterschiedliche Parameter, so wird als Bewertungsoperator oft das punktweise arithmetische Mittel, also

$$B_{\text{MiWe}}(i_1, i_2)(x) := \frac{i_1(x) + i_2(x)}{2} \quad (2.46)$$

gewählt (vgl. Abbildung 2.10). Allerdings erkennt ist in der Abbildung zu erkennen, dass das Ergebnis der Anwendung auf zwei Ein-Peak-Funktionen

eine Zwei-Peak-Funktion ist, welche sich als Eingangsfunktion für Ordnungsfunktionale als nicht in jedem Fall optimal herausstellen wird. Wenn dies vermieden werden soll, so bietet sich die Verwendung eines anderen Bewertungsoperators an:

Faltungsmittelwert In diesem Fall gilt:

$$B_{\text{Falt}}(i_1, i_2)(x) := 2 * \int i_1(t) * i_2(2x - t) dt \quad (2.47)$$

(vgl. Abbildung 2.11).

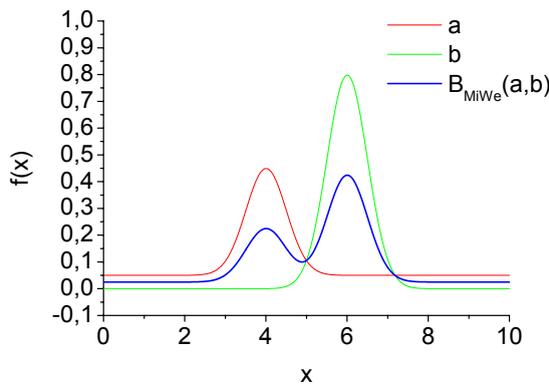


ABBILDUNG 2.10: Der Mittelwertoperator $B_{\text{MiWe}}(a, b)$ angewendet auf die Informationsfunktionen a und b .

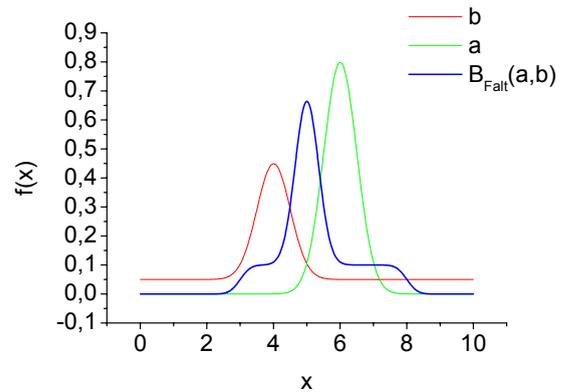


ABBILDUNG 2.11: Der Faltungsmittelwert $B_{\text{Falt}}(a, b)$ zweier Ein-Peak-Eingangsfunktionen a und b .

Der wesentliche Unterschied zwischen beiden Formen ist im Fall $i_1 = i_2 = i$ für eine Wahrscheinlichkeitsdichte i zu erkennen. Während trivialerweise

$$B_{\text{MiWe}}(i, i) = i \quad (2.48)$$

ist und somit

$$EW(B_{\text{MiWe}}(i, i)) = EW(i) \quad (2.49)$$

$$\text{Var}(B_{\text{MiWe}}(i, i)) = \text{Var}(i), \quad (2.50)$$

gilt für den Faltungsmittelwert:

Satz 2 (Eigenschaften des Faltungsmittelwertes)

Seien

$$B_{\text{Falt}}(i_1, i_2)(x) := 2 * \int i_1(t) * i_2(2x - t) dt \quad (2.51)$$

der Faltungsmittelwert und i eine Wahrscheinlichkeitsdichte, so gilt

$$EW(B(i, i)) = EW(i) \quad (2.52)$$

$$Var(B(i, i)) = \frac{1}{2}Var(i), \quad (2.53)$$

wenn $E(x)$ den Erwartungswert und $Var(x)$ die Varianz bezeichnet.

Beweis. Die Vertauschung der Integrationsreihenfolge ist jeweils möglich, da alle auftretenden Teilintegrale existieren. Daneben gilt $\int i(x)dx = 1$, da i eine Wahrscheinlichkeitsdichte ist.

$$\begin{aligned} EW(B(i, i)) &= \int x * B(i, i) dx \\ &= \int x \int 2i(t)i(2x - t) dt dx \\ &= \int i(t) \int (2x)i(2x - t) dx dt \\ &= \int i(t) 0.5 * \int (x + t)i(x) dx dt \\ &= 0.5 * \int i(t) \left[\int xi(x) dx + \int ti(x) dx \right] dt \\ &= 0.5 * \int i(t) [EW(i) + t] dt \\ &= 0.5 * \int i(t)EW(i) dt + 0.5 * \int i(t)t dt \\ &= 0.5 * EW(i) + 0.5 * EW(i) \\ &= EW(i) \end{aligned}$$

Unter Verwendung, dass $EW(B(i, i)) = EW(i) =: EW$, ergibt sich weiter

$$\begin{aligned} Var(B(i, i)) &= 2 \int i(t) \int i(2x - t) [x - EW]^2 dx dt \\ &= 2 \int i(t) 0.5 * \int i(x) [0.5(x + t) - EW]^2 dx dt \\ &= \int i(t) \int i(x) [(x + EW) + 0.5(t - x)]^2 dx dt \\ &= \int i(t) \int i(x) [(x + EW)^2 + (t - x) + 0.25(t - x)^2] dx dt \\ &= \int i(t) \left[Var(i) + \int i(x)t dx - \int i(x)x dx + \int i(x)0.25(t - x)^2 dx \right] dt \\ &= Var(i) + \int i(t) * \dots \\ &\quad \dots \left[t - EW + 0.25 \int i(x)t^2 dx - 0.5 \int i(x)t x dx + 0.25 \int i(x)x^2 dx \right] dt \\ &= Var(i) + EW - EW + \int i(t) [0.25t^2 - 0.5t * EW + 0.25EW(x^2)] dt \end{aligned}$$

$$\begin{aligned}
 &= \text{Var}(i) + 0.25EW(i^2) - 0.5EW^2 + 0.25EW(i^2) \\
 &= \text{Var}(i) - 0.5 * (EW^2 - EW(i^2)) \\
 &= \frac{1}{2}\text{Var}(i),
 \end{aligned}$$

wobei der letzte Schritt aus dem Verschiebungssatz folgt. ■

Zwei gleiche Informationen verringern also im ersten Fall die Varianz des Ergebnisses nicht, während sie im zweiten halbiert wird. Welcher von beiden Operatoren in einer konkreten Anwendung sinnvoller ist, hängt natürlich von der Quelle der Information ab. Bei zwei voneinander unabhängigen Messungen des gleichen Wertes wird allerdings der Faltungsmittelwert sinnvoller sein, da eine unabhängige Bestätigung die Unsicherheit über den wahren Wert verringern sollte.

Es ist offensichtlich, dass die Berechnung des Faltungsmittelwertes mit deutlich höherer Komplexität verbunden ist, als die des Mittelwertoperators. Auch wenn Fragen der Berechenbarkeit nicht näher betrachtet werden, ist doch zu erwähnen, dass sich der Faltungsmittelwert mit Hilfe der schnellen Fourier-Transformation (FFT) äußerst effizient berechnen lässt. Ist

$$[H(i)](x) := \int i(t)e^{2\pi ixt} dt \tag{2.54}$$

die Fourier-Transformierte von i und H^{-1} deren Inverse, so ergibt sich

$$B_{\text{Falt}}(i_1, i_2)(x) := 2 * \int i_1(t) * i_2(2x - t) dt \tag{2.55}$$

$$= 2 * H^{-1} [H(i_1) * H(i_2)] (2x). \tag{2.56}$$

2.3.2.2. Subjektive Wahrscheinlichkeiten

Die im Kalkül der subjektiven Wahrscheinlichkeiten auftretenden Vertrauenswerte werden als Gewichte in den unterschiedlichen Bewertungsoperatoren verwendet. Prinzipiell sind für die mit diesen Gewichten versehenen Informationsfunktionen alle Bewertungsoperatoren aus der klassischen Entscheidungstheorie anwendbar. Sind zwei Informationsfunktionen i_1 und i_2 mit Gewichten g_1 und g_2 gegeben, so ergibt sich aus den oben genannten Beispielen unmittelbar der bewertete Mittelwertoperator:

$$B(i_1, i_2)(x) := \frac{g_1 * i_1(x) + g_2 * i_2(x)}{g_1 + g_2} \tag{2.57}$$

Demgegenüber ist die Berücksichtigung der Gewichtungsfaktoren im Faltungsmittelwert nicht trivial. Eine Möglichkeit stellt die Verwendung der relativen Gewichte $\frac{g_1}{g_1+g_2}$ als Exponenten der Funktionen dar:

$$\int i_1^{\frac{g_1}{g_1+g_2}}(t) * i_2^{\frac{g_2}{g_1+g_2}}(t - x) dt \tag{2.58}$$

Allerdings ist eine Anwendung des gewichteten Faltungsmittelwert bisher in der Literatur kaum zu finden, wohl auch, da die Berechnungskomplexität des Integralausdrucks relativ hoch ist.

2.3.2.3. Fuzzy Decision Theory

Die FDT kennt eine Vielzahl unterschiedlicher Bewertungsoperatoren, von denen aber nur wenige auch in der praktischen Anwendung zu finden sind. Einen Überblick über die große Vielzahl bisher vorgeschlagener Verknüpfungsoperatoren für Fuzzy-Mengen aus denen sich jeweils Bewertungsoperatoren ableiten lassen findet sich z.B. in [75, 112, 128]. Das Problem der Auswahl eines geeigneten Operators für ein gegebenes Problem ist in der Fuzzy-Entscheidungstheorie bisher nicht allgemein gültig gelöst. Dem Anwender bleibt im konkreten Fall nur übrig, auszuprobieren, welche Form für seine spezifische Anwendung zufriedenstellende Resultate liefert. Mit dem „Zadeh-Plus“-Operator und dem „Gamma“-Operator werden im Folgenden zwei sehr gebräuchliche Vertreter näher untersucht.

Zadeh-Plus-Operator Das von Zadeh [175] schon in seiner Originalarbeit vorgeschlagenen Erweiterungsprinzip für klassische Operationen ergibt als eine mögliche Erweiterung der Mittelwertbildung auf Fuzzy-Mengen den „Zadeh-Plus“-Operator (vgl. Abbildung 2.12):

$$B_{\text{ZaPlu}}(i_1, i_2)(x) = \sup_{\frac{x_1+x_2}{2}=x} [\min(i_1(x_1), i_2(x_2))]. \quad (2.59)$$

Er neigt jedoch aufgrund der enthaltenen Minimumbildung zu einer eher geringen Einschätzung des Zugehörigkeitswertes. Wenn z.B. i_1 durch einen bestimmten Wert \hat{i} nach oben beschränkt ist, so ist es auch $B_{\text{ZaPlu}}(i_1, i_2)$ für jedes beliebige i_2 . Daher schlägt z.B. Rommelfanger [126, S.34ff] vor, statt \min und \sup ein beliebiges anderes t-Norm/-Conorm Paar zu verwenden. Solche Operatoren finden sich allerdings aufgrund ihrer Komplexität kaum in praktischen Anwendungen.

Gamma-Operator Eine andere Klasse der so genannten „kompensatorischen Operatoren“, bei welchen kleine Werte des einen Operanden durch größere Werte des anderen kompensiert werden, findet ebenfalls Verwendung. Insbesondere die Gruppe der Gamma-Operatoren wird gerne in Regelungssystemen eingesetzt (vgl. [12]). Dabei ist für ein $\gamma \in [0, 1]$ (vgl. Abbildung 2.13):

$$B_{\gamma}(i_1, i_2) = (i_1 i_2)^{1-\gamma} (1 - (1 - i_1)(1 - i_2))^{\gamma}. \quad (2.60)$$

Der Gamma-Operator weist allerdings in der Anwendung ein sehr spezielles Verhalten auf, wenn sich die Entropie der beiden Parameter deutlich unterscheidet. Dadurch ist er nur in speziellen Fällen anwendbar. Hierauf wird in Abschnitt 4.1.3 detailliert eingegangen.

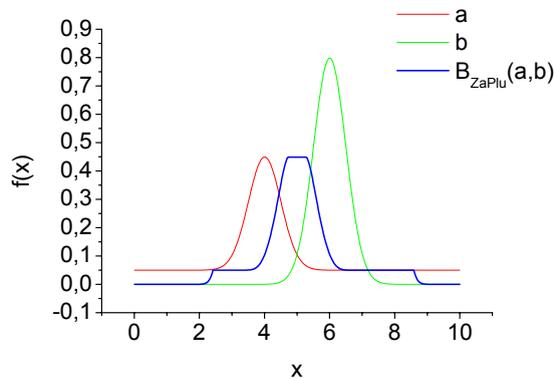


ABBILDUNG 2.12: Anwendung des Zadeh-Plus-Operators auf zwei Eingangsfunktionen a und b

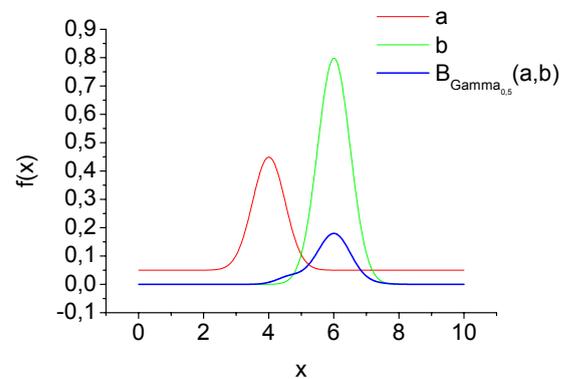


ABBILDUNG 2.13: Anwendung des Gamma-Operators auf zwei Eingangsfunktionen a und b

2.3.2.4. Evidenztheorie

Zadeh [176] erkannte die Evidenztheorie von Anfang an als eine mögliche Interpretation von Fuzzy-Zugehörigkeitsfunktionen (vgl auch [179]). In der Zwischenzeit wurde sie jedoch - losgelöst vom Fuzzy-Kontext - in mehreren Varianten weiterentwickelt, wobei die wichtigsten aktuellen Varianten wohl von Dubois und Prade [41], Smets [146, 147] sowie Kohlas und Monney [81] kommen.

Von den drei genannten Formulierungen ist die letzte am anwendungsnächsten. Bei ihr werden als Bewertungsoperatoren, welche hier „Kombinationsoperatoren für Hinweise“ genannt werden, fast durchgehend Faltungsmittelwerte verwendet [81]. Eine gesonderte Betrachtung spezieller Typen von Bewertungsoperatoren für die Evidenztheorie ist daher nicht notwendig.

2.3.2.5. Verwendete Bewertungsoperatoren

Aus den beschriebenen Bewertungsoperatoren können vier ausgewählt werden, die exemplarisch für ein breites Spektrum der in der Anwendung vertre-

tenen Operatoren stehen: Der punktweise Mittelwert, das Faltungsprodukt, der „Zadeh-Plus-Operator“ sowie der Gamma-Operator.

$$B_{\text{MiWe}}(i_1, i_2)(x) := \frac{i_1(x) + i_2(x)}{2}, \quad (2.61)$$

$$B_{\text{Falt}}(i_1, i_2)(x) := 2 * \int i_1(t) * i_2(2x - t) dt, \quad (2.62)$$

$$B_{\text{ZaPlu}}(i_1, i_2)(x) := \sup_{\frac{x_1+x_2}{2}=x} [\min(i_1(x_1), i_2(x_2))], \quad (2.63)$$

$$B_{\text{Gamma}}(i_1, i_2)(x) := (i_1(x)i_2(x))^{0.5}(1 - (1 - i_1(x))(1 - i_2(x)))^{0.5}. \quad (2.64)$$

Die vier ausgewählten Operatoren werden im Kapitel 4 numerisch auf ihre Störungssensitivität untersucht.

2.3.3. Ordnungsfunktionale

Nachdem durch Anwendung des Bewertungsoperators für jede Handlungsalternative eine Informationsfunktion über deren Qualität vorliegt, ist es Aufgabe des Ordnungsfunktionals, die Qualitätsaussage auf eine reelle Zahl zu reduzieren, um so die „beste“ Alternative bestimmen zu können.

Ein solches Funktional wird in der klassischen Entscheidungstheorie meist „Entscheidungskriterium“ genannt, im Kontext von Fuzzy-Entscheidungsmodellen wird in diesem Zusammenhang der Begriff „Defuzzifizierung“ verwendet.

Es gibt drei prinzipiell verschiedene Ansätze, um Ordnungsfunktionale zu definieren (vgl. [157]): Absolute Ordnung, Referenz-Mengen Methode und Relative Ordnung.

- *Absolute Ordnung: Jeder Funktion wird unmittelbar ein numerischer Wert zugeordnet, z.B. $\max_x(f(x) = 1)$ oder $\int_{\mathbb{R}} f(x)$.*
- *Referenz-Mengen Methode: Es wird ein Ähnlichkeitsfunktional definiert, welches die relative Abweichung von zwei Informationsfunktionen misst. Die Bewertung jeder Funktion ergibt sich dann aus der Abweichung zu einer optimalen Zielfunktion.*
- *Relative Ordnung: Es wird zunächst eine vollständige Vergleichsrelation für beschränkte Funktionen definiert, wie z.B. $f \geq g \Leftrightarrow \forall x \in \mathbb{R} : f(x) \geq g(x)$. In einem zweiten Schritt werden die Funktionen dann so „durchnummeriert“,*

dass die größten Funktionen unter der definierten Relation die höchste Zahl tragen³.

Relative Ordnungen sind in der Anwendung aufwändig, da ihre Berechnungskomplexität für n verschiedenen Handlungsalternativen selbst bei Verwendung von Quicksort-Algorithmen mindestens $O(n \log n)$ beträgt. Andererseits haben sie gerade bei soziologischen Entscheidungsmodellen eine bestimmte Berechtigung, da es für Menschen in der Regel schon schwer genug ist, eine konsistente, also insbesondere transitive Vergleichsrelation anzugeben (z.B. [87]). Da soziologische Fragestellungen aber nicht im Fokus der Arbeit stehen und darüber hinaus Debreu gezeigt hat [33], dass sich aus solch einer relativen Ordnung fast immer eine (stetige) absolute Ordnung entwickeln lässt, werden im Folgenden nur noch absolute Ordnungen betrachtet. Es soll nicht unerwähnt bleiben, dass gerade beim Vergleich von Fuzzy-ZGF in letzter Zeit sehr ausgefeilte relative Methoden entwickelt wurden (z.B. [31, 62, 154]), deren genaue Eigenschaften aber noch kaum erforscht sind.

Während sich im Bereich der klassischen Entscheidungstheorie im Laufe der Zeit drei Entscheidungskriterien herauskristallisiert haben, welche für alle Arten von Entscheidungsproblemen herangezogen werden können, herrscht in der Literatur eine große Vielfalt an Vorschlägen für solche Kriterien in Fuzzy-Modellen. Insbesondere fehlt hier auch noch eine stringente Methode, für ein konkretes Problem das passende Kriterium zu wählen. Es fällt aber auf, dass in letzter Zeit schwerpunktbasierte Kriterien, die letztendlich eine Übertragung des Erwartungswertes auf Fuzzy-Mengen darstellen, immer häufiger vorgeschlagen werden.

2.3.3.1. Klassische Entscheidungstheorie

Die gebräuchlichen⁴ Ordnungsfunktionale im klassischen Kontext sind Maximum-Likelihood-, μ und (μ, σ) -Prinzip.

Definition 20 (Average-Maximum-Likelihood-Prinzip)

(vgl. Abbildung 2.14)

$$O(\alpha_x) := \text{avg}_{x \in \mathbb{R}^X}, \quad \text{so dass} \quad \alpha_x(x) = \max_{y \in \mathbb{R}} \alpha_x(y) \quad (2.65)$$

³Im allgemeinen Fall von überabzählbar vielen Funktionen ist eine derartige Nummerierung nicht trivial. Sie ist aber in jedem Fall auch so möglich, dass die Funktionswerte beschränkt bleiben.

⁴Weitere in der Literatur auftretende Prinzipien, wie MaxMin, Hurwicz[64], Niehans-Savage [111, 138] und Laplace[89] sind lediglich Varianten der drei beschriebenen Kriterien, die sich durch Änderung des Bewertungsoperators erzielen lassen.

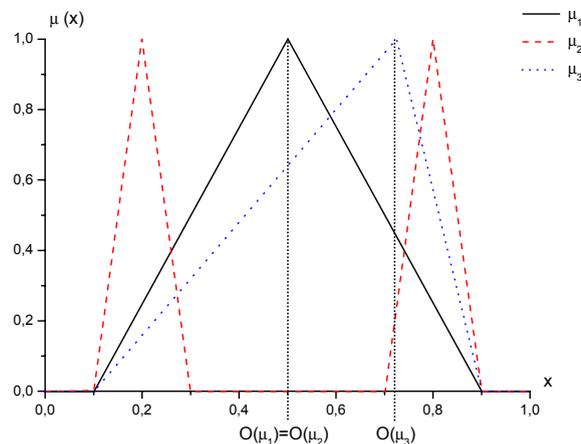


ABBILDUNG 2.14: Anwendung des Average-Maximum-Likelihood-Prinzips auf drei verschiedene Zugehörigkeitsfunktionen

Hierbei wird jede Handlungsalternative mit dem Ausgang bewertet, der am wahrscheinlichsten ist und zwar unabhängig von der absoluten Wahrscheinlichkeit, dass genau das wahrscheinlichste Ergebnis auch eintritt. Vorteil hiervon ist, dass sie sehr einfach durchführbar ist. Bei relativ symmetrischen Bewertungsfunktionen mit einer Spitze, also z.B. Normalverteilungen oder Dreiecksfunktionen, stimmt das Ergebnis darüber hinaus mit dem Erwartungswert überein. Deswegen ist der Maximum-Likelihood-Ansatz in Situationen, in denen es mehr auf hohe Berechnungseffizienz als auf exakt richtige Entscheidungen ankommt, häufig gerechtfertigt.

Definition 21 (Schwerpunktmethode⁵)

(vgl. Abbildung 2.15)

$$O(\alpha_x) := \frac{\int_{\mathbb{R}} x * \alpha_x(x) dx}{\int_{\mathbb{R}} \alpha_x(x) dx} \tag{2.66}$$

Im Falle dass α_x eine Wahrscheinlichkeitsdichte ist, vereinfacht sich die Berechnung zu:

$$O(\alpha_x) := \int_{\mathbb{R}} x * \alpha_x(x) dx \tag{2.67}$$

In klassischen Entscheidungsmodellen ist der Erwartungswert das vorherrschende Ordnungsfunktional. Das ist damit zu rechtfertigen, dass zumindest bei häufiger Wiederholung derselben Entscheidung, die Maximierung des Nutzenerwartungswertes bei jeder Einzelentscheidung die Maximierung des Gesamtnutzens bedeutet.

⁵Auch μ -, Bernoulli- oder Erwartungsnutzen-Prinzip genannt

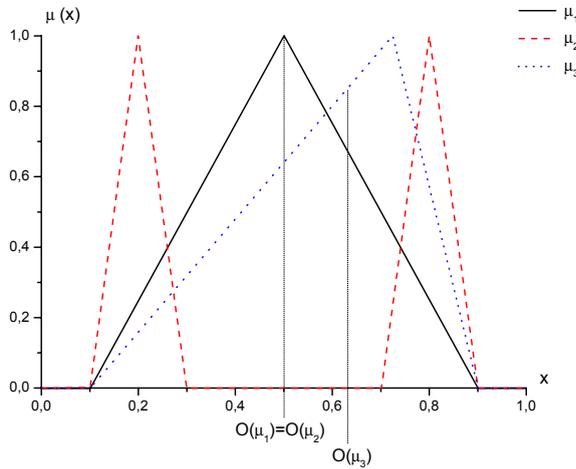


ABBILDUNG 2.15: Anwendung der Schwerpunktmethodede auf drei verschiedene Zugehörigkeitsfunktionen

Eventuelle Einwände gegen die Schwerpunktmethodede, in der Praxis spiegle der Erwartungswert oft nicht hinreichend die Problematik von sehr unerwünschten Ausgängen mit sehr kleinen Eintrittswahrscheinlichkeiten wider - meist mit dem Beispiel des "russischen Roulettes" unterlegt -, werden durch das Funktionalmodell wirksam entkräftet: Die Berücksichtigung solcher negativen Auswertungen ist Aufgabe des Bewertungsoperators, nicht des Ordnungsfunktional. Insgesamt ist das μ -Prinzip sicher das am weitesten verbreitete Entscheidungsprinzip im klassischen Kontext.

Definition 22 ((μ, σ)-Prinzip)

Für ein beliebiges $\tau \in [0; 1]$ ist

$$\bar{O}(\alpha_x) := \tau * \mu + (1 - \tau) * \sigma \tag{2.68}$$

$$\mu := \frac{\int_{\mathbb{R}} x * \alpha_x(x) dx}{\int_{\mathbb{R}} \alpha_x(x) dx} \tag{2.69}$$

$$\sigma^2 := \frac{\int_{\mathbb{R}} (x - \mu)^2 * \alpha_x(x) dx}{\int_{\mathbb{R}} \alpha_x(x) dx}. \tag{2.70}$$

Offensichtlich stimmt für $\tau = 1$ das (μ, σ) -Prinzip mit dem Bernoulli-Prinzip überein. Davon abweichend wird aber für alle $\tau < 1$ nicht nur der Erwartungswert, sondern auch die Standardabweichung des Ergebnisses explizit in die Entscheidung mit einbezogen. Das erscheint auf den ersten Blick widersinnig, da eine der fundamentalen Annahmen der klassischen Entscheidungstheorie ja gerade die Indifferenz des Entscheiders gegenüber einem Nutzen⁶ von 100

⁶wohlgermerkt ist hier von Nutzen, nicht von Auszahlung oder Ähnlichem die Rede

und einer 50%-Chance auf einen Nutzen von 200 war. Allerdings ist nicht von der Hand zu weisen, dass - keineswegs erst in neuerer Zeit in Strömungen zum Produktionsqualitätsmanagement wie Six-Sigma - auch die Verlässlichkeit von Entscheidungen ein wichtiges Qualitätskriterium ist. So ist es Anlegern meist wesentlich angenehmer, wenn sie jährlich etwa gleich hohe Renditen mit ihren Fonds-Anlagen erwirtschaften, als wenn der Fonds-Manager zwar im Mittel (und damit langfristig) eine bessere Performance liefert, aber das jährliche Ergebnis scheinbar zufällig schwankt. Auch in der Produktion von Standardgütern, die genormten Qualitätsanforderungen genügen müssen, ist es wichtiger, mit kleiner Qualitätsstreuung knapp über den Mindestanforderungen zu bleiben als eine im Mittel wesentlich höhere Qualität mit annähernd 50% Ausschuss zu produzieren.

2.3.3.2. Fuzzy Decision Theory

Aus den vielen in der Literatur vorgeschlagenen Entscheidungsprinzipien bzw. Defuzzifizierungs-Funktionalen seien hier nur die wichtigsten herausgegriffen. Übersichten wie auch weiterführende Vergleiche solcher Funktionale finden sich z.B. in [93, 135, 165, 166]. Zunächst ist die **Schwerpunktmethode**⁷ zur Defuzzifizierung zu nennen, die genau dem Bernoulli-Prinzip in der klassischen Entscheidungstheorie entspricht. Nach der probabilistischen Interpretation der Zugehörigkeitsfunktion ist ihre Verwendung auch durchaus nahe liegend. Allerdings hat sie einen gravierenden Nachteil im Zusammenhang mit Fuzzy-Entscheidungssystemen: Ist in einem konkreten Fall die Fuzzy-Modellierung gewählt worden, weil ein „exaktes“ klassisches Modell aufgrund Beschränkungen in der Berechnungszeit/-komplexität nicht implementierbar ist, so gehen durch den notwendigen Integrationsvorgang zur Bestimmung des Schwerpunktes viele der Zeitvorteile wieder verloren.

Yager und Filev schlagen in [171] darüber hinaus eine ganze Reihe von Abwandlungen der Schwerpunktmethode vor, die hier nicht im Einzelnen diskutiert werden können.

Definition 23 (ϵ -Quantil)

Für ein $\epsilon \in [0; 1]$ sei

$$O(\alpha_x) = x, \text{ so dass } \frac{\int_{-\infty}^x \alpha_x(x) dx}{\int_{\mathbb{R}} \alpha_x dx} = \epsilon. \quad (2.71)$$

⁷Zu finden in allen allgemeinen Lehrbüchern zur FDT, z.B. [75, 79]

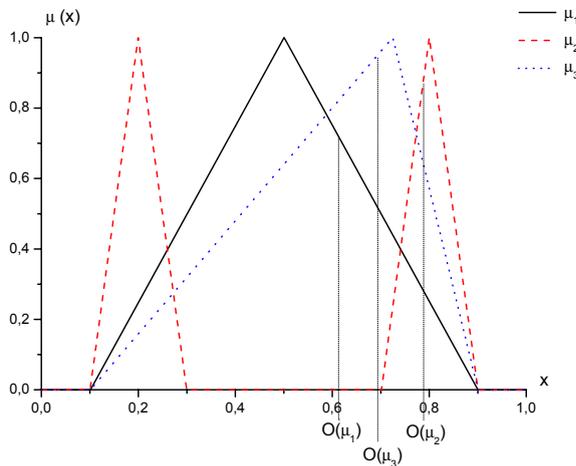


ABBILDUNG 2.16: Anwendung des $\epsilon (= 0.7)$ -Quantils auf drei verschiedene Zugehörigkeitsfunktionen

Das ϵ -Quantil hat seinen Ursprung in der beschreibenden Statistik, wo es häufig als Charakterisierung einer Wahrscheinlichkeitsverteilung verwendet wird. So ist das 0,5-Quantil auch als „Median“ bekannt.

Wird α_x als Wahrscheinlichkeitsverteilung über den Nutzen interpretiert, so gibt das ϵ -Quantil den Grenzwert an, unter dem das Ergebnis mit Wahrscheinlichkeit ϵ liegen wird. Es bietet sich als Ordnungsfunktional immer dann an, wenn der Entscheider genaue Vorgaben hat, welches Risiko er eingehen kann, wie z.B. der einzelne Devisenhändler in einer Großbank.

Definition 24 (Max- α -Schnitt Prinzip)

Für ein $\alpha_0 \in (0; 1]$ sei

$$O(\alpha_x) = \max_{x | \alpha_x(x) \geq \alpha_0} x. \quad (2.72)$$

Das Max- α -Schnitt-Prinzip stellt eine Verallgemeinerung des MaxMax- und des Maximum-Likelihood-Prinzips im klassischen Kontext dar⁸: Für normale ZGF ergibt sich bei $\alpha_0 = 1$ genau das Maximum-Likelihood-Prinzip, für beschränkte ZGF ergibt sich beim Grenzübergang zu $\alpha_0 = 0$ gerade das MaxMax-Prinzip. Damit erbt das Max- α -Schnitt Prinzip auch prinzipiell deren Schwächen, nämlich dass zur Beurteilung der Alternative nur ein einzelnes, nicht unbedingt repräsentatives Ergebnis herangezogen wird. Allerdings eignet es sich zur schnellen Vorsortierung der Handlungsalternativen, indem alle Möglichkeiten aussortiert werden, die nicht zumindest mit einem bestimmten α_0 -Level ein gewünschtes Resultat erreichen können.

⁸Ähnliche Erweiterungen finden sich z.B. bei [40].

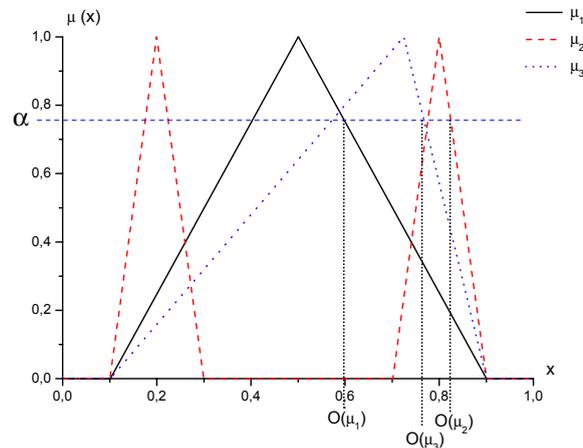


ABBILDUNG 2.17: Anwendung des Max- α -Schnitt Prinzip auf drei verschiedene Zugehörigkeitsfunktionen.

Zu einem ähnlichen Zweck wird auch oft das „Min- α -Schnitt-Prinzip“ herangezogen, mit dem alle Alternativen aussortiert werden, die ein gewünschtes Ergebnis nicht mindestens mit einem bestimmten α_0 -Level erreichen.

Definition 25 (Min- α -Schnitt Prinzip)

Für ein $\alpha_0 \in (0; 1]$ sei

$$O(\alpha_x) = \min_{x | \alpha_x(x) \geq \alpha_0} x. \tag{2.73}$$

Über die bereits vorgestellten Ordnungsfunktionale hinaus gibt es in der Literatur eine große Vielzahl weiterer Vorschläge (z.B. die Drehmomentmethode [76] oder das Inferenzfilter [77]) , die teilweise mit sehr spezifischen Eigenschaften für ganz spezielle Anwendungen konstruiert wurden (vgl. [20, 49, 130, 134, 170]). Eine Darstellung sämtlicher Varianten wird den Rahmen der Arbeit jedoch sprengen.

2.3.3.3. Für den Vergleich herangezogene Ordnungsfunktionale

Da sich für alle beschriebenen Funktionale Beispiele konstruieren lassen, in denen die daraus abgeleitete Reihenfolge von zwei Fuzzy-Mengen intuitiv falsch erscheint, werden in der Literatur eine große Anzahl von Alternativen vorgeschlagen. Die Auswahl eines für eine konkrete Anwendung angemessenen Defuzzifizierungsfunktionals ist eine in der Forschung viel diskutierte Frage der FDT. Hier scheint die FDT noch weit von der Kanonisierung der klassischen Theorie entfernt zu sein, bei welcher Ordnungsfunktionale von Erwartungswert-Typ weitgehend anerkannt sind.

Die folgenden drei Ordnungsfunktionale decken allerdings einen großen Teil der Anwendungen in beiden Kalkülen ab (siehe z.B. [75, 79, 90]) und werden daher im Folgenden untersucht:

- μ -Prinzip bzw. Schwerpunktmethod

$$O_1(i) := \frac{\int_{\mathbb{R}} x * i(x) dx}{\int_{\mathbb{R}} i(x) dx} \quad (2.74)$$

- Average-Maximum-Likelihood Prinzip:

$$O_2(i) := \text{avg}_{x \in \mathbb{R}} x, \text{ so dass } \alpha_x(x) = \max_{y \in \mathbb{R}} \alpha_x(y) \quad (2.75)$$

- ϵ -Quantil mit $\epsilon = 0,7$:

$$O_3(i) := x_0 \left| \int_{-\infty}^{x_0} i(x) dx = 0.7 * \int_{\mathbb{R}} i(x) dx \right. \quad (2.76)$$

2.4. Zusammenfassung

Mit dem in Abschnitt 2.2 entwickelten Funktionalmodell (Definition 8) steht nun ein einheitlicher Formalismus zur Verfügung, der die mathematische Beschreibung von Entscheidungsmodellen kalkülübergreifend erlaubt. Darüber hinaus wurde demonstriert, wie sich typische Strukturen aus Entscheidungsmodellen in diesem Modell darstellen lassen.

Das entwickelte Modell kann daher als Grundlage des Vergleichs der verschiedenen Entscheidungskalküle dienen. Um diesen Vergleich durchzuführen, wird im nächsten Abschnitt ein neues, auf Messung der Informationsnutzung basierendes Vergleichskriterium entwickelt, dass sich auf die beiden Bestandteile des Funktionalmodells anwenden lässt.

3. Neues Vergleichskriterium für Entscheidungsmodelle

Beim Vergleich der Qualität von Entscheidungsmodellen ergibt sich eine prinzipielle Herausforderung: Die Qualität des konkreten Ergebnisses der Anwendung eines Entscheidungsmodells sagt sehr wenig über die Qualität des Modells an sich aus. Das liegt insbesondere daran, dass auch die optimale Handlung (mit kleiner Wahrscheinlichkeit) ein relativ schlechtes Ergebnis nach sich ziehen kann.

Lediglich wenn die Entscheidung unter Laborbedingungen sehr häufig wiederholt werden kann, um so eine signifikante Statistik über die Ausgänge zu bilden, können verlässliche Aussagen getroffen werden. Dann besteht allerdings das Problem, dass zur Simulation der Ergebnisse gerade wieder das Modell des eigentlichen Vorgangs verwendet werden muss, welches ursprünglich zur Entwicklung des Entscheidungsmodells benutzt wurde.

Deswegen ist nicht offensichtlich, wie verschiedene Entscheidungskalküle miteinander verglichen werden können. Ribo und Pinz [125] modellieren ihr Entscheidungssystem zur Kartographierung der Hindernisse in einem Raum durch einen autonomen Roboter in drei verschiedenen Modellen und prüfen dann, wie genau die interne Karte des Roboters die wirkliche Raumeometrie widerspiegelt. Der Ansatz ist auf eine allgemeine Untersuchung nur schlecht übertragbar, da zum einen die „richtige“ Lösung des Entscheidungsproblems hier bekannt ist und zum anderen auf die Art und Weise die konkreten Modelle, nicht aber die Kalküle an sich verglichen werden können: Da es kein abstraktes Verifikationsverfahren für die Qualität der Modellierung gibt, ist stets unentscheidbar, ob das bessere Abschneiden eines Kalküls bei der Anwendung auf ein Referenzproblem durch die bessere Eignung des Kalküls an sich oder lediglich durch die Wahl eines nicht optimalen Modells aus den anderen Kalkülen begründet ist.

Im Folgenden wird deswegen ein neuer Ansatz entwickelt, bei dem nicht versucht wird, die Qualität des Ergebnisses der Entscheidung absolut zu bestimmen. Stattdessen wird in einer Art „Sensitivitätsanalyse“ ermittelt, inwieweit sich Fehler in den Eingangsdaten, wie sie in klassischen Modellen durch zu wenige Erhebungsstichproben oder in unscharfen Modellen durch die relativ willkürli-

che Festlegung von Zugehörigkeitsfunktionen entstehen, auf die Entscheidung auswirken, beziehungsweise inwieweit eine Verbesserung der Sicherheit der Eingangsinformation auch zu einer solchen Verbesserung des Ergebnisses beiträgt.

3.1. Messung des Informationsgehaltes

3.1.1. Entropie

Um den Informationsgehalt einer Informationsfunktion zu messen, wird in der Informationstheorie die Entropie der Funktion ermittelt (vgl. [2]). Die dabei am häufigsten benutzte Form ist:

Definition 26 (Hartley-Shannon-Entropie (HS-Entropie) [142])

Sei $X = \{x_1, \dots, x_n\}$ die Menge der Zustände eines Systems und $P = [p_1, \dots, p_n]$ eine Wahrscheinlichkeitsverteilung auf X , also $p_i \in [0; 1]$ und $\sum_i p_i = 1$, dann ist die „Entropie“ von P definiert¹ ([142]) als

$$H(P) = - \sum_{x_i \in X} p(x_i) \log p(x_i). \quad (3.1)$$

Im Fall eines vollständigen Zustandsraumes mit einer Wahrscheinlichkeitsverteilung μ darüber gilt analog:

$$H(\mu) = - \int_X \mu(x) \log \mu(x) dx. \quad (3.2)$$

Die so definierte Funktion hat sechs Eigenschaften, welche sie besonders geeignet erscheinen läßt, um den Informationsgehalt oder genauer den Gehalt an Unwissen einer Wahrscheinlichkeitsverteilung zu messen:

1. **SYMMETRIE:** $H(P)$ ist unabhängig von der Ordnung von X .
Diese Eigenschaft garantiert damit auch die weitgehende Unabhängigkeit der Entropie der Information von der Modellierung der Fragestellung
2. **ERWEITERBARKEIT:** Sei $X^* := X \cup \{x_{n+1}\}$ mit $P^* = [p_1, \dots, p_n, 0]$, so ist $H(P^*) = H(P)$.
Das Hinzufügen von unmöglichen Zuständen verändert also die Entropie nicht. Die Eigenschaft ist deswegen wichtig, weil hierdurch vernünftige

¹dabei sei stets $0 * \log(0) := 0$ gesetzt

Grenzen im Definitionsbereich der Zustandsvariablen eingeführt werden können. Die Entropie der Aussage, ein Mann sei mittelgroß gewesen, ist² damit unabhängig von der Frage, ob Männer zwischen 1.0 m und 2.2 m oder zwischen 0.5 m und 2.6 m groß sein können.

3. INFORMATIONSMASS: $H(P)$ ist genau dann maximal, wenn $p_i = 1/n$ für alle n bzw. $\omega(x) = \text{const.}$

Zusammen mit den ersten beiden Eigenschaften ist damit der Hauptzweck der Entropiefunktion sichergestellt: Ihr Wert ist umso kleiner, je sicherer der wahre Zustand des Systemes bekannt ist.

4. MONOTONIE: Seien $X := \{x_1, \dots, x_n\}$ mit $p_i = 1/n$ für alle i und $X^* := \{x_1^*, \dots, x_m^*\}$ mit $p_i^* = 1/m$ für alle i , so gilt: $H(P) > H(P^*) \Leftrightarrow n > m$.

Die Monotonie Eigenschaft ist sozusagen komplementär zur Erweiterbarkeit: Wird dem Zustandsraum ein Element hinzugefügt, welches genauso wahrscheinlich ist wie alle anderen Zustände auch, so ist die Entropie, also die Unwissenheit über den wahren Zustand, größer geworden.

5. STETIGKEIT: $H(P)$ ist stetig in jeder Komponente von P , bzw. stetiges Funktional über $B(\mathbb{R})$.

Die Stetigkeit beschreibt die intuitiv offensichtliche Tatsache, dass sich die Information über den wahren Zustand nur wenig ändert, wenn sich die Wahrscheinlichkeitsverteilung über die möglichen Zustände ebenfalls nur wenig ändert.

6. ADDITIVITÄT: Seien X und Y zwei unabhängige Zustandsräume und P_x, P_y ihre Verteilungen, so gilt: $H(P_{x \otimes y}) = H(P_x) + H(P_y)$.

Die - in physikalischen Anwendungen ausgesprochen wichtige - Bedingung der Additivität (siehe z.B. [61]) ist die Hauptursache für die logarithmischen Terme in der Entropieformel. In vielen informationstheoretischen Fragestellungen wird Additivität allerdings nicht benötigt.

Über die Hartley-Shannon-Entropie hinaus gibt es weitere Entropie-Funktionale mit ähnlichen Eigenschaften, insbesondere solche, die zur Verbesserung der Berechnungseffizienz auf die Additivität verzichten (vgl. [167]). Da im Folgenden der Berechnungszeitaufwand zweitrangig ist, wird auf die HS-Entropie zurückgegriffen.

²nach entsprechender Transformation in eine Wahrscheinlichkeitsaussage

3.1.2. Entropie im Entscheidungskontext

Das Konzept der Entropie wird bei der Modellierung von Entscheidungssystemen schon sehr lange verwandt. Das bereits in Huygens erstem Buch über Wahrscheinlichkeitstheorie "De Ratiociniis in Ludo Aleae" (vgl. [65], im Facsimile und englischer Übersetzung in [57]) im Jahr 1657 angewandte, später von Laplace (vgl. [88, 46]) zum Prinzip erhobene „Bernoulli-Prinzip“ lautet wie folgt:

Definition 27 (Bernoulli-Prinzip/ Prinzip des fehlenden Grundes)

Sind bei einem Versuch wie zum Beispiel dem Wurf eines Würfels mehrere Ausgänge möglich, so sind sie als gleich wahrscheinlich anzunehmen, es sei denn, es liegen Informationen vor, die dem widersprechen.

Zeitgemäß ausgedrückt bedeutet das nichts anderes als das von Jaynes (vgl. [70, 71]) formulierte „Prinzip der maximalen Entropie“:

Definition 28 (Prinzip der maximalen Entropie)

Die für Umweltparameter angenommene Wahrscheinlichkeitsverteilung sollte unter allen sich nicht mit den gegebenen Informationen widersprechenden Verteilungen diejenige mit der größten Entropie sein.

Das Prinzip der maximalen Entropie ist heute eines der vorherrschenden Designparadigmen bei klassischen Entscheidungssystemen und findet dementsprechend vielfältige Anwendungen in allen Bereichen automatischer Entscheidungsfindung und -unterstützung (z.B. [1, 5, 92, 143, 164]).

De Luca und Termini übertrugen das Konzept der Entropie als Maß für den Informationsgehalt einer Funktion auch auf Fuzzy-ZGF bzw. -Entscheidungssysteme ([96]). Da es dort allerdings nicht durch die zugrunde liegende „natürliche“ Erklärung mit dem Bernoulli-Prinzip begründet ist - zumindest, wenn Fuzzy-ZGF nicht als Wahrscheinlichkeitsinformationen interpretiert werden -, findet es heute zwar auch bei solchen Systemen breite Anwendung (vgl. z.B. [123, 148, 151]), konkurriert aber mit einigen ähnlichen Methoden zur Messung der „Unschärfe“ einer Fuzzy-Menge (vgl. z.B. [99, 156, 167, 168]). Durch ihre Anwendung in beiden wichtigen Entscheidungskalkülen ist sie dennoch hervorragend zu einem Vergleich der Kalküle untereinander geeignet.

Der neue Aspekt bei der Verwendung von Entropiemaßen ist nun, dass sie nicht für die Modellierung der Eingangsinformation, sondern für die quantitative Beurteilung von Bewertungsoperatoren und Ordnungsfunktionalen herangezogen

wird.

Damit wird die Frage untersucht, inwieweit sich eine Verbesserung der Eingangsinformation - welche sich nach dem Prinzip der maximalen Entropie in einer Verringerung der Entropie der Informationsfunktionen auswirkt - auch in einer Verbesserung des Ergebnisses niederschlägt. Da die Methoden zur Messung der Ergebnisverbesserung bei Bewertungsoperatoren und Ordnungsfunktionalen unterschiedlich ausfallen müssen - schließlich ist das Ergebnis eines Bewertungsoperators eine Funktion, während das des Ordnungsfunktionalen eine einzelne reelle Zahl ist - werden die Methoden in zwei separaten Abschnitten getrennt behandelt (3.2 und 3.3).

Zunächst müssen jedoch einige Herausforderungen diskutiert werden, die beim kalkülübergreifenden Vergleich von Entropien auftreten.

3.1.3. Herausforderungen bei Verwendung des Entropiefunktionalen

Da das Entropiefunktional (Def. 26) für alle Formen von Informationsfunktionen definiert ist, eignet es sich prinzipiell zum Vergleich der verschiedenen Entscheidungskalküle. Dabei treten allerdings drei Herausforderungen auf:

1. Die absolute Entropie des Ergebnisses der Anwendung des Bewertungsoperators auf gegebene Eingangsfunktionen ist wenig aussagekräftig, da der Bewertungsoperator in einem konkreten Entscheidungsmodell immer auch den Grad von Optimismus ausdrückt, den der Designer des Modells in die Qualität der Eingangsdaten hatte.

Deswegen wird im Folgenden nicht die absolute Entropie des Ergebnisses, sondern nur dessen Sensitivität auf Änderung der Eingangsentropie betrachtet.

2. Die Sensitivität ist abhängig von der Gewichtung der einzelnen Informationen im Entscheidungsproblem. Da in den meisten realen Entscheidungsproblemen verschiedene Parameter auch ein unterschiedliches Gewicht haben, ist zu erwarten, dass die Auswirkungen von Entropieänderungen in verschiedenen Eingangsfunktionen unterschiedlich groß sind. Die Problematik stellt sich im Rahmen der Arbeit allerdings nicht, da alle darin betrachteten Bewertungsoperatoren symmetrisch bezüglich der beiden Eingangsfunktionen sind.

In der Praxis lässt sich diese Beobachtung aber zur Analyse bestehender Entscheidungsmodelle einsetzen: Ist die Sensitivität des Ergebnisses auf

Variationen der Entropie einer Eingangsfunktion sehr niedrig, so ist darüber nachzudenken, ob die explizite Berücksichtigung der Variablen überhaupt sinnvoll erscheint.

3. Der Vergleich der Entropie zweier Funktionen mit unterschiedlichem Inhalt ist, wie im folgenden Beispiel demonstriert wird, nicht immer aussagekräftig.

Beispiel 3 (Entropie und Normierung)

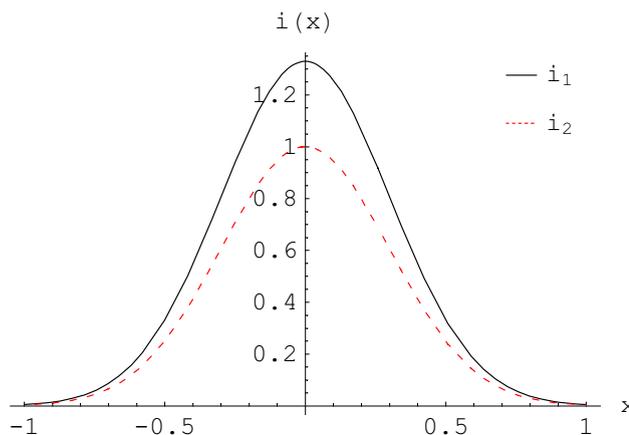
Gegeben sind die beiden auf $x \in [-1, 1]$ definierten Informationsfunktionen (vgl. Abbildung 3.1)

$$i_1(x) := \frac{1}{0.3\sqrt{2\pi}} * e^{-\frac{x^2}{2*0.3^2}} \tag{3.3}$$

$$i_2(x) := e^{-\frac{x^2}{2*0.3^2}}, \tag{3.4}$$

welche als Normalverteilung $N_{0,0.3}$ und deren Renormierung zu einer Fuzzy-ZGF aufgefasst werden können. Interpretiert das Entscheidungsmodell die beiden

ABBILDUNG 3.1: Zwei Informationsfunktionen, die sich nur durch ihre jeweilige Normierung (i_1 auf $\int_{\mathbb{R}} i_1 = 1$; i_2 auf $\max_{\mathbb{R}} i_2 = 1$) unterscheiden.



Funktionen entsprechend, so wird erwartet, dass sie beide denselben Informationsgehalt haben. Die HS-Entropie der beiden Funktionen ist aber sehr unterschiedlich:

$$\mathcal{H}(i_1) = - \int_{-1}^1 i_1(x) \log i_1(x) dx = 0.21497 \tag{3.5}$$

$$\mathcal{H}(i_2) = - \int_{-1}^1 i_2(x) \log i_2(x) dx = 0.37599. \tag{3.6}$$

Daher werden die Ergebnisse der Anwendung des Bewertungsoperators vor der Berechnung der Entropie folgendermaßen normiert:

$$B^*(i_1, i_2) = \frac{B(i_1, i_2)}{\int B(i_1, i_2)}. \tag{3.7}$$

3.2. Entropiesensitivität als neues Vergleichskriterium für Bewertungsoperatoren

Sind Fragen der Berechnungseffizienz außen vor gelassen, so kann sich ein Vergleich von Bewertungsoperatoren prinzipiell auf zwei Eigenschaften beziehen:

- Wie gut beschreibt der Bewertungsoperator das tatsächlich gewünschte Modell?
- Wie gut wird die vorhandene Information über Umweltzustände ausgenutzt?

Während die erste Frage nur im konkreten Einzelfall durch Experimente beurteilt werden kann (z.B. [125]), erschließt sich die zweite Frage einer allgemeinen Untersuchung. Sie ist außerdem von besonderer Bedeutung, da die Genauigkeit der Eingangsinformation in der praktischen Anwendung unmittelbar in Geld zu messen ist, sei es durch die Instrumentierung mit genaueren Sensoren oder die Erweiterung des Expertenwissens durch weitere Interviews etc. Kann an einem konkreten Entscheidungsmodell beobachtet werden, dass eine Verbesserung oder Verschlechterung der Information über einen Parameter keinen Einfluss auf das Endergebnis hat, dann kann die Messung des Parameters auch eingespart oder zumindest auf die kostengünstigste Variante reduziert werden.

Da Bewertungsoperatoren sehr spezifisch für einzelne Entscheidungsprobleme sind, erscheint es zunächst nicht trivial, ein geeignetes quantitatives Vergleichskriterium zu finden. In der Literatur finden sich einerseits eher qualitative Vergleiche (z.B. [162]), bei denen meist die Meinung vertreten wird, dass sich die klassische Theorie insbesondere dann eigne, wenn eine Vielzahl von Informationen (Messwerten) vorliege, während neuere (unscharfe) Methoden in Anwendungen mit weniger Informationen Vorteile hätten. Allerdings werden solche Aussagen kaum quantifiziert.

Andererseits sind sehr konkrete, auf einen spezifischen Anwendungsfall zugeschnittene Vergleiche (z.B. [125]) zu finden, deren Aussagekraft aber stark davon abhängt, ob die gewählte Modellierung in jedem der Kalküle auch tatsächlich die bestmögliche darstellt.

Eine grundlegende Forderung an Bewertungsoperatoren liefert jedoch einen Ansatz zum quantitativen Vergleich verschiedener Kalküle auf theoretischer Ebene:

Forderung (Bewertungsoperatoren und Entropie): Wird die Qualität der Eingangsinformation erhöht, was in der Praxis stets mit höherem technischen

Aufwand und daher mit Kosten verbunden ist, sollte sich auch die Qualität, also insbesondere die Sicherheit des Ergebnisses, erhöhen.

Hierbei wird davon ausgegangen, dass bei der Modellierung des Entscheidungssystems wie auch der Eingangsinformationen stets *sämtliches* verfügbare Wissen eingeht, es also kein zusätzliches a priori Wissen gibt, dass ohne Zusatzkosten zur Verbesserung des Systems genutzt werden kann.

Mathematisch gesehen geht es also um folgende Fragestellung: Seien i_1, i_2 zwei Informationsfunktionen, welche ohne Beschränkung aus \mathcal{J}_I^1 für ein geeignetes Intervall I gewählt und mit

$$\int_I i_1(x) dx = \int_I i_2(x) dx = 1 \quad (3.8)$$

normiert sind.

Seien B ein Bewertungsoperator und $s(i)$ ein beliebiges Störungsmodell. Wenn nun mit

$$H[f(x)] = - \int_I f(x) \log[f(x)] dx \quad (3.9)$$

die Hartley-Shannon-Entropie von $f(x)$ bezeichnet wird, wie verhält sich dann

$$H[B(s(i_1), i_2)] \quad (3.10)$$

bei Änderungen von

$$H[s(i_1)] ? \quad (3.11)$$

Diese Fragestellung legt nahe, als Vergleichskriterium für Bewertungsoperatoren die Entropiesensitivität

$$S_h := \frac{\partial H[B(s(i_1), i_2)]}{\partial H[s(i_1)]} \quad (3.12)$$

heran zu ziehen. Dabei ist zu beachten, dass es sich dabei um einen relativen Vergleichs- und kein absolutes Qualitätskriterium handelt. Selbst das wünschenswerte Vorzeichen von S_h hängt im Allgemeinen vom Anwendungskontext ab (siehe hierzu auch Abschnitt 4.1.3).

Um verschiedene Bewertungsoperatoren unter dem Aspekt der Entropiesensitivität zu untersuchen, werden in Abschnitt 4.1.1 zwei verschiedene Arten Störungsmodelle betrachtet, stochastische und deterministische, und die entsprechenden Entropien für verschiedene Bewertungsoperatoren gegenüber gestellt. Darüber hinaus wird die gleiche Untersuchung für

$$H[B(s(i_1), s(i_2))] \quad (3.13)$$

durchgeführt, um die Sensitivität des Bewertungsoperators auf unterschiedliche Entropien in den Eingangsfunktionen analysieren zu können.

3.3. Störungssensitivität als neues Vergleichskriterium für Ordnungsfunktionale

Die Vergleichsmethodik für Bewertungsoperatoren lässt sich nicht eins zu eins auf Ordnungsfunktionale übertragen, da das Ergebnis der Anwendung eines Ordnungsfunktional keine Funktion ist, für die eine Entropie berechenbar ist, sondern ein einzelner reeller Wert.

In der Literatur findet sich aber eine große Vielzahl anderer Kriterien, nach denen solche Funktionale verglichen werden. Die meisten dieser Kriterien lassen sich jedoch in zwei Gruppen zusammenfassen:

Intuitive Bewertung Die intuitiven Vergleiche (z.B. [38]) beruhen stets darauf, dass zwei ZGF konstruiert werden, deren „korrekte“ Rangfolge für den jeweiligen Autor offensichtlich erscheint, die aber von den gängigen Ordnungsfunktionalen entweder gleich oder gerade umgekehrt bewertet werden. Eine solche Beobachtung wird zum Anlass genommen, ein neues Ordnungsfunktional zu konstruieren, welches genau das betrachtete ZGF-Paar in die richtige Reihenfolge bringt, ohne dabei andere ZGF offensichtlich „falsch“ zu ordnen.

Einem solchen Vorgehen ist entgegen zu halten, dass es in jeder Ordnung von Funktionen Spezialfälle gibt, die nicht unmittelbar intuitiv sind. Das liegt daran, dass ein Ordnungsfunktional als Abbildung

$$\mathcal{B}_+^c(\mathbb{R}) \rightarrow \mathbb{R} \tag{3.14}$$

stets unendlich viele ZGF auf einen einzelnen Punkt abbilden und somit gleich bewerten muss³.

Formale Bewertung Andere Arbeiten (z.B. [93, 133, 158, 159]) vergleichen Ordnungsfunktionale, indem sie die Erfüllung bestimmter Eigenschaften (wie z.B. Translationsinvarianz, Stetigkeit) überprüfen. Die Auswahl eines Ordnungsfunktional für eine bestimmte Anwendung aufgrund wünschenswerter oder weniger wünschenswerter Kriterien wird dabei dem Leser überlassen.

Hierbei ist zu beachten, dass die Auswahl der erwünschten Kriterien insbesondere von der Interpretation der ZGF im konkreten Modell abhängt.

³Hierbei bezeichne \mathcal{B}_+^c den unendlich dimensionalen Vektorraum der positiven Funktionen mit kompaktem Träger.

So demonstriert z.B. Runkler in [133], dass die Forderung von Leekwijck/Kerre ([93]) nach

$$O(\mu) \in \text{Träger}\{\mu\}, \quad (3.15)$$

also dass für den Wert des Ordnungsfunktional nur eine Zahl in Frage kommt, die mit einem Zugehörigkeitswert größer 0 in der zu bewertenden Fuzzy-Menge enthalten ist, nur bei manchen in Regelungssystemen auftretenden Fuzzy-Bewertungen sinnvoll ist, während sie bei typischerweise in Expertensystemen auftretenden ZGF zu unsinnigen Ergebnissen führen kann.

Für die praktische Anwendung ist aber neben solch abstrakten Eigenschaften wie Translationsinvarianz, Skaleninvarianz, Monotonie etc. genau wie bei Bewertungsoperatoren auch die Anfälligkeit eines Entscheidungssystems gegenüber Fehlern in der Eingangsinformation entscheidend.

Es ist schließlich wünschenswert, dass das Ergebnis des Entscheidungsprozesses möglichst unabhängig von statistischen Fehlern in Messwerten oder stets ungenauen Umsetzungen von Expertenaussagen ist. Die Abhängigkeit des Ergebnisses von stochastischen Fehlern der Eingangsinformation kann dabei durch die Varianz der Ergebnisse bei mehrmaliger Anwendung des Ordnungsfunktional auf stochastisch gestörte Eingangsfunktionen gemessen werden. Daher wird vorgeschlagen, die Sensitivität der Entscheidung auf stochastische⁴ Fehler in den Eingangsinformationen explizit als Bewertungskriterium für Ordnungsfunktionale zu verwenden.

Mathematisch gesehen geht es also um folgende Fragestellung: Seien $i(x)$ eine Informationsfunktion, die ohne Beschränkung aus \mathcal{J}_1^1 für ein geeignetes Intervall I gewählt werden kann, und O ein Ordnungsfunktional.

Ist nun s ein beliebiges Störungsmodell, wie verhält sich dann

$$O(s(i)), \quad (3.16)$$

insbesondere wie groß ist

$$\text{Var} [O(s(i))] \quad (3.17)$$

in Abhängigkeit von der Stärke der Störung s ?

Da sich zeigen wird, dass die Varianz nicht nur vom gewählten Ordnungsfunktional, sondern in entscheidender Weise auch von der Form der Informationsfunktion abhängt, wird die Fragestellung in Abschnitt 4.2 für verschiedene Ordnungsfunktionale und Eingangsfunktionen durchgeführt.

⁴Die Sensitivität des Ordnungsfunktional auf deterministische Fehler der Eingangsfunktion zu betrachten ist wenig zielführend, weil es bei separater Betrachtung des Ordnungsfunktional keine deterministisch „falschen“ Eingangswerte gibt.

3.4. Zusammenfassung

Mit den in diesem Kapitel entwickelten Konzept der Verwendung der Störungssensitivität zur Beurteilung von Entscheidungsmodellen steht nun für beide wesentlichen Bestandteile des Funktionalmodells ein neues Vergleichskriterium zur Verfügung, das zum Vergleich der verschiedenen Entscheidungskalküle im nächsten Absatz herangezogen wird.

Die entwickelten Vergleichskriterien werden im folgenden Kapitel auf einige beispielhafte Vertreter der verschiedenen Kalküle angewandt werden, um so eine Aussage über die Informationsnutzung der einzelnen Kalküle ableiten zu können.

4. Ergebnisse numerischer Vergleiche

Nachdem in Kapitel 2 ein einheitliches Funktionalmodell zur Beschreibung von Entscheidungsmodellen entwickelt wurde und in Kapitel 3 für die beiden Hauptbestandteile des Funktionalmodells, Bewertungsoperatoren und Ordnungsfunktionale, jeweils ein Vergleichskriterium entwickelt wurde, das die Effizienz der Informationsausnutzung beschreibt, wird in diesem Kapitel der Vergleich der Entscheidungskalküle durchgeführt.

Der Vergleich wird - für Bewertungsoperatoren (4.1) und Ordnungsfunktionale (4.2) auf Grund der unterschiedlichen Methodik getrennt - anhand der in den Abschnitten 2.3.2.5 bzw. 2.3.3.3 beschriebenen typischen Vertreter der einzelnen Bestandteile vorgenommen. Dabei werden jeweils zunächst die Berechnungsmethode, dann die Ergebnisse der numerischen Simulation dargestellt und schließlich die Ergebnisse inhaltlich interpretiert.

4.1. Bewertungsoperatoren

Es werden die vier in Abschnitt 2.3.2.5 definierten Bewertungsoperatoren (vgl. Tabelle 4.1) miteinander verglichen, indem die Entropie der Eingangsfunktionen auf drei mögliche Arten schrittweise erhöht wird:

1. Die Entropie einer Variablen wird durch Überlagerung mit weißem Rauschen stochastisch erhöht.
2. Die Entropie einer Variablen wird durch Überlagerung mit einer Gleichverteilung „deterministisch“ erhöht (vgl. Abbildung 4.1.1.2).
3. Die Entropie beider Variablen wird durch Überlagerung mit einer Gleichverteilung „deterministisch“ erhöht.

Die hieraus resultierenden zwölf Kurven für mögliche Kombinationen von Bewertungsoperator und Störungsmodell werden im Abschnitt 4.1.2 dargestellt

$$B_{\text{MiWe}}(i_1, i_2)(x) := \frac{i_1(x) + i_2(x)}{2}, \quad (4.1)$$

$$B_{\text{Falt}}(i_1, i_2)(x) := 2 * \int i_1(t) * i_2(2x - t) dt, \quad (4.2)$$

$$B_{\text{ZaPlu}}(i_1, i_2)(x) := \sup_{\frac{x_1+x_2}{2}=x} [\min(i_1(x_1), i_2(x_2))], \quad (4.3)$$

$$B_{\text{Gamma}}(i_1, i_2)(x) := (i_1(x)i_2(x))^{0.5}(1 - (1 - i_1(x))(1 - i_2(x)))^{0.5}. \quad (4.4)$$

TABELLE 4.1: Bewertungsoperatoren

und erläutert. Da der Gamma-Operator bei Erhöhung der Eingangsentropie einer Variablen das zunächst überraschende Verhalten zeigt, die Ausgangsentropie zu verringern, ist ihm ein eigener Abschnitt (4.1.3) gewidmet.

4.1.1. Berechnungsmethode

4.1.1.1. Eingangsfunktionen

Zur Überprüfung ihrer Sensitivität wurden die ausgewählten vier Bewertungsoperatoren auf die beiden Eingangsfunktionen

$$i_1(x) = \frac{1}{0.5\sqrt{2\pi}} * e^{-\frac{(x-4)^2}{2*0.5^2}} \quad (4.5)$$

$$i_2(x) = \frac{1}{0.5\sqrt{2\pi}} * e^{-\frac{(x-6)^2}{2*0.5^2}} \quad (4.6)$$

angewendet, wobei die Funktionen jeweils auf das Intervall $[0, 10]$ beschnitten wurden¹.

Die HS-Entropie der Eingangsfunktionen beträgt im ungestörten Fall

$$\text{HSE}(i_1) := - \int_{\mathcal{X}} i_1(x) \log [i_1(x)] dx \quad (4.7)$$

$$= - \int_{\mathcal{X}} \frac{1}{0.5\sqrt{2\pi}} * e^{-\frac{(x-4)^2}{2*0.5^2}} \log \left[\frac{1}{0.5\sqrt{2\pi}} * e^{-\frac{(x-4)^2}{2*0.5^2}} \right] dx \quad (4.8)$$

$$\approx 0.315 \quad (4.9)$$

$$= \text{HSE}(i_2). \quad (4.10)$$

4.1.1.2. Störungsmodelle

Die drei verschiedenen Störungsmodelle wurden folgendermaßen dargestellt (siehe auch Abbildung 4.1.1.2):

¹d.h. $i(x) = 0 \forall x \notin [0, 10]$

- *Stochastische Entropieerhöhung einer Variablen:*

Zum einen wurde i_1 durch ein normal verteiltes Rauschen überlagert:

$$i^* = |(1 - \epsilon) * i + \epsilon * X_{\text{Norm}(0,1)}| \quad (4.11)$$

$$i_{\text{ges}} = \frac{i^*}{\int i^*(x) dx}. \quad (4.12)$$

Hierbei ist $X_{\text{Norm}(0,1)}$ eine $N_{0,1}$ -verteilte Zufallsvariable, und der Parameter ϵ wurde zwischen 0 und 0.2 variiert.

- *Deterministische Entropieerhöhung einer Variablen:*

Zum anderen wurde die Entropie durch Überlagerung mit einer Gleichverteilung erhöht:

$$i^*(x) = (1 - \epsilon) * i(x) + \epsilon * 0.1 \quad (4.13)$$

$$i_{\text{ges}} = \frac{i^*}{\int i^*(x) dx}, \quad (4.14)$$

wobei hier ϵ zwischen 0 und 1 variiert wurde.

- *Deterministische Entropieerhöhung beider Variablen:*

In einem zweiten Rechenlauf wurde die oben beschriebene deterministische Entropieerhöhung nicht nur auf eine, sondern gleichzeitig auf beide Variablen angewandt.

Als Störungsmodell wurde dabei bewusst eine additive Überlagerung der Störfunktion an Stelle einer multiplikativen gewählt: Für jede Informationsfunktion i und Störungsfunktion σ gilt stets

$$\text{Supp}(i * \sigma) = \text{Supp}(i)$$

wenn S den Träger der Funktion bezeichnet. Daher ist die maximal durch eine multiplikativ überlagerte Störung erreichbare Entropie auch durch die Entropie der Gleichverteilung auf $S(i)$ beschränkt. Hingegen ist bei der gewählten additiven Störung mit dem deterministischen Modell bei Verwendung des Parameters $\epsilon = 1$ die maximal erzielbare Entropie die der Entropie der Gleichverteilung auf dem gesamten Ereignisraum.

Die vier beschriebenen Bewertungsoperatoren wurden auf die verschieden gestörten Eingangsfunktionen angewandt. Die Wertebereiche des Störungsparameters ϵ wurden dabei jeweils so gewählt, dass durch eine weitere Vergrößerung von ϵ keine weitere Erhöhung der HS-Entropie der Eingangsfunktionen mehr zu erzielen war.

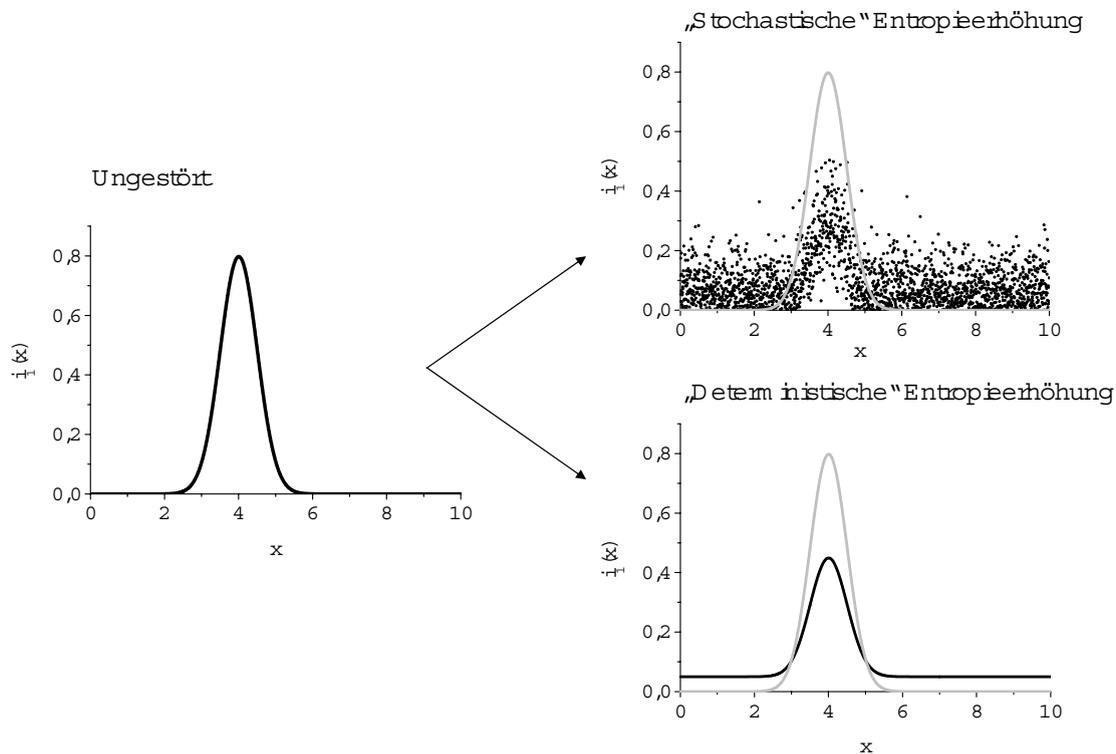


ABBILDUNG 4.1: Verwendete Störungsmodelle für die Eingangsfunktionen des Bewertungsoperators

4.1.1.3. Numerische Berechnung

Die Berechnungen erfolgten mit Mathematica 5.0, wobei die auftretenden Integrale durch numerische Summation über ein Gitter mit Gitterweite $\Delta = 0.0005$ bei stochastischer Störung und Gitterweite $\Delta = 0.005$ bei deterministischer Störung berechnet wurden. Bei allen Entropien erfolgte im ungestörten Fall eine Normierung zu 0. Die verwendeten Quelltexte finden sich im Anhang im Abschnitt C.

4.1.2. Ergebnisse

4.1.2.1. Auswirkungen der verschiedenen Störungsmodelle auf einzelne Operatoren

Die Abbildungen 4.2 bis 4.5 zeigen für jeden der vier betrachteten Bewertungsoperatoren jeweils die Auswirkung der drei verschiedenen Störungsmodelle. Auf der horizontalen Achse ist dabei jeweils die Entropieerhöhung der Eingangsfunk-

tion(en) dargestellt, auf der vertikalen die Entropieerhöhung der Ausgangsfunktion. Im Einzelnen ist folgendes Verhalten der Operatoren erkennbar:

- **FALTUNGSOPERATOR:**

Die Ergebnisse der Simulationsrechnung sind in Abbildung 4.2 dargestellt. Auf die deterministische Störung einer Variablen reagiert der Faltungsoperator mit einer linearen Erhöhung der Entropie der Ausgangsfunktion.

Bei der stochastischen Störung ist die Reaktion etwas stärker. Die Störungssensitivität ist bei diesem Operator also von der Art der vorliegenden Störung abhängig und nicht nur von der Entropie der Eingangsfunktionen. Für die Praxis ist dabei allerdings zu beachten, dass der von der Art der Störung abhängige Faktor selbst dann sehr klein ist, wenn bei einer stochastischen Erhöhung der Entropie um mehr als 0,5 die Eingangsfunktion im Funktionsgraphen schon nicht mehr von vollständigem weißem Rauschen zu unterscheiden ist ($\Delta HSE_{\text{Ausgang}} \approx 0.05$ bei $HSE(a) = 0.5$).

Bei gleichzeitiger Erhöhung der Entropie beider Eingangsvariablen steigt die Ausgangsentropie erwartungsgemäß stärker an, als im Fall nur einer gestörten Variablen und ist leicht nichtlinear.

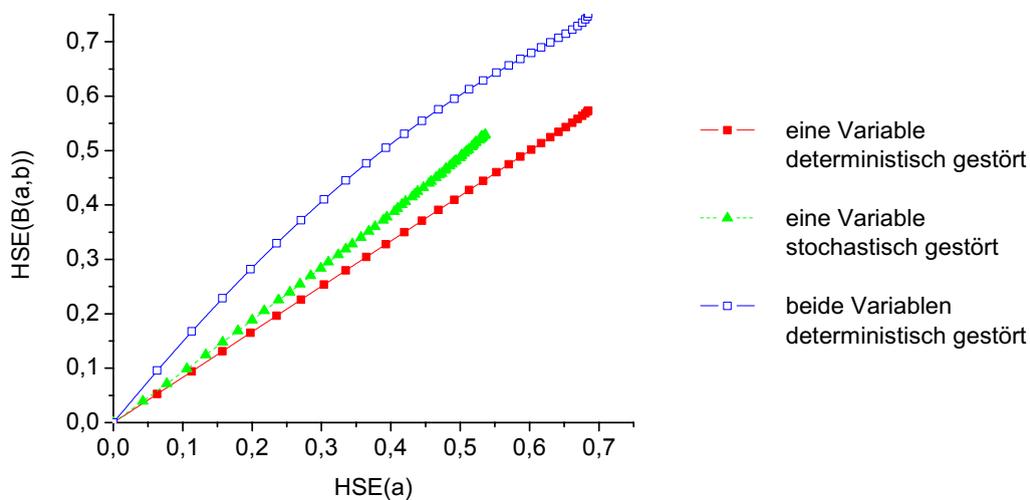


ABBILDUNG 4.2: Anwendung der drei verschiedenen Störungsmethoden beim Faltungsoperator.

- **MITTELWERTOPERATOR:**

Die Ergebnisse der Simulationsrechnung sind in Abbildung 4.3 dargestellt.

Auf die deterministische Störung einer Variablen reagiert der Mittelwertoperator ebenso linear wie der Faltungsoperator, aber insgesamt deutlich schwächer.

Zudem ist bei ihm kein Unterschied in der Reaktion auf deterministische oder stochastische Störung festzustellen. Seine Störungssensitivität ist nicht vom vorliegenden Störungsmodell abhängig.

Auch die Reaktion auf eine Störung beider Eingangsvariablen ist erwartungsgemäß ebenso linear und deutlich stärker als bei nur einer gestörten Variablen, wobei die Werte allgemein unterhalb derer des Faltungsoperators liegen.

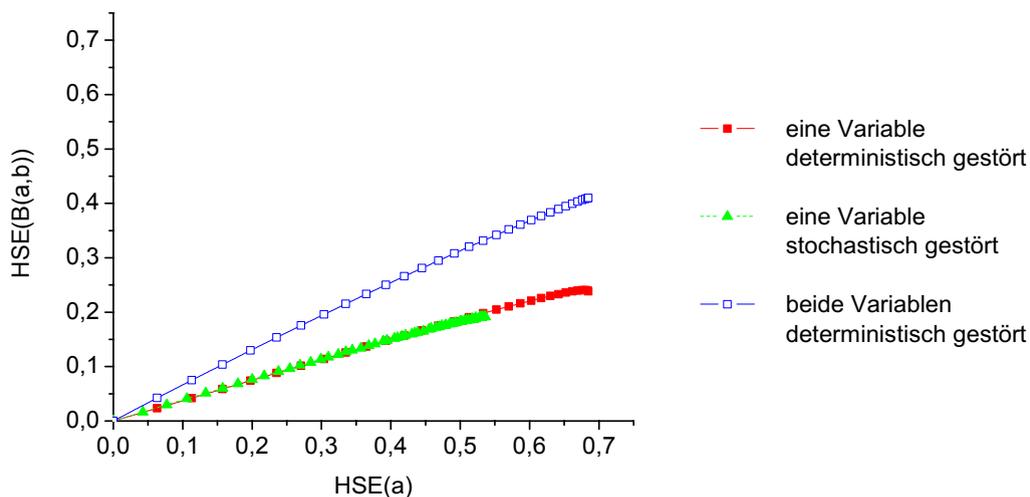


ABBILDUNG 4.3: Anwendung der drei verschiedenen Störungsmethoden beim Mittelwertoperator.

- **ZADEH-PLUS-OPERATOR:**

Die Ergebnisse der Simulationsrechnung sind in Abbildung 4.4 dargestellt. Im deterministischen Störungsmodell wächst die Entropie der Ausgangsfunktion deutlich nichtlinear mit der Entropie der Eingangsfunktion.

Auf eine stochastische Störung reagiert dieser Operator extrem empfindlich. Die Ausgangsentropie steigt schon bei kleinsten Störungen *einer* Variablen durch Rauschen stärker an als bei deterministischer Störung *beider* Variablen. Das ist damit zu erklären, dass die enthaltene Minimumbildung sehr anfällig gegen einzelne Ausreißerwerte ist. Im weiteren Verlauf flacht die Kurve als Folge des gewählten Störungsmodells deutlich ab. Die Erhöhung der Ausgangsentropie scheint gegen einen Wert von etwas über

0,5 zu konvergieren. Das lässt sich dadurch erklären, dass die maximale Entropierhöhung, die im Modell mit weißem Rauschen erzielbar ist, also die Differenz der Entropie der ungestörten Funktionen zur Entropie des weißen Rauschens, bei ca. 0,55 liegt. Daher ist auch die Entropieerhöhung im stochastischen Störungsmodell, bei dem sich die Ausgangsfunktion des Zadeh-Plus-Operators ebenfalls dem weißen Rauschen annähert durch diesen Wert begrenzt. Hierin ist auch die Ursache dafür zu sehen, dass die Sensitivitätskurve für die stochastische Störung konkav, die der deterministischen Störung dagegen leicht konvex ist.

Der Schnittpunkt der betrachteten Kurve mit der Sensitivitätskurve für die deterministische Störung beider Variablen kommt dadurch zustande, dass sich die Ausgangsfunktion in letzterem Modell der Gleichverteilung annähert und somit eine höhere maximale Entropie hat.

Auf die deterministische Störung beider Variablen reagiert der Operator dann auch linear, wobei die Steigung der Sensitivitätskurve annähernd 1 beträgt.

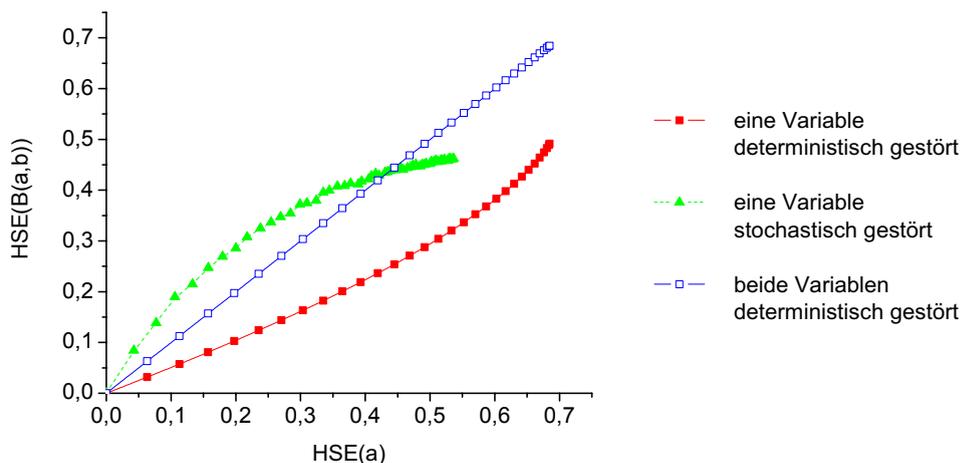


ABBILDUNG 4.4: Anwendung der drei verschiedenen Störungsmethoden beim Zadeh-Plus-Operator.

- GAMMA-OPERATOR:

Die Ergebnisse der Simulationsrechnung sind in Abbildung 4.5 dargestellt. Der Gamma-Operator reagiert bei Störung nur einer Variablen unabhängig vom Störungsmodell gänzlich anderes als die bisher betrachteten Operatoren: An Stelle eines Anstieges der Ausgangs- bei Erhöhung der Eingangsentropie verringert sich diese beim Gamma-Operator. Das verwendete

te Störungsmodell spielt hierbei grundsätzlich keine Rolle, wenngleich sich die Sensitivitätskurven bei den beiden Modellen leicht unterscheiden. Da dieses Verhalten sich grundlegend von dem der anderen Operatoren unterscheidet, wird es in Abschnitt 4.1.3 anhand detaillierterer Berechnungen separat diskutiert.

Bei der deterministischen Störung beider Variablen ist die Reaktion dagegen „normal“, wobei schon eine kleine anfängliche Störung einen deutlichen Entropieanstieg des Ausgangs verursacht, danach aber die Kurve linear ansteigt.

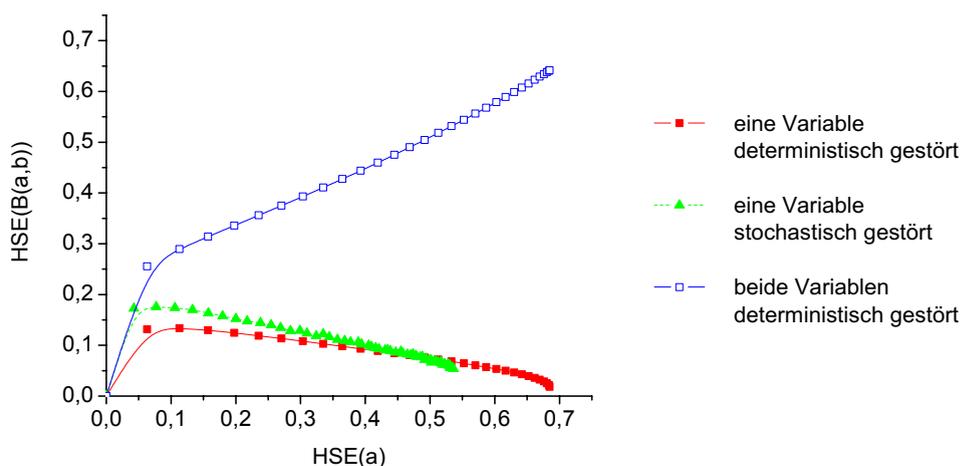


ABBILDUNG 4.5: Anwendung der drei verschiedenen Störungsmethoden beim Gamma-Operator.

4.1.2.2. Vergleich der vier Bewertungsoperatoren

Die Ergebnisse der Berechnungen für die vier Bewertungsoperatoren bei jedem einzelnen Störungsmodell sind in den Abbildungen (4.6 bis 4.8) dargestellt.

Werden zunächst nur die Ergebnisse für den Faltungsoperator B_{Falt} und den Zadeh-Plus-Operator B_{ZaPlu} miteinander verglichen, so ist festzustellen, dass

- bei der Variation der Entropie einer einzelnen Eingangsvariablen die Entropie des Faltungsmittelwertes linear reagiert und zwar unabhängig davon, auf welche Weise die Entropieänderung erzeugt wurde.
- der Zadeh-Plus-Operator sehr stark auf anfängliche stochastische Erhöhungen der Eingangsentropie einer Variablen reagiert, da einzelne Ausrei-

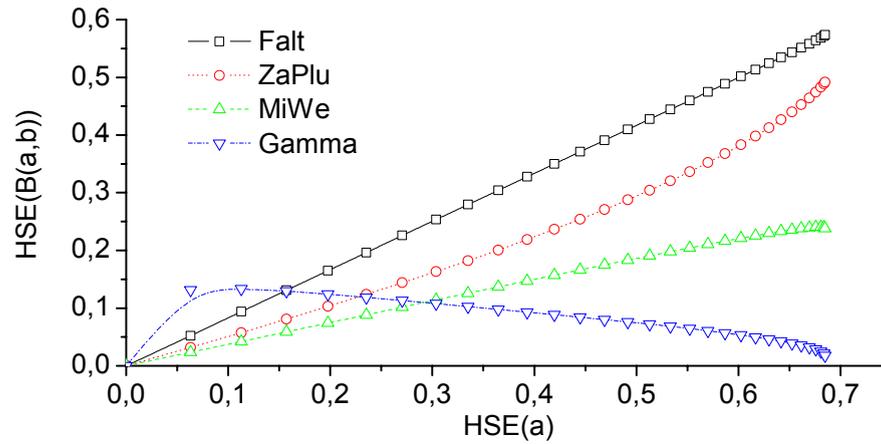


ABBILDUNG 4.6: Vergleich der Operatoren bei deterministischer Erhöhung der Entropie einer Variablen.

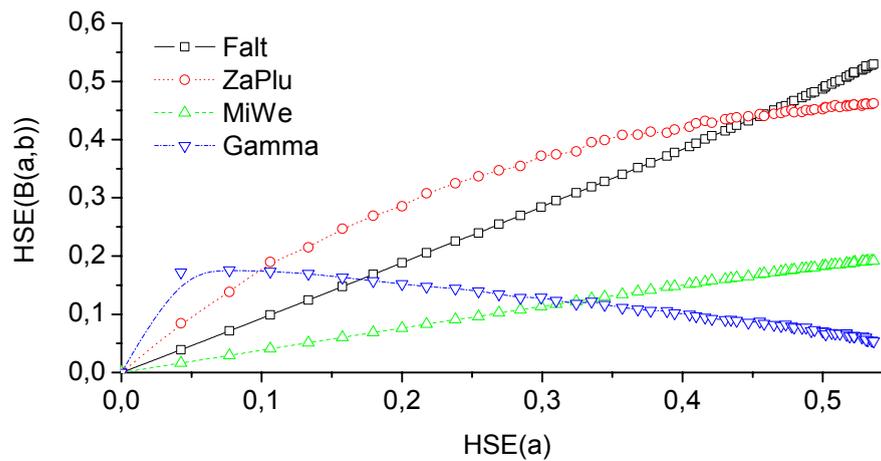


ABBILDUNG 4.7: Vergleich der Operatoren bei stochastischer Erhöhung der Entropie einer Variablen.

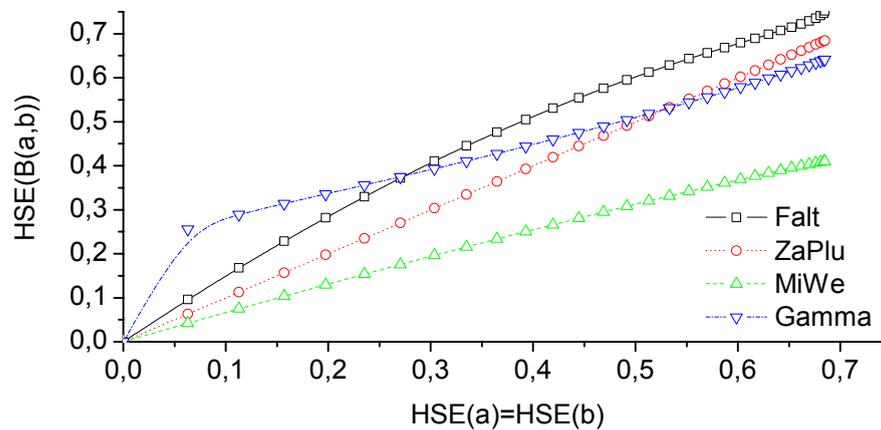


ABBILDUNG 4.8: Vergleich der Operatoren bei deterministischer Erhöhung der Entropie beider Variablen.

über nach oben hier stärkeren Einfluss auf das Ergebnis haben als beim mittelnden Faltungsintegral. Bei der deterministischen Entropieerhöhung reagiert er dagegen deutlich schwächer.

- auch bei gleichzeitiger deterministischer Variation beider Eingangsentropien der Zadeh-Operator etwas schwächer reagiert als der Faltungsmittelwert.
- insgesamt jedoch festzustellen ist, dass innerhalb der natürlichen Schwankung zwischen optimistischer und pessimistischer Interpretation der Messdaten beide Operatoren vernünftige Ergebnisse liefern. Da in der Praxis in vielen Anwendungen davon ausgegangen werden kann, dass eine geringere Entropie in der Eingangsinformation (also besseres Wissen über den wahren Wert des Parameters) durch höheren messtechnischen Aufwand erzielbar ist, kann mit beiden Operatoren die Balance zwischen Genauigkeit (= Kosten) der Eingangsinformation und Sicherheit der Ergebnisse fein gesteuert werden.

Werden zusätzlich die beiden anderen Operatoren in den Vergleich einbezogen, so ist zu erkennen, dass

- die Sensitivität des dritten Operators B_{MiWe} gerade im Fall der stochastischen Entropieerhöhung mit $\frac{\Delta HSE(ein)}{\Delta HSE(aus)} \approx 0.35$ sehr gering ist. Der Einsatz des Operators in praktischen Anwendungen sollte nur bei sehr knapper Rechenleistung in Betracht gezogen werden. Das ist insbesondere auch des-

wegen der Fall, da er dazu neigt, aus zwei Ein-Peak-Eingangsverteilungen eine Zwei-Peak-Verteilung zu erzeugen (siehe Abbildung 2.10, S. 41). Solche Mehr-Peak-Verteilungen sind aber für den nachfolgenden Schritt im Entscheidungssystem, dem Ordnungsfunktional, eher ungünstig.

- der Gamma-Operator bei Störung nur einer Eingangsvariablen unabhängig vom gewählten Störungsmodell grundlegend anders reagiert, indem die Ausgangsentropie bei steigender Eingangsentropie sinkt, während er bei deterministischer Störung beider Variablen eine den anderen drei Operatoren vergleichbare Sensitivitätskurve aufweist.

Die beobachtete Besonderheit des Gamma-Operators wird im nächsten Abschnitt detailliert analysiert.

4.1.3. „Anomalie“ des Gamma-Operators

Der Gamma-Operator verhält sich - insbesondere bei der Variation einer Variablen - ganz anders, als es zunächst erwartet wird: Anstelle einer Erhöhung der Ergebnisentropie verringert sie sich nach einem anfänglichen Anstieg kontinuierlich. In anderen Worten: Je unsicherer die Eingangsinformation für einen der Parameter ist, umso sicherer fällt die Bewertung des Gesamtsystems aus.

Die Ursache hierfür wird aus den in Abbildung 4.9 dargestellten Ergebnissfunktionen der Anwendung des Gamma-Operators auf verschieden gestörte Eingangsfunktionen deutlich:

Für zwei ungestörte, gaußförmige Eingangsfunktionen a und b (dicke graue Linien) ist $B_{\text{Gamma}}(a, b)$ eine symmetrisch zwischen diesen beiden Eingangskurven liegende Kurve (dünne schwarze Linie). Wird nun die im Plot linke der beiden Kurven (a) deterministisch gestört und erhöht somit ihre Entropie von 0,32 auf 0,39, so ist die Ausgangskurve (blau gestrichelt) schon deutlich in Richtung der ungestörten rechten Kurve verschoben. Die Verschiebung wird umso stärker, je mehr die Entropie von a erhöht wird (im Plot mit wachsender Spitze bei $x = 6$). Damit wird erklärbar, warum die Ausgangsentropie mit steigender Eingangsentropie einer Variablen fällt: Je größer der Entropieunterschied beider Eingangsvariablen ist, umso näher liegt die Ausgangsfunktion bei der Eingangsfunktion mit der kleineren Entropie. Wird die Entropie als Unsicherheit über den wahren Wert interpretiert, so bewirkt der Gamma-Operator bei zwei verschiedenen Eingangsfunktionen, dass diejenige mit größerer Sicherheit für richtig gehalten wird, bei der die Eingangsentropie am geringsten ist. Hierbei ist zu beachten,

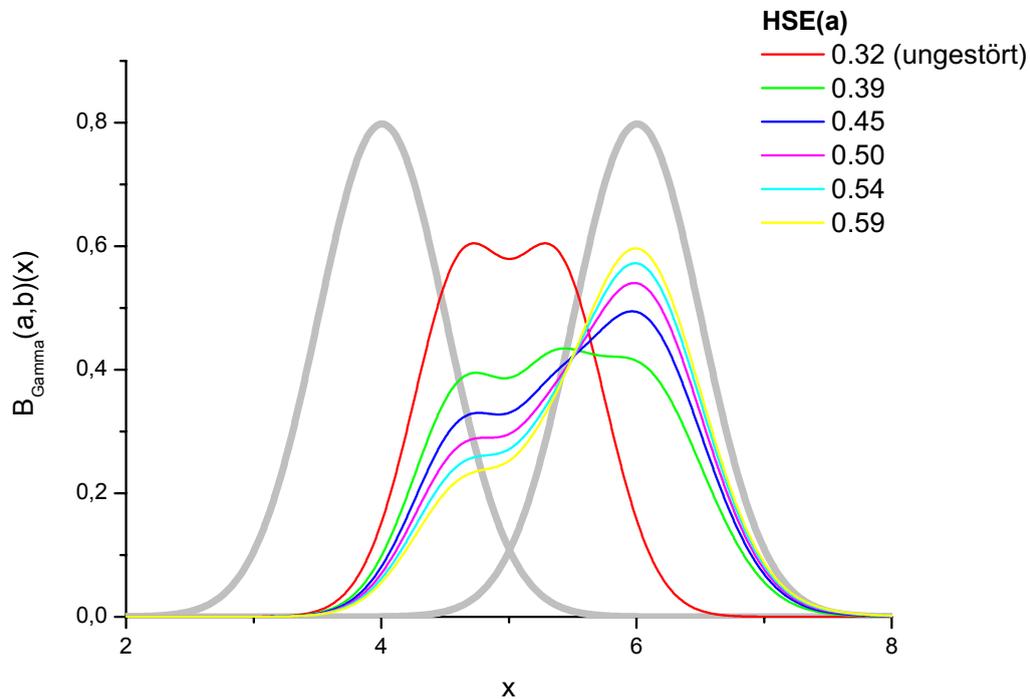


ABBILDUNG 4.9: Ergebnisfunktionen der Anwendung des Gamma-Operators auf zwei gaußförmige Eingangsfunktionen

dass die Erhöhung der Entropie nur einer Eingangsfunktion stets zu einer Vergrößerung der Entropiedifferenz der beiden Funktionen führt.

Die obige Interpretation des Verhaltens des Gamma-Operators wird bestärkt durch eine weitere Beobachtung: Je größer der Unterschied in den Mittelwerten der Eingangsfunktionen ist, um so deutlicher ist die inverse Reaktion auf Erhöhung der Eingangsentropie einer Variablen, wie es aus Abbildung 4.10 ersichtlich ist.

In der Abbildung sind verschiedene Sensitivitätskurven des Gamma-Operators bei deterministischer Störung der Variablen a abgetragen. Dabei wurden die Abstände der Erwartungswerte (bzw. Flächenschwerpunkte) durch horizontale Verschiebungen der Funktion b variiert. Bei $HSE(a)=0.1$ sind von oben nach unten die Kurven für die Erwartungswertdifferenz 2.0, 1.9, ..., 1.1 dargestellt.

Bei allen Kurven ist im Anfangsverlauf zunächst ein „normaler“ Anstieg der Ausgangsentropie bei steigender Entropie von a zu erkennen. Je größer allerdings der Abstand der Erwartungswerte der beiden Eingangsfunktionen ist, bei um so kleinerer Entropiedifferenz fällt die Sensitivitätskurve wieder ab. Bei fast übereinstimmenden Eingangsfunktionen zeigt der Gamma-Operator ein relativ

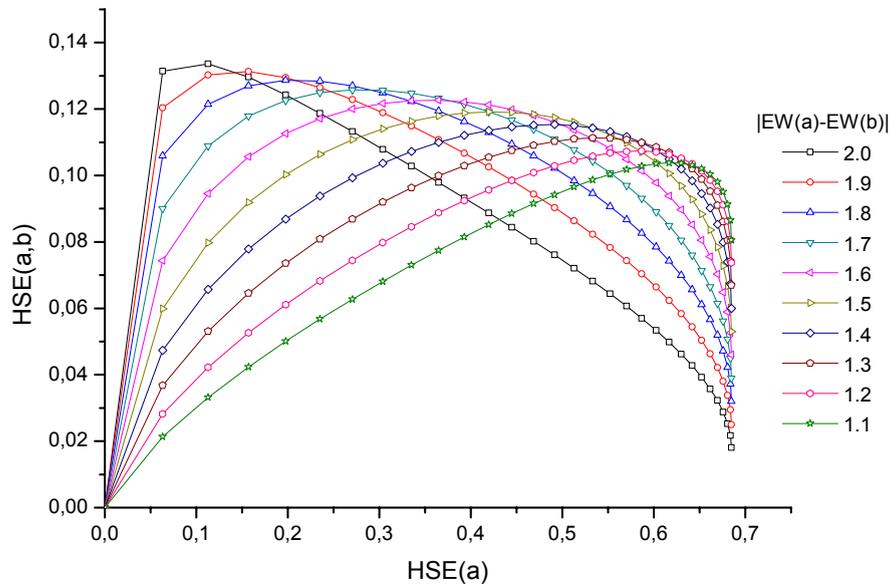


ABBILDUNG 4.10: Sensitivität der Entropie der Ausgangsfunktion des Gamma-Operators auf Änderungen der Entropie einer Eingangsfunktion bei unterschiedlichem Abstand der Erwartungswerte der beiden Eingangsfunktionen

normales Verhalten. Das bedeutet, dass der ungewissere Wert um so eher als falsch angenommen wird, je stärker er vom sichereren Wert abweicht.

Ebenso passt zu der Interpretation des Verhaltens des Gamma-Operators die Tatsache, dass er bei der simultanen Störung beider Variablen ein „normales“ Verhalten zeigt, also auf eine Erhöhung der Eingangsentropie mit einer Erhöhung der Ausgangsentropie reagiert (vgl. Abbildung 4.5). Wenn nämlich die Unsicherheit beider Eingangswerte gleichmäßig steigt, ergibt sich daraus kein Hinweis darauf, dass einer der beiden Eingangswerte wahrscheinlicher richtig ist. Somit lässt sich keine größere Sicherheit für den Ausgangswert ableiten.

Aufgrund seiner speziellen Charakteristik ist der Gamma-Operator besonders geeignet für eine bestimmte Anwendung, in welcher menschliches Verhalten auf Messwerte modelliert wird, wie an folgendem Beispiel zu erkennen ist:

Beispiel 4

„WIDERSPRÜCHLICHE“ MESSWERTE

Da die Ist-Temperatur einen entscheidenden Einfluss auf die Prozessqualität eines gegebenen Prozesses und damit auf die Entscheidungen zur Regelung einer

Heizung hat, wird sie mit zwei voneinander unabhängigen Messgeräten gemessen.

Erhält der Prozessführer bei diesen beiden Messungen nun sehr unterschiedliche Werte, hat aber keine Gelegenheit zu weiteren Kontrollmessungen, so wird er im Zweifelsfall den Mittelwert der beiden als wahre Temperatur annehmen, sich aber darüber sehr unsicher sein, es sei denn er weiß, dass einer der beiden Messwerte wesentlich unzuverlässiger ist als der andere. Je größer die Differenz zwischen den beiden Messwerten und je größer der Unterschied in der Zuverlässigkeit ist, umso mehr wird der Prozessführer dazu neigen, den Wert des zuverlässigeren Messgerätes als wahren Wert anzunehmen und den des anderen Gerätes zu verwerfen.

Werden naheliegenderweise die Zuverlässigkeit der Messgeräte durch eine größere oder kleinere Entropie der von ihnen erzeugten Informationsfunktionen modelliert, so ist zu erkennen, dass der Gamma-Operator genau dieselbe Charakteristik zeigt wie der menschliche Entscheider im Beispiel: Je unzuverlässiger das eine Messgerät und je größer die Messwertdifferenz, umso näher liegt die Ausgangsfunktion an der nicht gestörten Eingangsfunktion.

4.1.4. Vergleich der Bewertungsoperatoren

Über die in den vorangehenden beiden Abschnitten beschriebenen Detailergebnisse lassen sich drei wesentliche Erkenntnisse aus dem durchgeführten Vergleich von Bewertungsoperatoren gewinnen:

1. Mit den betrachteten Eingangsfunktionen zeigen der Zadeh-Plus-Operator und der Faltungsmittelwert als typische Vertreter von FDT und CDT sehr ähnliche Ergebnisse, wobei der erste die Information etwas pessimistischer beurteilt, als der zweite.
2. Der Mittelwertoperator reagiert deutlich schwächer auf Entropieänderungen der Eingangsfunktionen als die anderen Operatoren. In einem hiermit modellierten Entscheidungssystem hat eine Veränderung der Qualität der Eingangsinformation (z.B. durch bessere oder schlechtere Sensorik) die geringste Auswirkung.
3. Der Gamma-Operator zeigt bei den hier betrachteten Beispielen ein grundlegend anderes Verhalten bei Entropieänderungen einer Variablen als die anderen drei und ist daher für spezielle Anwendungen nutzbar, bei welchen

einem Messwert umso mehr vertraut werden soll, je kleiner die Entropie des Messgerätes ist.

Vor allem die dritte Erkenntnis ist überraschend, weil gängige Literatur nicht auf die Besonderheit des Gamma-Operators hinweist und er vielfach ohne spezielle Berücksichtigung seines Verhaltens eingesetzt wird.

4.2. Ordnungsfunktionale

Nach dem Vergleich der Bewertungsoperatoren werden nun die drei im Abschnitt 2.3.3.3 ausgewählten Ordnungsfunktionale miteinander verglichen (vgl. Tabelle 4.2). Teile der hier beschriebenen Ergebnisse sind bereits in [139] veröffentlicht.

$$O_1(i) := \frac{\int_{\mathbb{R}} x * i(x) dx}{\int_{\mathbb{R}} i(x) dx} \quad (4.15)$$

$$O_2(i) := \overline{\left\{ x \in \mathbb{R} \mid i(x) = \max_{y \in \mathbb{R}} i(y) \right\}} \quad (4.16)$$

$$\text{mit } \overline{\{x_1, x_2, x_3, \dots, x_n\}} := \sum_{i=1}^n \frac{x_i}{n}$$

$$O_3(i) := x_0 \left| \int_{-\infty}^{x_0} i(x) dx = 0.7 * \int_{\mathbb{R}} i(x) dx \right. \quad (4.17)$$

TABELLE 4.2: Zum Vergleich herangezogene Ordnungsfunktionale

Um den Vergleich durchzuführen, werden vier verschiedene Eingangsfunktionen stochastisch gestört und die Varianz des Ergebnisses der Anwendung jedes Ordnungsfunktional ausgewertet.

4.2.1. Berechnungsmethode

4.2.1.1. Eingangsfunktionen

Vier typische Informationsfunktionen (IF) werden zum Vergleich der Fehleranfälligkeit der Ordnungsfunktionale herangezogen (vgl. Abbildung 4.11), nämlich

- eine Glockenkurve ($i_1(x) = \frac{1}{1.962} e^{-0.82*(x-5)^2}$),

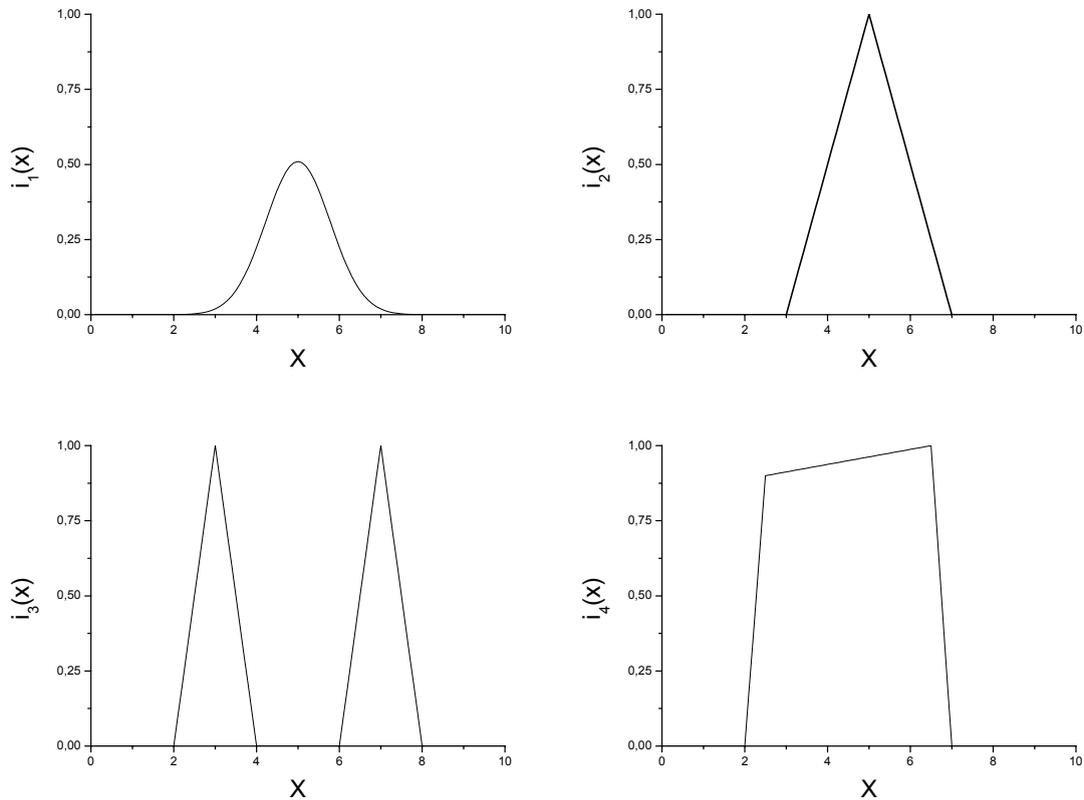


ABBILDUNG 4.11: Vier für den Vergleich der Ordnungsfunktionale herangezogenen Eingangsfunktionen

- eine Dreiecksfunktion (i_2),
- eine Zwei-Peak-Funktion (i_3),
- sowie eine fast trapezförmige Funktion (i_4).

Da die Funktionen als typische Vertreter der klassischen Wahrscheinlichkeitstheorie bzw. der Fuzzy-Theorie ausgesucht wurden, wurde auch ihre Normierung entsprechend gewählt: i_1 ist als klassische Wahrscheinlichkeitsdichte auf $\int_{\mathbb{R}} i_1(x) dx = 1$ normiert, während die anderen drei Funktionen, die regelmäßig im Fuzzy-Kontext auftreten, entsprechend auf $\max_{\mathbb{R}} i(x) = 1$ normiert sind. Die genauen Definitionen der Funktionen finden sich im Anhang B.

Alle IF wurden für die Berechnung auf das Intervall $[0; 10]$ beschnitten, um einen kompakten Träger zu erhalten. Das stellt auch für die Normalverteilung keine wesentliche Einschränkung dar, da der Flächeninhalt des abgeschnittenen Bereichs kleiner als 0.0001 ist.

4.2.1.2. Störungsmodell

Zur Nachbildung der Störung werden die IF nun jeweils mit Gaußschem weißen Rauschen multiplikativ überlagert:

Sei X eine $N_{0;0.2}$ -verteilte Zufallsvariable, so ist

$$\dot{i}_{\text{gestört}}(x) := i(x) * (1 + X). \quad (4.18)$$

Eine solche Störung der Informationsfunktionen entspricht bei den verwendeten Informationsfunktionen einer Erhöhung der Eingangsentropie um etwa 0,1. Wie in Abbildung 4.12 anhand der Eingangsfunktion i_1 dargestellt wird, ist die

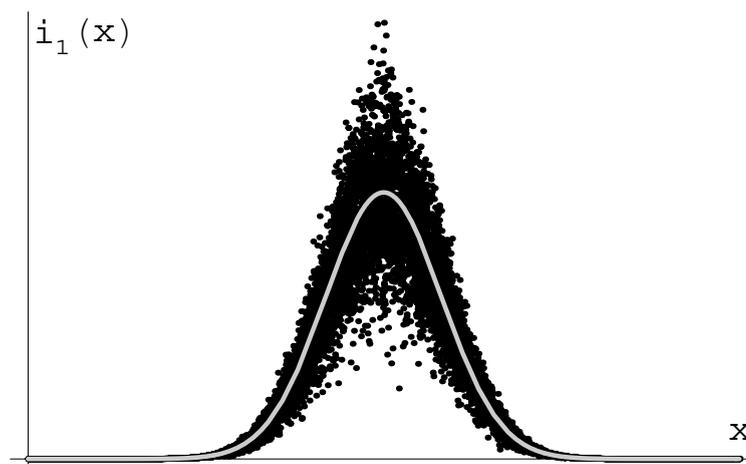


ABBILDUNG 4.12: Visualisierung der Wirkung des angewandten Störungsmodells auf die Eingangsfunktion i_1 (schwarze Punkte).

ursprüngliche Kurvenform zwar in der gestörten Kurve (schwarze Punkte) noch erkennbar, die Information aber schon stark verrauscht.

Durch die Verwendung einer multiplikativen Störung müssen die Informationsfunktionen nicht auf gleiche Maximalwerte normiert werden. Wenn stattdessen - wie in der Informationstheorie üblich - additive Störungen verwendet werden, so ist der relative Fehler bei der auf $\int_{\mathbb{R}} i_1 = 1$ normierten Normalverteilung etwa doppelt so groß wie bei den auf $\max_{\mathbb{R}} i = 1$ normierten anderen Informationsfunktionen. Die mit additiven Störungen auf Grund des Superpositionsprinzips erzielbare Vereinfachung der Berechnungen ist demgegenüber zu vernachlässigen, da der dann notwendige Normierungsschritt stets mehr Rechenleistung benötigt als die Durchführung der Multiplikation. Eine durchgeführte Kontrollrechnung mit einer additiven Störung führte darüber hinaus zu im Rahmen der statistischen Signifikanz gleichen Ergebnissen wie sie Tabelle 4.3 für das multiplikative Störungsmodell zeigt.

Zur Variation der Eingangsentropie bei der genaueren Untersuchung des Maximumoperators in Abschnitt 4.2.2.5 muss die Intensität der Störung variiert werden können. Da das beim multiplikativen Störungsmodell nur in engeren Grenzen möglich ist², wird hier zu einem additiven Modell übergegangen:

$$i_{\text{gestört}}(x) := \frac{|i(x) + (k * X)|}{\int_{\mathbb{R}} |i(x) + (k * X)|}. \quad (4.19)$$

Der Wertebereich des Parameters k ist dabei so gewählt, dass durch eine weitere Vergrößerung keine Erhöhung der Entropie mehr zu erzielen ist.

Ein deterministisches Störungsmodell, wie es zum Vergleich der Bewertungsoperatoren zusätzlich herangezogen wurde, ist nicht zweckmäßig, da eine solche deterministische Störung zwar eventuell zu einem anderen Ergebnis des Bewertungsoperators führt, aber auch bei wiederholter Anwendung stets zum selben. Daher ist die Varianz des Ergebnisses bei deterministischer Störung bei jedem Ordnungsfunktional stets 0.

4.2.1.3. Numerische Berechnung

Die Auswertung wurde numerisch für jede Kombination Informationsfunktion/Ordnungsfunktional 30.000-mal vorgenommen. Die Berechnungen erfolgten mit Mathematica 5.0, wobei die auftretenden Integrale durch Summenbildung über einer Gitterweite von $\Delta = 0.0005$ ermittelt wurden. Die verwendeten Quelltexte finden sich im Anhang in Abschnitt C.

4.2.2. Ergebnisse

4.2.2.1. Vergleich aller vier Eingangsfunktionen bei gleicher Störung

Die Ergebnisse sind in Tabelle 4.3 dargestellt. Für jedes der drei betrachteten Ordnungsfunktionale (Zeilen) und für jede der vier Informationsfunktionen (Spalten) wurden zunächst der Wert des Funktionalwertes für die ungestörte Funktion bestimmt und anschließend Mittelwert μ und Varianz σ des Funktionalwertes für die gestörte Funktion. Die Varianzwerte für die Ordnungsfunktionale O_1 und O_3 liegen dabei in der Größenordnung 10^{-6} .

Die Berechnungsergebnisse werden in den folgenden Abschnitten bezüglich dreier verschiedener Aspekte analysiert:

1. Vergleich der Verwendung von Glockenkurven (CDT) und Dreiecksfunktionen (FDT) zur Repräsentation des Wissens (4.2.2.2).

²da stets gilt $i_{\text{gestört}}(x) = 0$, wenn $i_{\text{ungestört}}(x) = 0$

		i_1		i_2		i_3		i_4	
		ungestört	gestört	ungestört	gestört	ungestört	gestört	ungestört	gestört
O_1	μ	5.00	5.00	5.00	5.00	5.00	5.00	4.54	4.54
	$10^6 \cdot \sigma$	-	2.316	-	2.676	-	27.633	-	6.672
O_2	μ	5.00	5.00	5.00	5.00	5.00	5.01	6.5	5.47
	σ	-	0.023	-	0.008	-	4.002	-	0.790
O_3	μ	5.43	5.43	5.45	5.45	6.90	6.90	5.45	5.45
	$10^6 \cdot \sigma$	-	7.027	-	8.266	-	33.988	-	38.681

TABELLE 4.3: *Ergebnisse der numerischen Berechnungen zur Störungssensitivität von Ordnungsfunktionalen.*

2. Auswirkung von ungünstigen Formen von Informationsfunktionen wie solchen mit mehreren Maxima oder sehr flachen Kurven (4.2.2.2).
3. Vergleich aller drei Ordnungsfunktionale insbesondere unter Berücksichtigung der Unterschiede zwischen integralbasierten-Methoden (O_1 und O_3) und extremwertbasierten (O_2) (4.2.2.2).

Da sich zeigt, dass eine signifikante Störungsanfälligkeit nur bei der Maximum-Ordnung festzustellen ist, wird die Wirkung der Maximumordnung auf die verschiedenen Eingangsfunktionen in einem weiteren Abschnitt (4.2.2.5) vertieft betrachtet.

4.2.2.2. Vergleich von Dreiecks- und Glockenkurve zur Repräsentation des Wissens

Im Vergleich der Ergebnisse für i_1 und i_2 , welche jeweils die typischen Formen der Darstellung von Wissen in der CDT und FDT sind, ist zu erkennen, dass

- die beiden integral-basierten Ordnungsfunktionale O_1 und O_3 eine um mehrere Größenordnungen kleinere Störanfälligkeit aufweisen als die Maximum-Methode.

- es nur bei der Maximum-Methode einen merklichen Unterschied der Störanfälligkeit zwischen Dreiecks- und Glockenform gibt, welcher durch die kleinere Ableitung von i_1 in der Umgebung des Maximalpunktes zu erklären ist.
- selbst die Varianz der Maximum-Methode weniger als 0,5 Prozent beträgt und damit in vielen einfachen regelungstechnischen Anwendungen unterhalb der Sensorgenauigkeit liegt. So hat z.B. ein Platin-Widerstandsthermometer wie der häufig verwendete Typ Pt100 der Klasse A nach DIN EN 60751 in diesem Bereich nur eine Genauigkeit von $\pm 0,16$ also einen maximalen Fehler von ca. 3 Prozent.

Insgesamt ist festzustellen, dass aus der Störanfälligkeit von Ordnungsfunktionalen keine Präferenz für das eine oder andere Kalkül abgeleitet werden kann.

4.2.2.3. Ungünstigere Informationsfunktionen

Für Funktionen vom Typ i_3 oder i_4 zeigt sich ein grundlegend anderes Bild. Während die Varianz unter O_1 und O_3 zwar signifikant größer, aber letztendlich immer noch vernachlässigbar ist, erweist sich die Maximum-Methode als unbrauchbar:

- Die Eingangsfunktion i_3 hat zwei Maxima bei 3 und 7 in deren Umgebung der Graph jeweils steil ansteigt bzw. abfällt. Durch Addition³ eines zufälligen Wertes an jeder Stelle der Funktion wird der Funktionswert an einer der beiden Maximalstellen etwas größer werden, als an der anderen. Daher springt die Bewertung stochastisch zwischen ~ 3 und ~ 7 . Die Varianz des Ergebniswertes ist daher 4.

Die in der Ergebnistabelle dargestellte Verschiebung des Mittelwertes ist auf Grund Ungenauigkeiten bei der numerischen Berechnung entstanden und nicht signifikant.

- Da die i_4 (als einzige der betrachteten Funktionen) nicht symmetrisch um ihr einziges Maximum ist, ergibt sich durch die Störung eine systematische Verschiebung der Bewertung. Weil zudem die Ableitung links des Maximums deutlich kleiner ist als rechts davon, verschiebt sich der Wert nach links. Ein Vergleich mit einer IF anderer Gestalt ist somit nicht mehr möglich.

Bei Verwendung eines additiven Störungsmodells ergibt sich eine ähnliche

³oder Multiplikation wenn ein multiplikatives Störungsmodell verwendet wird

Verschiebung auch bei den anderen beiden betrachteten Ordnungsfunktionalen, da sich der Schwerpunkt der Fläche hin zum Mittelpunkt des Integrationsintervalls verschiebt.

Solange die Eingangsfunktionen für das Ordnungsfunktional nicht glocken- oder dreiecksförmig sind - in der praktischen Anwendung in Entscheidungssystemen wird dies selten der Fall sein, weil die Eingangsfunktion ihrerseits Ergebnis eines Bewertungsoperators ist - kann eine ungünstige Wahl des Bewertungsoperators deshalb die Qualität der Entscheidung sowohl systematisch (Verschiebung des Erwartungswertes) als auch stochastisch (Erhöhung der Varianz) negativ beeinflussen.

4.2.2.4. Vergleich der Ordnungsfunktionale miteinander

Im Vergleich der drei Ordnungsfunktionale miteinander ist das Folgende zu erkennen:

- Die beiden integralbasierten Funktionale O_1 und O_3 zeigen unabhängig von der gewählten Form der Eingangsfunktion eine sehr geringe Sensitivität gegen Störungen der Eingangsinformation.
- Die Maximum-Methode ist schon bei „günstigen“ Eingangsfunktionen wie i_1 und i_2 um vier Größenordnungen störanfälliger. Allerdings ist auch hier der Fehler nur knapp oberhalb der Rechengenauigkeit bei einer Integrationsgitterweite von 0,0005 und ist daher bei vielen praktischen Anwendungen zu vernachlässigen.
- Bei „ungünstigeren“ Eingangsfunktionen wie i_3 und i_4 , die keine einzelne klare Spitze haben, liefert die Maximum-Methode allerdings sehr stark gestörte, teilweise sogar unbrauchbare Ergebnisse.

Der wesentlich höhere Berechnungsaufwand einer integral-basierten Methode - die Berechnung von O_1 dauert etwa 10-mal so lang wie die von O_2 - lässt sich daher kaum mit einer geringeren Störanfälligkeit rechtfertigen, solange nur IF vom Typ i_1 oder i_2 miteinander verglichen werden sollen.

Es zeigt sich somit, dass die Verwendung einer Maximum-Ordnung ohne genaue Kenntnis der Gestalt der IF zwei Gefahren birgt: Zum einen ergeben sich systematische Verschiebungen der Bewertung zu (Un-)Gunsten bestimmter Formen von IF. Diese Verschiebung ist um so größer, je flacher die Eingangsinformation

auf einer Seite des Maximalwertes und je größer die Steigungsdifferenz auf beiden Seiten ist.

Zum anderen droht eine hohe Varianz der Ergebnisse, welche in folgendem Abschnitt genauer untersucht wird.

4.2.2.5. Auswirkung der Störungsintensität beim der Maximum-Methode

Da die Maximum-Methode als Einzige mit einer wesentlichen Varianz des Ergebnisses auf eine stochastische Störung der Eingangsfunktion reagiert, wird für diese die Abhängigkeit zwischen Störungsintensität und Ergebnisvarianz detaillierter betrachtet. Eine solche Analyse führt für die anderen beiden Ordnungsfunktionale zu keinem wesentlichen Ergebnis, da dort die Varianz in der Größenordnung $10^{-5} - 10^{-6}$ und somit im Bereich der verwendeten Integrationsgenauigkeit liegt.

Um die Abhängigkeit zu untersuchen, wird das in Gleichung (4.17) beschriebene Störungsmodell leicht abgewandelt. An Stelle von

$$i_{\text{gestört}}(x) := i(x) * (1 + X), \quad (4.20)$$

wobei X eine $N_{0,0.2}$ -verteilte Zufallsvariable ist, wird nun das etwas aufwändiger zu berechnende Störungsmodell

$$i_{\text{gestört}}(x) := \frac{|i(x) + (k * X)|}{\int_{\mathbb{R}} |i(x) + (k * X)|} \quad (4.21)$$

verwendet, welches es ermöglicht, durch Variation des Parameters k in einem Bereich zwischen 0 und 1 die Entropie der gestörten Eingangsfunktion fein zu steuern.

In Abbildung 4.13 ist für die drei Eingangsfunktionen i_1 , i_2 und i_4 jeweils die Varianz der Ausgangswerte in Abhängigkeit der durch die Störung verursachten Entropieerhöhung dargestellt. Eine vergleichbare Darstellung für die Eingangsfunktion mit zwei Spitzen, i_3 , ist nicht möglich, da die Varianz des Ergebnisses hier schon bei kleinsten Störungen bei etwa 4 liegt. Dies ist damit zu erklären, dass der Wert von $O_2(i_{3,\text{gestört}})$ bei einer minimalen Störung zufällig gleichverteilt zwischen ~ 3 und ~ 7 springt. Aus den Kurven ist Folgendes abzulesen:

- Bei den günstigen Formen der Eingangsfunktion (Dreiecks- und Glockenform) führt eine schwache Störung zu keiner nennenswerten Erhöhung der Varianz der Werte des Funktionals. Erst wenn die Eingangsfunktion fast vollständig von weißem Rauschen überlagert ist, ihre Entropie also der des weißen Rauschens entspricht, steigt auch die Varianz drastisch an.

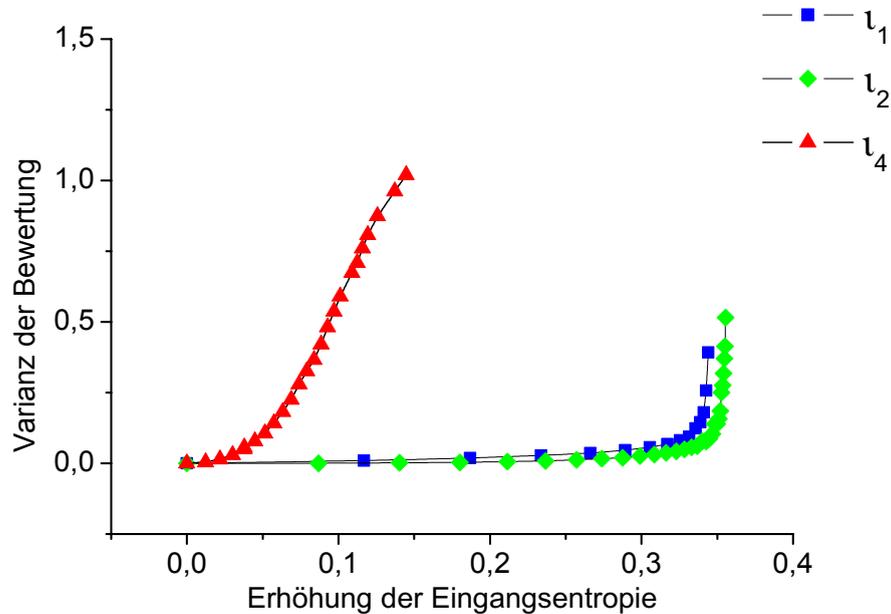


ABBILDUNG 4.13: Sensitivität des Maximumoperators auf Erhöhung der Entropie der Eingangsfunktion.

- Es lässt sich kein signifikanter Unterschied zwischen Glocken- und Dreiecksform ableiten, wobei die leicht höhere Varianz bei gleicher Entropieerhöhung mit dem flacheren Kurvenverlauf der Glockenkurve im Bereich des Maximums zu erklären ist.
- Bei der fast trapezförmigen Eingangsfunktion steigt die Varianz schon bei einer geringen Störung der Eingangsfunktion stark an, da deren Graph auf der linken Seite des Maximums über einen weiten Bereich relativ flach verläuft, was die Wahrscheinlichkeit erhöht, dass einer der Funktionswerte in diesem Bereich durch die Störung über das Maximum erhöht wird.
- Für die Eingangsfunktion mit zwei Maxima ergibt sich schließlich, dass schon kleinste Störungen die Varianz sprunghaft auf einen Wert von ca. 4 ansteigen lassen. Die Ursache hierfür ist, dass das Maximum der gestörten Funktion stets in der Umgebung *einer* der beiden Spitzen liegt.

4.2.3. Vergleich der Ordnungsfunktionale

Die vorliegenden Ergebnisse haben über die reine Bewertung der konkreten Funktionale hinaus Folgendes veranschaulicht:

- Die Unterscheidung zwischen klassischen Entscheidungs- und Fuzzy-Modellen ist für viele Fragestellungen unerheblich. Unter in der Praxis vernachlässigbaren Einschränkungen (insb. einer Kompaktifizierung des Trägers der IF) sind Methoden der klassischen Wahrscheinlichkeitstheorie auf allgemeine Fuzzy-ZGF übertragbar, da die verwendeten Funktionale letztendlich unter verschiedenen Namen weitgehend identisch sind.
- Die Vereinfachung, an Stelle von integral-basierten Methoden Extremwert-Methoden zu verwenden, geht stets mit einer deutlichen Zunahme der Fehleranfälligkeit des Gesamtsystems einher. Ob die erhöhte Fehleranfälligkeit im konkreten Fall tolerabel ist, hängt entscheidend davon ab, welche Gestalt die zu vergleichenden IF haben: Je deutlicher die Informationsfunktion ein einzelnes Maximum aufweist und je steiler der Funktionsgraph im Bereich dieses Maximums ist, um so kleiner ist die Auswirkung auf die Störungssensitivität.
- Die Repräsentation des Wissens durch eine Dreiecksfunktion an Stelle der im Sinne einer statistischen Interpretation von ZGF eigentlich angebrachten Gaußschen Glockenfunktion spielt bei den hier betrachteten Eigenschaften keine Rolle. Das entspricht Ergebnissen, wonach der statistische Fehler in der Schätzung der Varianz bei praktisch relevanten Stichprobenumfängen größer ist als der Fehler durch partielle Linearisierung der Funktion (vgl [6]).

Bei der Optimierung der Recheneffizienz des Ordnungsfunktionals kommt es also insbesondere darauf an, ob in den vorhergegangenen Stufen des Entscheidungsmodells sicher gestellt werden kann, dass die resultierenden Bewertungen der einzelnen Alternativen hinreichend gutartig sind, also im besten Fall Dreiecks- oder Glockenfunktionen.

Insgesamt zeigt sich: Werden zur Modellierung des Fuzzy-Systems die „richtigen“ Methoden gewählt, wird also insbesondere auf die Verwendung von Extremwert-Methoden verzichtet, wenn die Form der Eingangsfunktionen nicht bekannt oder nicht kontrollierbar ist, so gibt es keinen qualitativen Unterschied zwischen unscharfen und statistischen Modellen.

Hierdurch wird ein in der Literatur zu beobachtender Trend (z.B. [10]) bestätigt:

Ist die Genauigkeit der Entscheidungsfindung (im Gegensatz z.B. zur maximalen Berechnungseffizienz) wichtigstes Kriterium bei der Entwicklung eines Entscheidungsmodells, so nähern sich die verwendeten Fuzzy-Methoden immer stärker denen der Wahrscheinlichkeitstheorie an.

Die Fuzzy-Entscheidungstheorie ist allerdings der klassischen Theorie insofern überlegen, als dass sie es ermöglicht, verschiedene Genauigkeitsstufen in einem Modell zu vereinen und dass über die Regelbasis eine Interpretierbarkeit der Ergebnisse möglich wird.

5. Zusammenfassung und Ausblick

5.1. Allgemeine Erkenntnisse

Im Rahmen der Arbeit hat sich gezeigt, dass sich beim Versuch, unterschiedliche Entscheidungskalküle mathematisch-quantitativ zu vergleichen, drei prinzipielle Herausforderungen stellen:

- Das Fehlen einer etablierten Methodik bzw. eines anerkannten Katalogs von Kriterien zum Vergleich von Entscheidungsmodellen und -kalkülen.
- Der Mangel an einer ausreichenden Kanonisierung im Bereich der Fuzzy-Entscheidungstheorie und neuerer Methoden, welcher einerseits eine Abgrenzung der Methoden zur klassischen Theorie erschwert und andererseits den Betrachtungsgegenstand „Fuzzy-Entscheidungstheorie“ mit ihren zahlreichen derzeit in der Diskussion befindlichen Methoden sehr unhandlich erscheinen lässt.
- Die starke Vorprägung zu Gunsten der klassischen Wahrscheinlichkeitstheorie durch die traditionelle naturwissenschaftliche Ausbildung. Hierdurch erscheinen deren Konzepte sehr intuitiv, wohingegen es schon schwierig ist, den Informationsgehalt einer Fuzzy-ZGF anschaulich zu beschreiben, ohne auf Vokabeln der Wahrscheinlichkeitstheorie zurückzugreifen.

Fehlen einer etablierten Vergleichsmethodik

Ein quantitativer Vergleich von Entscheidungsmodellen anhand ihrer Ergebnisse ist - wie in Abschnitt 1.2.3 dargestellt - selbst nach deren Implementierung nur in wenigen Spezialfällen möglich, so z.B. bei sehr repetitiven Entscheidungen wie in Regelungssystemen oder wenn durch ein theoretisch verifizierbares Prozessmodell eine Simulation entsprechend vieler Entscheidungen möglich ist. Wird das Entscheidungssystem für wenige Einzelentscheidungen eingesetzt, wie z.B. bei typischen Expertensystemen, ist im Allgemeinen nicht einmal nach der

getroffenen Entscheidung klar, ob das System gut oder schlecht funktioniert hat. Daher haben verschiedene Autoren - teilweise aus praktischen Anforderungen ihrer speziellen Fragestellung heraus - die unterschiedlichsten Vergleichskriterien vorgeschlagen und angewandt. Die wenigen Entscheidungsmodell und -kalkül übergreifenden Kriterien beschränken sich weitgehend auf Fragen der Implementierbarkeit und Komplexität, kaum auf solche der Entscheidungsqualität.

Um aber nicht nur verschiedene Entscheidungskalküle miteinander vergleichen, sondern auch innerhalb der einzelnen Kalküle neue Ideen und Methoden zuverlässig beurteilen zu können, bedarf es eines allgemein anerkannten Kriterienkataloges, der sowohl Fragen der Implementierung und Berechnungseffizienz als auch der Wirtschaftlichkeit und Ergebnisqualität mit einschließt.

Das hier neu entwickelte Vergleichskriterium kann Bestandteil eines solchen Kataloges sein.

Mangelnde Kanonisierung der Fuzzy-Entscheidungstheorie

Verglichen mit der klassischen Entscheidungstheorie, die spätestens seit Mitte der 40er Jahre des 20. Jahrhunderts als eigenständiges wissenschaftliches Untersuchungsfeld etabliert ist und auf eine mehrere hundert Jahre hindurch entwickelte mathematische Basis aufbaut, ist die Fuzzy-Entscheidungstheorie eine sehr junge Disziplin. Sie hat sich zwar auf Grund der großen Akzeptanz, insbesondere in der regelungstechnischen Anwendung, sehr schnell entwickelt, trotzdem fehlt es ihr noch an einer gewissen Stabilisierung. Daher ist es derzeit kaum möglich, eine engere Abgrenzung des Begriffes „Fuzzy-Entscheidungssystem“ zu geben als: *„Ein Entscheidungssystem, bei welchem Fuzzy-Zugehörigkeitsfunktionen irgendwie eine Rolle spielen.“*

Viele inhaltliche Fragen, deren Beantwortung häufig zum erfolgreichen praktischen Einsatz nicht einmal notwendig sind, wie zum Beispiel der zentrale Streit um den Bedeutungsinhalt von Zugehörigkeitsfunktionen, werden teilweise leidenschaftlich diskutiert. Diese Situation erinnert ein wenig an den Mitte der 70er Jahre sehr engagiert geführten Streit zur Rationalität des Bernoulli-Prinzips in der klassischen Entscheidungstheorie¹.

Ein Vergleich *der* klassischen mit *der* Fuzzy-Entscheidungstheorie muss sich entweder auf sehr allgemeine und damit wenig aussagekräftige Kriterien beschränken oder aus der Vielzahl der diskutierten Konzepte der Fuzzy-Entscheidungstheorie diejenigen auswählen, von denen wahrscheinlich ist, dass

¹siehe hierzu z.B. [28, 91, 140]

sie auch in einigen Jahren noch zum dann etablierten Kanon der Disziplin gehören werden. Eine umfassende Betrachtung aller in der Literatur publizierten Vorschläge scheint ausgeschlossen.

Vorprägung zu Gunsten klassischer Entscheidungstheorie

Bei der praktischen Durchführung und insbesondere bei der Formulierung der Ergebnisse eines Vergleiches, wie er hier unternommen wurde, ergibt sich eine weitere Herausforderung:

Das in der naturwissenschaftlichen Grundbildung deutlich vorherrschende Modell zur Beschreibung unsicherer Zustände ist das Wahrscheinlichkeitskalkül. Fuzzy-Methoden werden, wenn überhaupt, erst in Spezialvorlesungen im Hauptstudium unterrichtet und selbst dann fast ausschließlich sehr anwendungsbezogen im regelungstechnischen Kontext. Die mathematischen Grundlagen der Fuzzy-Entscheidungstheorie, welche im klassischen Fall in Form von Maß- und Integrationstheorie fast schon zum Standardkanon der mathematisch-naturwissenschaftlichen Ausbildung gehören, bleiben einigen Spezialisten vorbehalten.

Das hat nicht zuletzt Auswirkungen auf die Sprache, in welcher über entscheidungstheoretische Zusammenhänge nachgedacht wird: Werden Praktiker darüber befragt, was z.B. eine dreiecksförmige Zugehörigkeitsfunktion verschiedener Längenmaße zur linguistischen Aussage „großer Mann“ bedeutet, dann werden fast immer Begriffe der Wahrscheinlichkeitstheorie genannt. So betrachtet bleibt aber von der FDT nicht viel mehr übrig als eine scheinbar um einige mathematische Zwänge bereinigte Variante des klassischen Kalküls, welche insbesondere den praktischen Anwender von mühseligen Integraltransformationen und Konvergenzüberlegungen befreit.

Zwar könnte in der Anwendung noch damit gelebt werden, dass durch eine solche Beschränkung auf eine mögliche Interpretation eben nur ein Teil des Potentials der unscharfen Theorie genutzt wird, insoweit sich das praktische Problem schon damit befriedigend lösen lässt. Für einen umfassenden Vergleich des Potentials der verschiedenen Kalküle ist solch eine Beschränkung natürlich nicht akzeptabel. Um sie zu vermeiden, muss deswegen auch und gerade bei der Wortwahl darauf geachtet werden, nicht bestimmte Interpretationen zu präjudizieren, wenn sie nicht notwendigerweise für ein bestimmtes Kalkül benötigt werden.

Das hat natürlich den Nachteil, dass einige Formulierungen eher vage bleiben müssen und es z.B. sehr schwierig ist, umfassend zu beschreiben, was die in einer „Informationsfunktion“ enthaltenen Informationen genau bedeuten. Solch

eine inhaltliche Interpretation ist daher immer nur im Kontext eines bestimmten Entscheidungsproblem und -modells möglich.

5.2. Wichtigste Ergebnisse der Arbeit

Ziel der Arbeit war es, verschiedene Entscheidungskalküle bezüglich ihrer Effizienz in der Nutzung vorhandener Information zu vergleichen. Um einen solchen Vergleich zu ermöglichen, wurde in drei Schritten vorgegangen. Zunächst wurden ein neues übergreifendes Modell für Entscheidungssysteme entwickelt und die bestehenden Kalküle im neu entwickelten Modell beschrieben. Im nächsten Schritt wurde eine allgemeine Methodik entwickelt, um die Informationsnutzung von Entscheidungssystemen zu messen, und schließlich wurde die Methodik auf typische Vertreter der einzelnen Entscheidungskalküle angewendet. Die wesentlichen Ergebnisse werden im Folgenden kurz dargestellt.

1. Entwurf eines allgemeinen Funktionalmodells für Entscheidungssysteme
Aufbauend auf in der Literatur bestehenden Ansätzen, insbesondere [112], wurde in Kapitel 2
 - ein neues mathematisches Modell für Entscheidungssysteme dargestellt, welches unabhängig vom gewählten Entscheidungskalkül ist. Hierzu war es insbesondere notwendig, mit dem neu eingeführten Begriff der „Informationsfunktion“ einen interpretationsunabhängigen Formalismus für die mathematische Beschreibung der Eingangsinformation in ein solches Modell zu etablieren.
 - anhand eines Beispiels (Beispiel 1) demonstriert, dass dieses mathematische Modell exakt zu dem natürlichen Vorgehen bei der Modellierung eines Entscheidungssystems passt und somit in der praktischen Anwendung keinen Zusatzaufwand erfordert.
 - schließlich gezeigt, wie sich übliche Strukturen aus verschiedenen Entscheidungskalkülen im entwickelten Modell darstellen und somit der Nachweis erbracht, dass das Modell tatsächlich in der Lage ist, als einheitlicher Rahmen für die Beschreibung von Entscheidungsmodellen zu dienen.
2. Entwicklung einer Methodik zur Messung der Informationseffizienz
Um die Effizienz der Informationsnutzung durch verschiedene Kalküle vergleichen zu können, wurde in Kapitel 3

- ein informationstheoretischer Ansatz zur quantitativen Messung der Informationsnutzung entwickelt. Eine solche auf der Entropie basierende Methode ist neu und findet sich in der bekannten Literatur derzeit nicht.
 - gezeigt, dass der Ansatz auch bei Verwendung unterschiedlicher Informationsdarstellungen durch die verschiedenen Kalküle zu aussagekräftigen Ergebnissen führt.
 - dargestellt, wie die entwickelte Methodik nicht nur zum abstrakten Vergleich verschiedener Entscheidungskalküle, sondern auch zur Beurteilung realer Entscheidungsmodelle verwendet werden kann. Das trägt insbesondere bei zur Verbesserung der Wirtschaftlichkeit bei der Entwicklung von automatischen Entscheidungssystemen, wenn die Genauigkeit und damit der Preis zu verwendender Sensorik gegen die Störungssensitivität des Entscheidungsmodells selbst abgewogen wird.
3. Messung der Informationsnutzung durch einzelne Entscheidungskalküle
- Schließlich wurde in Kapitel 4 mit der Anwendung der entwickelten Konzepte auf typische Vertreter der betrachteten Entscheidungskalküle gezeigt, dass
- es keinen signifikanten Unterschied zwischen den verschiedenen Entscheidungskalkülen bezüglich der Nutzung von Eingangsinformationen gibt. Die in der Literatur vorherrschende Wahrnehmung, dass unscharfe Methoden Vorteile bei Problemstellungen mit wenig Information haben, lässt sich somit quantitativ nicht nachvollziehen.
 - es aber innerhalb der einzelnen Entscheidungskalküle durchaus gravierende Unterschiede zwischen verschiedenen Modellen gibt. Häufig hat der Entwickler des Modells dabei eine Abwägung zwischen der Effizienz der Informationsnutzung und der Berechnungskomplexität zu treffen. Insgesamt schneiden, unabhängig vom gewählten Entscheidungskalkül, sowohl bei den Bewertungsoperatoren als auch bei den Ordnungsfunktionalen integralbasierte Methoden besser ab als einzelpunktbasierte.
 - einzelne Entscheidungsmodelle (in dieser Arbeit z.B. der Gamma-Operator) auf eine Verbesserung der Qualität der Eingangsinformation mit einer scheinbaren Verschlechterung der Ergebnisqualität reagieren. Dieses Verhalten kann in speziellen Anwendungen gewünscht

sein, z .B. bei der Verwendung verfälschter und sich dadurch widersprechender Messwerte. Der Anwender sollte sich aber darüber bewusst sein, dass ein solches Verhalten grundlegend anders ist als das Verhalten der meisten anderen Entscheidungsmodelle.

Da die Frage der Informationseffizienz wie in Abschnitt 1.3 dargestellt stets auch die Frage nach der wirtschaftlichen Ressourcenverwendung in automatischen Entscheidungs- und Entscheidungsunterstützungssystemen beinhaltet, trägt die Arbeit dazu bei, die Wirtschaftlichkeit solcher Systeme zum Entwurfszeitpunkt zu beurteilen und zu verbessern.

5.3. Ausblick

Aus dem neuen Konzept der quantitativen Beurteilung von Entscheidungssystemen anhand der Effizienz der Informationsnutzung mit Hilfe von Entropiemaßen ergeben sich darüber hinaus Ansatzpunkte für weiterführende Untersuchungen, wie z.B.

- die Quantifizierung der dargestellten wirtschaftlichen Abwägung zwischen einer mit Verbesserung der Informationsnutzung verbundenen Erhöhung des Bedarfs an Berechnungsressourcen und den Kosten für eine Verbesserung der Eingangsinformationen durch verbesserte Sensorik.
- die Anwendung des hier entwickelten Ansatzes auf komplexe, real existierende Entscheidungssysteme und die Verfeinerung der verwendeten Störungsmodelle für die Eingangsinformation in praktischer Anwendung.
- aufbauend auf der hier dargestellten *Messung* der Informationseffizienz die Entwicklung eines Verfahrens zur *Verbesserung* der Informationseffizienz in bestehenden Entscheidungssystemen.

Literaturverzeichnis

- [1] ABBAS, A. E.: *An Entropy Approach for Utility Assignment in Decision Analysis*. AIP conference proceedings, 659:328–338, 2003.
- [2] ACZEL, J. und Z. DAROCZY: *On Measures of Information and their Characterization*. Academic Press, 1975.
- [3] ALBRECHT, J.: *Vorausschauende Optimale Steuer- und Regelstrategien zur Verbesserung der Kraftwerksführung*. In: *Fortschritts-Berichte VDI Reihe*, Band 616. VDI-Verlag, Düsseldorf, 1997.
- [4] ANGSTENBERGER, L.: *Dynamic fuzzy pattern recognition with applications to finance and engineering*. Kluwer, 2001.
- [5] ARMSTRONG, N., W. KALCEFF, J. P. CLINE und J. BONEVICH: *A Bayesian/Maximum Entropy Method for the Certification of a Nanocrystallite-Size; NIST Standard Reference Material*. Springer Series in Material Sciences, 68:187–228, 2004.
- [6] ARRENBURG, J.: *Schätzung der Varianz von Mittelwertschätzern in endlichen Populationen*, Band 45 der Reihe *Angewandte Statistik und Ökonometrie*. Vandenhoeck und Rupprecht, 1998.
- [7] BAMBERG, G. und A. COENENBERG: *Betriebswirtschaftliche Entscheidungslehre*. WiSo-Kurzlehrbücher. Vahlen, 11. Auflage, 2002.
- [8] BAUER, H.: *Maß- und Integrationstheorie*. de Gruyter, Berlin, New York, 2. Auflage, 1992.
- [9] BAUER, H.: *Wahrscheinlichkeitstheorie*. de Gruyter, Berlin, 5. Auflage, 2002.
- [10] BELIAKOV, G.: *Fuzzy Sets and Membership Functions Based on Probabilities*. Information Sciences, 91(1-2):95–111, 1996.
- [11] BELLMAN, R.E. und L.A. ZADEH: *Decision Making in a Fuzzy Environment*. Management Science, 17:141–163, 1970.
- [12] BERNHARD, T.: *Ein Beitrag zur Gewichteten Multikriteriellen Optimierung von Heizungs- und Lüftungsregelkreisen auf Grundlage des Fuzzy Decision Making*. Doktorarbeit, Universität Karlsruhe, 2000.

- [13] BERNOULLI, D.: *Specimen Theoriae Novae de Mensura Sortis*. Commentarii Academiae Scientiarum Imperialis Petropolitanae, 5:175–192, 1738.
- [14] BEYERER, J.: *Verfahren zur quantitativen statistischen Bewertung von Zusatzwissen in der Meßtechnik*. Nummer 783 in *Fortschritt-Berichte VDI*. VDI-Verlag, 1999.
- [15] BEZDEK, J., B. SPILLMAN und R. SPILLMAN: *Fuzzy Relations Space for Group Decision Theory*. *Fuzzy Sets and Systems*, 1:255–268, 1978.
- [16] BEZDEK, J. C., J. KELLER, R. KRISNAPURAM und N. R. PAL: *Fuzzy models and algorithms for pattern recognition and image processing*. Kluwer, 1999.
- [17] BIETHAHN, J.: *Fuzzy-Set-Theorie in Betriebswirtschaftlichen Anwendungen*. Vahlen, 1997.
- [18] BOCKLISCH, S.: *Verarbeitung multisensorieller Informationen mit Fuzzy-Techniken*. ITG Fachbericht, 126:25–33, 1994.
- [19] BOROS, E. und P. L. HAMMER (Herausgeber): *Discrete optimization*. North-Holland, 2003.
- [20] BOUCHON-MEUNIER, B., M. RIFQI und S. BOTHOREL: *Towards general measures of comparison of objects*. *Fuzzy Sets and Systems*, 84(2):143–153, 1996.
- [21] BRETTHAUER, G. und B. STRAUBE: *Einsatzmöglichkeiten Für Fuzzy-Regler*. Studie SMWK 4-7541.82-EPS/1. Fraunhofer Institut ITB-EPS, Dresden, 1992.
- [22] BUTLER, M., H. P. WILLIAMS und L.-A. YARROW: *The Two-Period Traveling Salesman Problem Applied to Milk Collection in Ireland*. *Computational Optimization and Applications*, 7(3):291–306, 1997.
- [23] CASILLAS, J., O. CORDON, F. HERRERA und L. MAGDALENA: *Trade-off between Accuracy and Interpretability In Fuzzy Rule-Based Modelling*. *Studies in Fuzziness and Soft Computing*. Physica, 2002.
- [24] CAZEMIER, D.R., P. LAGACHERIE und R. MARTIN-CLOUAIRE: *A possibility theory approach for estimating available water capacity from imprecise information contained in soil databases*. *Georama*, 103:113–132, 2001.
- [25] CHEN, M.-Y. und D.A. LINKENS: *Rules-base self-generation and simplification for data-driven fuzzy models*. *Fuzzy Sets and Systems*, 142(2):243–265, 2004.
- [26] CLAUDEL, S., C. FONTEIX, J.-P. LECLERC und H.-G. LINTZ: *Application of the possibility theory to the compartment modelling of flow pattern in industrial processes*. *Chemical Engineering Science*, 58:4005–4016, 2003.

- [27] COCHRANE, E. M. und J. E. BEASLEY: *The co-adaptive neural network approach to the Euclidean Travelling Salesman Problem*. Neural Networks, Oxford, 16(10):1499–1525, 2003.
- [28] COENENBERG, A. G. und R. KLEINE-DOEPKE: *Zur Abbildung der Risikopräferenz durch Nutzenfunktionen. Stellungnahme zur Kritik Jacobs und Lebers am Bernoulli-Prinzip*. ZfB, 45:663–665, 1975.
- [29] COLLINS, L. M., YAN ZHANG, JING LI und ET AL.: *A Comparison of the Performance of Statistical and Fuzzy Algorithms for Unexploded Ordnance Detection*. IEEE Transactions on Fuzzy Systems, 9(1):17–30, Feb. 2001.
- [30] CORK, D. J. und A. TOGUEM: *Using Fuzzy Logic to Confirm the Integrity of a Pattern Recognition Algorithm for Long Genomic Sequences: The W-Curve*. Annals of the New York Academy of Sciences, 980:32–40, 2002.
- [31] CZYZAK, P. und R. SLOWINSKI: *Possibilistic Construction of Fuzzy Outranking Relation for Multiple Criteria Ranking*. Fuzzy sets and systems, 81:123–131, 1996.
- [32] D'AGOSTINI, G.: *Role and meaning of subjective probability*. In: MOHAMMAD-DJAFARI (Herausgeber): *Bayesian inference and maximum entropy methods in science and engineering*. American institute of physics, 2001.
- [33] DEBREU, G.: *Representation of a Preference Ordering by a Numerical Function*. In: AL, ROBERT M. THRALL ET (Herausgeber): *Decision Processes*, Seiten 159–165. Wiley, New York, 1954.
- [34] DELGADO, M., J. L. VERDEGAY und M. A VILA: *A Model for Linguistic Partial Information in Decision-Making Problems*. International Journal of Intelligent Systems, 9(4):365–378, 1994.
- [35] DELMOTTE, F.: *Comparison of the Performances of Decision Aimed Algorithms with Bayesian and Beliefs Basis*. International Journal of intelligent systems, 16:963–981, 2001.
- [36] DELMOTTE, F. und P. BORNE: *Modeling of reliability with possibility theory*. IEEE Transactions on System, Man and Cybernetics - Part A: Systems and Humans, 28(1):78–88, Jan 1998.
- [37] DEMPSTER, A.P.: *A Generalization of Bayesian Inference*. Journal of the Royal Statistical Society, 30:205–247, 1968.
- [38] DENG, HEPU und CHUNG-HSING YEH: *Fuzzy utilities comparison in multi-criteria analysis*. Lecture notes in computer science, 1625:401–410, 1999.
- [39] DRAKOPOULOS, J. A.: *Probabilities, possibilities and fuzzy sets*. Fuzzy Sets and Systems, 75(1):1–15, 1995.

- [40] DUBOIS, D., H. FARGIER und H. PRADE: *Refinements of the maximin approach to decision-making in a fuzzy environment*. Fuzzy Sets and Systems, 81(1):103–122, 1996.
- [41] DUBOIS, D. und H. PRADE: *Théorie Des Possibilités*. Masson, Paris, 1985.
- [42] DUBOIS, D. und H. PRADE: *The three semantics of fuzzy sets*. Fuzzy Sets and Systems, 90(2):141–150, 1997.
- [43] DUBOIS, D. und H. PRADE: *Possibility theory, probability theory and multiple-valued logics: A clarification*. Annals of mathematics and artificial intelligence, 32:35–66, 2001.
- [44] DUBOIS, D. und H. PRADE: *Fuzzy set and possibility theory-based methods in artificial intelligence*. Artificial intelligence, 148:1–9, 2003.
- [45] DUMITRIU, B., R. MIKUT, G. BRETTHAUER, G. WERFEL, S. BOETTGER und M. SIEWIOR: *Inbetriebnahme einer fuzzy-adaptiven Giessspiegelregelung in einem Stahlwerk der Baogang Steel*. Stahl und Eisen, Seiten 35–38, 1999.
- [46] DUPONT, P.: *Laplace and the Indifference Principle in the 'Essai philosophique des probabilités'*. Rend. Sem. Mat. Univ. Politec. Torino, 36:125–137, 1977.
- [47] DUVENHAGE, R.: *The Nature of Information in Quantum Mechanics*. Foundations of Physics, 32(9):1399–1417, 2002.
- [48] FENG, CHU: *Fuzzy multicriteria decision-making in distribution of factories: an application of approximate reasoning*. Fuzzy Sets and Systems, 71(2):197–205, 1995.
- [49] FILEV, D.P. und R. R. YAGER: *A generalized defuzzification method via bad distributions*. International Journal of Intelligent Systems, 6:687–697, 1991.
- [50] FISCHER, A. und J. FLIEGE: *Diskrete Optimierung*. Ergebnisberichte angewandte Mathematik. Univ. Dortmund, Fachbereich Mathematik, 2001.
- [51] FISHER, R.A.: *A Theory of Statistical Estimation*. Proceedings Cambridge Philosophical Society, 22:700–725, 1925.
- [52] GIL, M. A. und M. T. LOPEZ: *Statistical Management of Fuzzy Elements in Random Experiments. Part 2: The Fisher Information Associated with a Fuzzy Information System*. Information Sciences, 69(3):243, 1993.
- [53] GIL, MARIA A. und PRAMOD JAIN: *Comparison of experiments in statistical decision problems with fuzzy utilities*. IEEE Transactions on Systmes, Man and Cybernetics, 22(4):662–670, July/August 1992.

- [54] GRUNWALD, P. D. und P. M. VITANYI: *Kolmogorov Complexity and Information Theory. With an Interpretation in Terms of Questions and Answers.* Journal of Logic, Language and Information, 12(4):497–529, 2003.
- [55] GUAN, J. W. und D. A. BELL: *Approximate Reasoning and Evidence Theory.* Information Sciences, 96(3-4):207–235, 1997.
- [56] GUO, SHU-XIANG und ZHEN-ZHOU LU: *Procedure for computing the possibility and fuzzy probability of failure of structures.* Applied Mathematics and Mechanics, 24(3):338–343, Mar 2003. English Edition.
- [57] HALD, A.: *A history of probability and statistics and their applications before 1750.* Wiley series in probability and mathematical statistics. Wiley, 1990.
- [58] HAMMER, P.L.: *Discrete Optimization.* Elsevier Science, 1979.
- [59] HAUSDORFF, F.: *Grundzüge der Mengenlehre.* Veit und Comp., Leipzig, 1914.
- [60] HAZOD, W.: *Stetige Faltungshalbgruppen von Wahrscheinlichkeitsmaßen und erzeugende Distributionen.* Springer, 1997.
- [61] HONERKAMP, J.: *Statistical Physics.* Advanced Texts in Physics. Springer, 1. Auflage, 2000.
- [62] HSU, HSI-MEI und CHEN-TUNG CHEN: *Fuzzy Credibility Relation Method for Multiple Criteria Decision-Making Problems.* Information Sciences, 96(1-2):79–91, 1997.
- [63] HUANG, H. P. und Y.-H. LIU: *Fuzzy Support Vector Machines for Pattern Recognition and Data Mining.* International Journal of Fuzzy Systems, 2002.
- [64] HURWICZ, L.: *Optimality Criteria for Decision Making under Ignorance,* Band 370 der Reihe *Cowles Commission Discussion Papers.* Cowles foundation, 1951.
- [65] HUYGENS, CH.: *De Ratiociniis in Ludo Aleae.* S. Keimer, 1657. in [66].
- [66] HUYGENS, CH.: *Oeuvres complètes,* Band 14: *Calcul des probabilités.* Travaux d mathématiques pures : 1655-1666. publ. par la Société Hollandaise des Sciences, 1920. *Originalexemplar vorhanden in der Bibliothek des Mathematischen Institut der Universität Freiburg.*
- [67] IKONEN, E.: *Algorithms for process modelling using fuzzy neural networks.* Acta Universtatis Ouluensis. Oulun Yliopisto, 1996.
- [68] IP, C. K. W., C. K. KWONG, H. BAI und Y. C. TSIM: *The process modelling of epoxy dispensing for microchip encapsulation using fuzzy linear regression with fuzzy intervals.* International Journal of Advanced Manufacturing Technology, 22(5/6):417–423, 2003.

- [69] JANG, JYH-SHING R., CHUEN-TSAI SUN und E. MIZUTANI: *Neuro-fuzzy and soft computing*. Matlab curriculum series. Prentice Hall, Upper Saddle River, NJ, 1997.
- [70] JAYNES, E.T.: *Information Theory and Statistical Mechanics*. Physical Review, 106:620–639, 1957.
- [71] JAYNES, E.T.: *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.
- [72] JÄKEL, J.: *Linguistische Fuzzy-Systeme mit verallgemeinerten Konklusionen und ihre Anwendung zur Modellbildung und Regelung*. Nummer 793 in *Fortschritt-Berichte*. VDI-Verlag, 1999.
- [73] JUMARIE, G.: *Further Results on the Mathematical Relations between Probability, Possibility and Fuzzy Logic. The Rationale of Subjectivity via Invariance of Information Loss*. Systems Analysis Modelling Simulation, 28(1/4):107–148, 1996.
- [74] KIENDL, H.: *Erweiterter Anwendungsbereich von Fuzzy Control durch Hyperinferenz und Hyperdefuzzifizierung*. VDI BERICHTE, 1113:319–328, 1994.
- [75] KIENDL, H.: *Fuzzy Control Methodenorientiert*. Oldenbourg Verlag, München, 1997.
- [76] KIENDL, H.: *Non-translation-invariant Defuzzification*. Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, Seiten 737–742, 1997.
- [77] KIENDL, H.: *Decision Analysis by Advanced Fuzzy Systems*. In: ZADEH, L. (Herausgeber): *Computing with words in Information/Intelligent Systems*, Seiten 223–242. Physica-Verlag, 1999.
- [78] KIENDL, H. und M. KRABS: *Ein Verfahren zur Generierung regelbasierter Modelle für dynamische Systeme*. at - Automatisierungstechnik, 37(11):423–430, 1989.
- [79] KOCH, M., T. KUHN und J. WERNSTEDT: *Fuzzy Control. Optimale Nachbildung und Entwurf optimaler Entscheidungen*. Oldenbourg, München, 1996.
- [80] KOEHLER, D. J., C. M. WHITE und R. GRONDIN: *An evidential support accumulation model of subjective probability*. Cognitive Psychology, 46:152–197, 2003.
- [81] KOHLAS, J. und P.-A. MONNEY: *A Mathematical Theory of Hints*, Band 425 der Reihe *Lecture Notes in Economics and Mathematical Systems*. Springer, 1995.

- [82] KOLMOGOROV, A.N.: *Three approaches to the quantitative definition of information*. Problems of information transmission, 1:4–7, 1965.
- [83] KOUTSOUKIS, N.-S.: *Decision modelling and information systems*. Kluwer, 2003.
- [84] KRABS, M.: *Das ROSA-Verfahren zur Modellierung dynamischer Systeme durch Regeln mit statistischer Relevanzbewertung*. Fortschrittsberichte des VDI, 8(404), 1994.
- [85] KRABS, M. und H. KIENDL: *Anwendungsfelder der automatischen Regelgenerierung mit dem ROSA-Verfahren*. at - Automatisierungstechnik, 43(6):269–276, 1995.
- [86] KREBS, V. und E. SCHAEFERS: *Dynamische Fuzzy-Systeme zur qualitativen Prozessmodellierung*. VDI Berichte, 1381:115–136, 1998.
- [87] KRELLE, W.: *Preistheorie*. Mohr [u.a.], Tübingen-Zürich, 1961.
- [88] LAPLACE, P. S. DE: *Essai Philosophique sur les Probabilités*. Bachelier, Paris, 1814. In Oeuvres complètes, Bd 1., 1878; Originalexemplar vorhanden in der Bibliothek des Mathematischen Institut der Universität Freiburg.
- [89] LAPLACE, P. S. DE: *Théorie Analytique Des Probabilités*. Courcier, Paris, 1814. 2. éd., rev. et augmentée; Originalexemplar vorhanden in der Bibliothek des Mathematischen Institut der Universität Freiburg.
- [90] LAUX, H.: *Entscheidungstheorie*. Springer, 5 Auflage, 2003.
- [91] LEBER, W.: *Zur Rationalität von Entscheidungskriterien bei Unsicherheit*. ZfB, 45:493–496, 1975.
- [92] LEE, E. W., R. K. YUEN, S. M. LO und K. C. LAM: *Probabilistic inference with maximum entropy for prediction of flashover in single compartment fire*. Advanced Engineering Informatics, 16(3):179–191, 2002.
- [93] LEEKWIJCK, W. VAN und E. E. KERRE: *Defuzzification: Criteria and Classification*. Fuzzy sets and systems, 108:159–178, 1999.
- [94] LIU, X.: *On the Methods of Decision Making under Uncertainty with Probability Information*. International Journal of Intelligent Systems, 19(12):1217–1238, 2004.
- [95] LOMBARDI, O.: *What is Information?* Foundations of science, 9(2):105–134, 2004.
- [96] LUCA, A. DE und S. TERMINI: *A definition of non-probabilistic entropy in the setting of fuzzy sets theory*. Information and Control, 20:301–312, 1972.

- [97] LUO, S.: *Maximum Shannon Entropy, Minimum Fisher Information, and an Elementary Game*. *Foundations of Physics*, 32(11):1757–1772, 2002.
- [98] MAO-JIUN, J. WANG und CHANG TIEN-CHIEN: *Tool steel materials selection under fuzzy environment*. *Fuzzy Sets and Systems*, 72(3):263–270, 1995.
- [99] MARICHAL, J.-L. und M. ROUBENS: *Entropy of Discrete Fuzzy Measures*. *International Journal of Uncertainty Fuzziness and Knowledge Based Systems*, 8(6):625–640, 2000.
- [100] MIKUT, R., J. JÄKEL und L. GRÖLL: *Interpretability issues in data-based learning of fuzzy systems*. Preprint submitted to Elsevier Science, Jan 2005.
- [101] MIKUT, R., J. JÄKEL, H. MALBERG und G. BRETTHAUER: *Datenbasierter Entwurf von Fuzzy-Systemen für medizinische Diagnoseaufgaben*. at - Automatisierungstechnik, 48:1–9, 2000.
- [102] MÖLLER, B., M. BEER, W. GRAF und A. HOFFMANN: *Possibility theory based safety assessment*. *Computer-Aided Civil and Infrastructure Engineering*, 14:81–91, 1999.
- [103] MOIVRE, A. DE: *The Doctrine of Chances: a Method of Calculating the Probabilities of Events in Play*. eigen, London, 3 Auflage, 1756.
- [104] MOON, BYUNG SOO: *Equivalence between fuzzy logic controllers and PI controllers for single input systems*. *Fuzzy Sets and Systems*, 69(2):105–113, 1995.
- [105] NAUCK, D. und R. KRUSE: *Fuzzy Systeme und Soft Computing*. In: BIETHAHN, J. (Herausgeber): *Fuzzy Set-Tehory in Betriebswirtschaftlichen Anwendungen*, Seiten 3–21. Verlag Vahlen, 1997.
- [106] NAUCK, D. und R. KRUSE: *Obtaining interpretable fuzzy classification rules from medical data*. *Artificial Intelligence in Medicine*, 16(2):149 – 169, 1999.
- [107] NEUMANN, J. und O. MORGENSTERN: *Theory of games and economic behavior*. Princeton, 1944.
- [108] NEUMANN, U.: *Integrierte Instandhaltungsplanung für elektrische Energieübertragungssysteme*. Doktorarbeit, Universität Dortmund, 1997.
- [109] NGUYEN, H. T.: *On Modeling of Linguistic Information Using Random sets*. *Information Sciences*, 34:265–274, 1984.
- [110] NGUYEN, H. T.: *Fuzzy sets and probability*. *Fuzzy Sets and Systems*, 90(2):129–132, 1997.
- [111] NIEHANS, H.: *Zur Preisbildung bei Ungewissen Erwartungen*. *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 84:433–456, 1948.

- [112] OTT, N.: *Unsicherheit, Unschärfe und rationales Entscheiden*. Wirtschaftswissenschaftliche Beiträge. Physica-Verlag, 2001.
- [113] PAMMER, H., R. FELIX und J. KÜHNEN: *Reihenfolgeplanung in der Automobilmontage mit FuzzyDecisionDesk*. In: *Anwendersymposium der Fuzzy-Neuro-Initiative NRW*, Seiten 39–43, Düsseldorf, 1998.
- [114] PANTELLA, M. und A. RIZZI: *From Circuits to Neurofuzzy Networks*. *Journal of Circuits, Systems and Computers*, 13(1):205–236, 2004.
- [115] PAPOULIS, A.: *Probability, Random Variables and Stochastic Processes*, Kapitel 3-5 und 4-4, Seiten 38–39, 78–81, 112–114. McGraw-Hill, 2. Auflage, 1984.
- [116] PASCAL, B. und P. DE FERMAT: *La Correspondance de Blaise Pascal et de Pierre de Fermat*, Band 32 der Reihe *Les Cahiers de Fontenay*. ENS de Fontenay-aux-Roses, 1983.
- [117] PYTEV, YU P.: *The methods of possibility theory in the problem of optimal estimation and decision making: I. Measures of possibility and necessity and integration with respect to possibility and necessity*. *Pattern recognition and image analysis*, 7(3):338–346, 1997.
- [118] PYTEV, YU P.: *The methods of possibility theory in the problem of optimal estimation and decision making: III. Fuzzy elements, independence, conditional distributions and optimal estimation*. *Pattern recognition and image analysis*, 9(3):416–426, 1999.
- [119] PYTEV, YU P.: *The methods of possibility theory in the problem of optimal estimation and decision making: IV. The methods of measurement reduction*. *Pattern recognition and image analysis*, 10(1):43–52, 2000.
- [120] PYTEV, YU P.: *The methods of possibility theory in the problem of optimal estimation and decision making: VI. Fuzzy sets. Independence. P-Completion. Methods for estimating fuzzy sets and their parameters*. *Pattern recognition and image analysis*, 12(2):107–115, 2002.
- [121] PYTEV, YU P.: *Stochastic models of possibility*. *Pattern recognition and image analysis*, 12(4):376–396, 2002.
- [122] PYTEV, YU P. und O. V. ZHUCHKO: *The methods of possibility theory in the problem of optimal estimation and decision making: VII. Reconstruction of functional dependences from experimental data*. *Pattern recognition and image analysis*, 12(2):116–129, 2002.
- [123] QIDONG, Y.: *Gas Field Development Program Optimizing by Urstoff Evaluation of Fuzzy Data Entropy*. *Natural Gas Industry*, 23(2):85–86, 2003.

- [124] RAMSEY, F.: *Truth and probability*. In: PAUL, K. (Herausgeber): *The Foundations of Mathematics and other Logical Essays*. Trench, Trubner and Co., London, 1931.
- [125] RIBO, M. und A. PINZ: *A Comparison of Three Uncertainty Calculi for Building Sonar-Based Occupancy Grids*. *Robotics and Autonomous Systems*, 35:201–209, 2001.
- [126] ROMMELFANGER, H.: *Entscheiden bei Unschärfe*. Springer, 1988.
- [127] ROMMELFANGER, H.: *Fuzzy-Logik Basierte Verarbeitung von Expertenregeln*. *OR-Spektrum*, 15:31–42, 1993.
- [128] ROMMELFANGER, H.: *Fuzzy decision support systeme*. Springer, Heidelberg, 1994.
- [129] ROMMELFANGER, H. und S. EICKEMEIER: *Entscheidungstheorie - Klassische Konzepte und Fuzzy Erweiterungen*. Springer, 2002.
- [130] RONG-JUN, LI: *Basis of fuzzy decision - Comparison and ranking of fuzzy sets*. *Control and Decision*, 18(2):221–224, March 2003.
- [131] ROUBENS, M.: *Fuzzy sets and decision analysis*. *Fuzzy Sets and Systems*, 90(2):199–206, 1997.
- [132] ROWLEY, J.: *What is information?* *Information Services and Use*, 18(4):243–254, 1998.
- [133] RUNKLER, T. A.: *Selection of appropriate defuzzification methods using application specific properties*. *IEEE Transactions on Fuzzy Systems*, 5(1):72–79, 1997.
- [134] RUNKLER, T. A. und M. GLESNER: *Defuzzification with improved static and dynamic behaviour: Extended center of area*. In: *European Congress on Fuzzy Intelligent Technology*, Seiten 845–851, Aachen, September 1993.
- [135] SAADE, J. J.: *A Unifying Approach to Defuzzification and Comparison of the Outputs of Fuzzy Controllers*. *IEEE Transactions on Fuzzy Systems*, 4(3):227–237, 1996.
- [136] SALIGER, E.: *Betriebswirtschaftliche Entscheidungstheorie*. Oldenbourg Verlag, München, 2003.
- [137] SAVAGE, L. J.: *The foundations of statistics*. Dover, 1954.
- [138] SAVAGE, L.J.: *The Theory of Statistical Decision*. *Journal of the American statistical association*, 46:55–67, 1951.
- [139] SCHAUDEL, F. und G. BRETTHAUER: *Störungssensitivität von Ordnungsfunktionalen*. Tagungsband des 49. IWK, Ilmenau, 2004.

- [140] SCHILDBACH, TH.: *Diskussion über das Bernoulli Prinzip in Deutschland und im Ausland*. ZfB, 59(7):766–778, 1989.
- [141] SHAFER, G.: *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [142] SHANNON, C. L. und W. WEAVER: *The Mathematical Theory of Communication*. Univ. of Illinois Press, 1964.
- [143] SINGH, V. P.: *The entropy theory as a tool for modelling and decision-making in environmental and water resources*. Water SA, 26(1):1–12, 2000.
- [144] SLAWINSKI, T.: *Analyse und effiziente Generierung von relevanten Fuzzy-Regeln in hochdimensionalen Suchräumen*. Nummer 686 in *Fortschritts-Berichte VDI*. VDI-Verlag, 2001.
- [145] SLAWINSKI, T., A. KRONE und H. KIENDL: *Automatisierung durch datenbasierte Fuzzy-Modellierung von Prozessbedienern*. VDI BERICHTE, 1381:203–220, 1998.
- [146] SMETS, P.: *Belief Functions*. In: SMETS, P. (Herausgeber): *Non Standard Logics for Automated Reasoning*, Seiten 253–286. Academic Press, London, 1988.
- [147] SMETS, P.: *Decision Making in Context where Uncertainty is Represented by Belief Functions*. In: SRIVASTAVA, RAJENDRA P. und THEODORE J. MOCK (Herausgeber): *Belief Functions in Business Decisions*, Seiten 17–61. Physica, 2002.
- [148] SON, C.: *Optimal control planning strategies with fuzzy entropy and sensor fusion for robotic part assembly tasks*. International Journal of Machine Tools and Manufacture, 42(12):1335–1344, 2002.
- [149] SORGER, G.: *Entscheidungstheorie bei Unsicherheit*. UTB für Wissenschaft. Lucius und Lucius, 2000.
- [150] SURAJ, Z., J.F. PETERS und W. RZASA: *A Comparison of Different Decision Algorithms used in Volumetric Storm Cell Classification*. Fundamenta Informaticae, 51:201.214, 2002.
- [151] SUZUKI, T., T. KODAMA, T. FURUHASHI und H. TSUTSUI: *Fuzzy modeling using genetic algorithms with fuzzy entropy as conciseness measure*. Information Sciences, 136(1-4):53–67, 2001.
- [152] TANAKA, H., T. OKUDA und K. ASAI: *A Formulation of fuzzy decision problems and its application to an investment problem*. Kybernetes, 5:25–30, 1976.

- [153] TEODOROVI, D. und G. PAVKOVIC: *The fuzzy set theory approach to the vehicle routing problem when demand at nodes is uncertain*. Fuzzy Sets and Systems, 82(3):307–317, 1996.
- [154] TRAN, L. und L. DUCKSTEIN: *Comparison of Fuzzy Numbers Using a Fuzzy Distance Measure*. Fuzzy sets and systems, 130:331–341, 2002.
- [155] WANG, L.-L., J. WANG und L.-T. JI: *A Study of Travelling Salesman Problem*. Journal - China Universities of Posts and Telecommunications, 8(1):15–19, 2001.
- [156] WANG, W.-J. und C.-H. CHIU: *Entropy and information energy for fuzzy sets*. Fuzzy Sets and Systems, 108(3):333–339, 1999.
- [157] WANG, X. und E. E. KERRE: *On the classification and dependencies of the ordering methods*. In: RUAN, D. (Herausgeber): *Fuzzy logic foundations and industrial applications*, Seiten 73–88. Kluwer Acad., 1996.
- [158] WANG, X. und E. E. KERRE: *Reasonable properties for the ordering of fuzzy quantities(III)*. Fuzzy sets and systems, 118:387–405, 2001.
- [159] WANG, X. und E. E. KERRE: *Reasonable properties for the ordering of fuzzy quantities (I)*. Fuzzy sets and systems, 118:375–385, 2001.
- [160] WATSON, S. R., J.J. WEISS und M.L. DONNELL: *Fuzzy Decision Analysis*. IEEE Transactions on systems, man and cybernetics, 9:1–9, 1979.
- [161] WEIRONG, X. und B. WEIDEMANN: *Fuzzy modelling and its application to magnetic bearing systems*. Fuzzy Sets and Systems, 73(2):201–217, 1995.
- [162] WHALEN, T: *Decision making under uncertainty with various assumptions about available information*. IEEE Transactions on Systems, Man and Cybernetics, 14:888–900, 1984.
- [163] WILHELM, K.: *Chevalier de Méré und sein Verhältnis zu Blaise Pascal*. Ebering, Berlin, 1936.
- [164] XU, Y.-Y. und X.-Z. ZHOU: *Chinese Sentence Parsing Based on Maximum Entropy Model*. Acta Electronica Sinica, 31(11):1608–1612, 2003.
- [165] XUZHU, WANG, B. DE BAETS und E. KERRE: *A comparative study of similarity measures*. Fuzzy Sets and Systems, 73(2):259–268, 1995.
- [166] YAGER, R. R.: *On choosing between fuzzy subsets*. Cybernetics, 9:151–154, 1980.
- [167] YAGER, R. R.: *Measures of Entropy and Fuzziness Related to Aggregation Operators*. Information Sciences, 82(3-4):147–166, 1995.

- [168] YAGER, R. R.: *On the Entropy of Fuzzy Measures*. IEEE Transactions on Fuzzy Systems, 8(4):453–461, 2000.
- [169] YAGER, R. R.: *Uncertainty Representation Using Fuzzy Measures*. IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics, 32(1):13–20, Feb. 2002.
- [170] YAGER, R. R., M. DETYNIĘCKI und B. BOUCHON-MEUNIER: *A context dependent method for ordering fuzzy numbers using probabilities*. Information sciences, 138:237–255, 2001.
- [171] YAGER, R. R. und D. FILEV: *On Ranking Fuzzy Numbers Using Valuations*. International journal of intelligent systems, 14:1249–1268, 1999.
- [172] YILMAZ, M. R.: *Subjective probability and information: a constructive perspective*. American Journal of Mathematical and Management Sciences, 23(1):7–36, 2003.
- [173] YIQIAN, H., L. JIANHONG und Z. TIEJUN: *A Kind of Adaptive Fuzzy Control Method and Its Application to Steam Temperature Control of a Boiler*. Proceedings of the chinese society of electricla engineering, 23-1:136–140, 2003.
- [174] YONG-GE, WU und LIU LEI-JIAN: *On The Evidence Inference Theory*. Information Sciences, 89(3-4):245–260, 1996.
- [175] ZADEH, L. A.: *Fuzzy Sets*. Information and Control, 8:338–353, 1965.
- [176] ZADEH, L. A.: *Fuzzy sets as a basis of a theory of possibility*. Fuzzy sets and systems, 1:3–28, 1978.
- [177] ZHANG, J. und J. MORRIS: *Process modelling and fault diagnosis using fuzzy neural networks*. Fuzzy Sets and Systems, 79(1):127–140, 1996.
- [178] ZHANG, ZHEN und YUNJUAN CAO: *First-order logic - The unity of fuzziness and randomness*. Fuzzy Sets and Systems, 93(2):185–195, 1998.
- [179] ZIMMERMANN, H. J.: *Fuzzy Set Theory and its Applications*. Kluwer Academic, 2 Auflage, 1990.

6. Anhang

A. Bezeichnungen und Abkürzungen

Verwendete Bezeichnungen

Zur Erhöhung der Verständlichkeit werden einige Bezeichnungen in einem strikteren Sinn verwendet als dies manchmal in der Literatur der Fall ist:

Operatoren und Funktionale Abbildungen von einem Funktionenraum in einen Funktionenraum werden als Operatoren bezeichnet, Abbildungen von einem Funktionenraum nach \mathbb{R} als Funktionale.

Entscheidungskalkül Ein Entscheidungskalkül ist eine grundlegende Systematik oder Sprache zur Modellierung von Entscheidungsmodellen, z.B. statistische Entscheidungstheorie, Fuzzy-Entscheidungstheorie, possibilistische Entscheidungstheorie, etc.

Entscheidungsproblem Ein Entscheidungsproblem ist eine Menge von Handlungsalternativen mit der Fragestellung, welche Alternative bei gegebenen Umständen die beste Handlung darstellt.

Entscheidungsmodell Ein Entscheidungsmodell ist ein mathematisches Modell zur Lösung eines Entscheidungsproblems, bestehend aus einer Menge von Handlungsalternativen, Informationen über Umweltzustände, einem Bewertungsoperator und einem Ordnungsfunktional.

Mathematische Abkürzungen

Abkürzung	Bedeutung
\mathbb{R}	Körper der reellen Zahlen
$\mathcal{P}(\mathbb{R})$	Potenzmenge von \mathbb{R}
M^c, M^{-1}	(\mathbb{R} -)Komplement der Menge M
$\mathcal{Q}, \mathcal{Q}^n$	Hyperquader im \mathbb{R}^n
$\mathcal{B}, \mathcal{B}(\mathbb{R})$	Raum der beschränkten, positiv definiten Funktionen über \mathbb{R}
$f _{\alpha}(x)$	die eindimensionale Funktion, welche durch Auswerten der zweidimensionalen Funktion $f(x, y)$ entlang der Geraden $y = \alpha$ entsteht
$a \circ b$	Nacheinanderausführung von Funktionalen: $[a \circ b](x) := a(b(x))$
$\int_{\mathbb{R}}, \int_{\mathbb{R}} dx$	Integral über die ganze reelle Achse
$(\vec{a})_i$	i -te Komponente des Vektors \vec{a}
$\mu, \mu(r)$	Wahrscheinlichkeitsverteilung oder Zugehörigkeitsfunktion
$EW(\mu_X)$	Erwartungswert der Zufallsvariablen mit der Dichte μ_X ; $EW(X) = \int_{\mathbb{R}} x d\mu_X$
$N_{\mu, \sigma}$	Normalverteilung mit Mittelwert μ und Varianz σ
$\mu \circledast \nu$	Summendichte (Faltprodukt): $\mu \circledast \nu := \int_{\mathbb{R}} \mu(x)\nu(z-x)dx$
$\mu \boxtimes \nu$	Produkttdichte: $\mu \boxtimes \nu := \int_{\mathbb{R}} 1/x \mu(z/x)\nu(x)dx$

Durchgängig verwendete Bezeichnungen

Abkürzung	Bedeutung
α, α_i	Handlungsalternativen
\mathbb{A}	Handlungsraum: Menge der möglichen Handlungsalternativen α
$\vec{\kappa}_{i,j}$	Ergebnisvektor für die Handlungsalternative i unter Annahme des Umweltzustandes j
$\nu, \nu(\alpha)$	Bewertungsfunktion für Handlungsalternativen
$\hat{\nu}, \hat{\nu}(\kappa)$	Bewertungsfunktion für Ergebnisvektoren
σ	Umweltzustände
\mathbb{S}	Zustandsraum: Menge sich gegenseitig ausschließender Umweltzustände σ
\vec{s}	Umweltvektor: Ein Umweltzustand σ_n aus einem parametrisierbaren Umweltraum ist eineindeutig beschrieben durch einen Wert (s_1, s_2, \dots, s_n) des Umweltvektors s_n .
\vec{i}, \vec{i}_s	Informationsfunktionen: Schar von Funktionen aus \mathcal{B} , die Informationen über Werteverteilungen in \vec{s} enthalten, also z.B. Wahrscheinlichkeitsverteilungen, Zugehörigkeitsfunktionen, ...
\mathcal{E}	Entscheidungsmodell $\mathcal{E} : (\mathbb{A}; (\mathcal{J}_{[a;b]}^i)^n) \rightarrow \mathbb{R}$
\mathcal{B}	Bewertungsoperator $\mathcal{B} : (\mathbb{R}; \mathcal{J}^n) \rightarrow \mathbb{R} \times \mathcal{J}$
\mathcal{O}	Ordnungsfunktional $\mathcal{O} : \mathcal{J} \rightarrow \mathbb{R}$

B. Funktionsdefinitionen

i_1 Die Dichte einer Normalverteilung¹ $N_{(5,0.8165)}$:

$$i_1(x) := \begin{cases} i_1(x) = \frac{1}{1.962} e^{-0.82*(x-5)^2} & \text{für } x \in [0; 10] \\ 0 & \text{sonst} \end{cases} \quad (6.1)$$

¹Die Varianz ist so gewählt, dass $\text{Var}(i_1) = \text{Var}(i_2)$.

i_2 Eine normale Dreiecks-ZGF:

$$i_2(x) := \begin{cases} 0 & \text{für } x \leq 3 \\ 0.5 * (x - 3) & \text{für } 3 < x \leq 5 \\ 0.5 * (7 - x) & \text{für } 5 < x \leq 7 \\ 0 & \text{für } x > 7 \end{cases} \quad (6.2)$$

i_3 Eine ZGF mit zwei dreieckigen Peaks:

$$i_3(x) := \begin{cases} 0 & \text{für } x \leq 2 \\ x - 2 & \text{für } 2 < x \leq 3 \\ 4 - x & \text{für } 3 < x \leq 4 \\ 0 & \text{für } 4 < x \leq 6 \\ x - 6 & \text{für } 6 < x \leq 7 \\ 8 - x & \text{für } 7 < x \leq 8 \\ 0 & \text{für } x > 8 \end{cases} \quad (6.3)$$

i_4 Eine fast trapezförmige ZGF:

$$i_4(x) := \begin{cases} 0 & \text{für } x \leq 2 \\ 1.8 * (x - 2) & \text{für } 2 < x \leq 2.5 \\ \frac{x-2.5}{40} + 0.9 & \text{für } 2.5 < x \leq 6.5 \\ 2 * (7 - x) & \text{für } 6.5 < x \leq 7 \\ 0 & \text{für } x > 7 \end{cases} \quad (6.4)$$

C. Programme

Bewertungsoperatoren

Laden der notwendigen Bibliotheken

```
Needs["Statistics`ContinuousDistributions`"]
Needs["Statistics`DataManipulation`"]
Needs["Calculus`Integration`"]
Needs["NumericalMath`ListIntegrate`"]
```

Definitionen

```
Normal4[x_, t_] := NormalDistribution[4, t];
NorVert4[x_, t_] = PDF[Normal4[x, t], x];
Normal6[x_, t_] := NormalDistribution[6, t];
NorVert6[x_, t_] = PDF[Normal6[x, t], x];
RandNormal002[x_] := Random[NormalDistribution[0, 0.2]];
Gleichvert[x_] := Random[UniformDistribution[0, 0.2]];

TableGen[fun_] := Table[fun, {x, 0, 10, Prec}];

lo01 = Log[0.1];
HSE[tab_] := Prec*Sum[tab[[n]]*
  If[tab[[n]] > 0, Log[tab[[n]]]/lo01, 0]
  , {n, 1, 10/Prec}];
MinkEnt[tab_] := 1 - (Prec*Sum[tab[[n]]*tab[[n]], {n, 1, 10/Prec});

ZaPlu[a_, b_] :=
  Module[{i, tem, t1, hilf1, hilf2},
    tem = TableGen[NorVert4[x, 1]];
    For[n = 1, n <= 10/Prec,
      t1 = 0;
      hilf1 = Min[10/Prec, 2*n - 1];
      For[i = Max[1, (2*n - 10/Prec)], i < hilf1,
        hilf2 = Min[a[[i]], b[[2*n - i]]];
        If[hilf2 > t1, t1 = hilf2];
        i++];
      tem[[n]] = t1;
      n++];
```

```

Return[tem]];

Falt2[a_, b_] :=
Module[{tem, ah, bh, la, lb, lf, f0, i},
  ah = PadRight[a, Round[20/Prec]];
  bh = PadRight[b, Round[20/Prec]];
  la = Fourier[ah];
  lb = Fourier[bh];
  lf = la*lb;
  f0 = InverseFourier[lf];
  f0 = f0/ListIntegrate[f0, Prec];
  tem = TableGen[NorVert4[x, 0.1]];
  For[i = 1, i <= 10/Prec, tem[[i]] = 2*f0[[2*i]]; i++];
  Return[tem];
];

Gamma05[a_, b_] := Module[{n, t1, t2, tem},
  tem = TableGen[NorVert4[x, 1]];
  For[n = 1, n < 10/Prec,
    t1 = Abs[a[[n]]];
    t2 = Abs[b[[n]]];
    tem[[n]] = Sqrt[t1*t2]*Sqrt[1 - (1 - t1)*(1 - t2)];
    n++];
  Return[tem]];

MiWe[a_, b_] := 0.5*a + 0.5*b;

```

Stochastische Störung

```

Prec = 0.0005;
DP = 60;
a = TableGen[NorVert4[x, 0.5]];
b = TableGen[NorVert6[x, 0.5]];
se1 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
se2 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
se3 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
se4 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
ABase = HSE[a];
Timing[For[i = 0, i <= DP,

```

```

Störung = Table[Random[NormalDistribution[0, 1]], {x, 0, 10, Prec}];
ar = Abs[(1 - (i/300))*a + Störung*i/300];
ar = ar/ListIntegrate[ar, Prec];
F1 = Falt2[ar, b];
F1 = F1/ListIntegrate[F1, Prec];
F2 = Gamma05[ar, b];
F2 = F2/ListIntegrate[F2, Prec];
F3 = MiWe[ar, b];
F3 = F3/ListIntegrate[F3, Prec];
F4 = ZaPlu[ar, b];
F4 = F4/ListIntegrate[F4, Prec];
EingangsEnt = HSE[ar] - ABase;
se1[[i*Rep, 1]] = EingangsEnt;
se1[[i*Rep, 2]] = HSE[F1];
se2[[i*Rep, 1]] = EingangsEnt;
se2[[i*Rep, 2]] = HSE[F2];
se3[[i*Rep, 1]] = EingangsEnt;
se3[[i*Rep, 2]] = HSE[F3];
se4[[i*Rep, 1]] = EingangsEnt;
se4[[i*Rep, 2]] = HSE[F4];
Print[i];
i++;]
Save["Falt2_SE_0005_50_5.m", se1];
Save["Gamma05_SE_0005_50_5.m", se2];
Save["MiWe_SE_0005_50_5.m", se3];
Save["ZaPlu_SE_0005_50_5.m", se4];

```

Deterministische Störung einer Variablen

```

Prec = 0.005;
DP = 30;
de1 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de2 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de3 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de4 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
a = TableGen[NorVert4[x, 0.5]];
b = TableGen[NorVert6[x, 0.5]];
even = TableGen[0.1*x/x];
ABase = HSE[a];

```

```

F1 = Falt2[a, b];
F1 = F1/ListIntegrate[F1, Prec];
F2 = Gamma05[a, b];
F2 = F2/ListIntegrate[F2, Prec];
F3 = MiWe[a, b];
F3 = F3/ListIntegrate[F3, Prec];
F4 = ZaPlu[a, b];
F4 = F4/ListIntegrate[F4, Prec];
F1Base = HSE[F1];
F2Base = HSE[F2];
F3Base = HSE[F3];
F4Base = HSE[F4];
Timing[For[i = 0, i <= DP,
    ar = (1 - i/DP)*a + i/DP*even;
    ar = ar/ListIntegrate[ar, Prec];
    br = b;
    F1 = Falt2[ar, br];
    F1 = F1/ListIntegrate[F1, Prec];
    F2 = Gamma05[ar, br];
    F2 = F2/ListIntegrate[F2, Prec];
    F3 = MiWe[ar, br];
    F3 = F3/ListIntegrate[F3, Prec];
    F4 = ZaPlu[ar, br];
    F4 = F4/ListIntegrate[F4, Prec];
    EingangsEnt = HSE[ar] - ABase;
    de1[[i*Rep + k, 1]] = EingangsEnt;
    de1[[i*Rep + k, 2]] = HSE[F1] - F1Base;
    de2[[i*Rep + k, 1]] = EingangsEnt;
    de2[[i*Rep + k, 2]] = HSE[F2] - F2Base;
    de3[[i*Rep + k, 1]] = EingangsEnt;
    de3[[i*Rep + k, 2]] = HSE[F3] - F3Base;
    de4[[i*Rep + k, 1]] = EingangsEnt;
    de4[[i*Rep + k, 2]] = HSE[F4] - F4Base;
    Print[i]
    i++;]
Export["Falt2_DE_005_30.txt", de1, "Table"];
Export["Gamma05_DE_005_30.txt", de2, "Table"];
Export["MiWe_DE_005_30.txt", de3, "Table"];

```

```
Export["ZaPlu_DE_005_30.txt", de4, "Table"];
```

Deterministische Störung zweier Variablen

```
Prec = 0.005;
DP = 30;
de1 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de2 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de3 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
de4 = Table[1, {j, 1, (DP + 1)}, {k, 1, 2}];
a = TableGen[NorVert4[x, 0.5]];
b = TableGen[NorVert6[x, 0.5]];
even = TableGen[0.1*x/x];
ABase = HSE[a];
F1 = Falt2[a, b];
F1 = F1/ListIntegrate[F1, Prec];
F2 = Gamma05[a, b];
F2 = F2/ListIntegrate[F2, Prec];
F3 = MiWe[a, b];
F3 = F3/ListIntegrate[F3, Prec];
F4 = ZaPlu[a, b];
F4 = F4/ListIntegrate[F4, Prec];
F1Base = HSE[F1];
F2Base = HSE[F2];
F3Base = HSE[F3];
F4Base = HSE[F4];
Timing[For[i = 0, i <= DP,
  ar = (1 - i/DP)*a + i/DP*even;
  ar = ar/ListIntegrate[ar, Prec];
  br = (1 - i/DP)*b + i/DP*even;
  br = br/ListIntegrate[br, Prec];
  F1 = Falt2[ar, br];
  F1 = F1/ListIntegrate[F1, Prec];
  F2 = Gamma05[ar, br];
  F2 = F2/ListIntegrate[F2, Prec];
  F3 = MiWe[ar, br];
  F3 = F3/ListIntegrate[F3, Prec];
  F4 = ZaPlu[ar, br];
  F4 = F4/ListIntegrate[F4, Prec];
```

```
EingangsEnt = HSE[ar] - ABase;
de1[[i*Rep + k, 1]] = EingangsEnt;
de1[[i*Rep + k, 2]] = HSE[F1] - F1Base;
de2[[i*Rep + k, 1]] = EingangsEnt;
de2[[i*Rep + k, 2]] = HSE[F2] - F2Base;
de3[[i*Rep + k, 1]] = EingangsEnt;
de3[[i*Rep + k, 2]] = HSE[F3] - F3Base;
de4[[i*Rep + k, 1]] = EingangsEnt;
de4[[i*Rep + k, 2]] = HSE[F4] - F4Base;
Print[i]
i++];]
Export["Falt2bo_DE_005_30.txt", de1, "Table"];
Export["Gamma05bo_DE_005_30.txt", de2, "Table"];
Export["MiWebo_DE_005_30.txt", de3, "Table"];
Export["ZaPlubo_DE_005_30.txt", de4, "Table"];
```

Ordnungsfunktionale

Laden der notwendigen Bibliotheken

```
<< Statistics`ContinuousDistributions`
<< Statistics`DataManipulation`
<< Graphics`Graphics`
<< Calculus`Integration`
<< NumericalMath`ListIntegrate`
```

Definitionen

```
Normal5d[x_] := NormalDistribution[5, 0.816497]
NorVert5d[x_] = PDF[Normal5d[x], x];
RandNormal002[x_] := 1 + Random[NormalDistribution[0, 0.2]]
Dreieck51[x_] := 0 /; x <= 3;
Dreieck51[x_] := (x - 3)/2 /; 3 < x <= 5;
Dreieck51[x_] := (7 - x)/2 /; 5 < x <= 7;
Dreieck51[x_] := 0 /; x > 7;
DPeak[x_] := 0 /; x <= 2;
DPeak[x_] := x - 2 /; 2 < x <= 3;
DPeak[x_] := 4 - x /; 3 < x <= 4;
DPeak[x_] := 0 /; 4 < x <= 6;
DPeak[x_] := x - 6 /; 6 < x <= 7;
DPeak[x_] := 8 - x /; 7 < x <= 8;
DPeak[x_] := 0 /; x > 8;
DFlat[x_] := 0 /; x <= 2;
DFlat[x_] := (x - 2)*1.8 /; 2 < x <= 2.5;
DFlat[x_] := (x - 2.5)/40 + 0.9 /; 2.5 < x <= 6.5;
DFlat[x_] := (7 - x)*2 /; 6.5 < x <= 7;
DFlat[x_] := 0 /; x > 7;
```

Berechnung

```
Prec := 0.00005;
Rep := 30000;
EWt[fun_] := Table[fun[x], {x, 0, 10, Prec}];
  NormTable = EWt[NorVert5d];
DrTable = EWt[Dreieck51];
DPTable = EWt[DPeak];
```

6. Anhang

```
F1Table = EWt[DFlat];
EWh[fun_] := Sum[fun[[i]]*Prec, {i, 1, 10/Prec}];
EW[fun_] :=
  Sum[fun[[i]]*(i - 1), {i, 1, 10/Prec}]/Sum[fun[[i]], {i, 1, 10/Prec}]*
  Prec;
MaxMean[fun_] := (Median[Position[fun, Max[fun]][[1]]] - 1)*Prec;
VD = Array[0, {3, Rep}];
VN = Array[0, {3, Rep}];
VP = Array[0, {3, Rep}];
VF = Array[0, {3, Rep}];
Timing[
  For[i = 1, i <= Rep,
    Stoe = Evaluate[Table[RandNormal002[x], {x, 0, 10, Prec}]];
    sNormTable = NormTable*Stoe;
    sDrTable = DrTable*Stoe;
    sDPTable = DPTable*Stoe;
    sF1Table = F1Table*Stoe;
    VN[[1, i]] = EW[sNormTable];
    VD[[1, i]] = EW[sDrTable];
    VP[[1, i]] = EW[sDPTable];
    VF[[1, i]] = EW[sF1Table];
    VN[[2, i]] = MaxMean[sNormTable];
    VD[[2, i]] = MaxMean[sDrTable];
    VP[[2, i]] = MaxMean[sDPTable];
    VF[[2, i]] = MaxMean[sF1Table];
    mN = Sum[sNormTable[[n]], {n, 1, 10/Prec}];
    mD = Sum[sDrTable[[n]], {n, 1, 10/Prec}];
    mP = Sum[sDPTable[[n]], {n, 1, 10/Prec}];
    mF = Sum[sF1Table[[n]], {n, 1, 10/Prec}];
    Tmp = 0;
    For[n = 1, Tmp < 0.7*mN, Tmp = Tmp + sNormTable[[n]]; n++];
    VN[[3, i]] = (n - 1)*Prec;
    Tmp = 0;
    For[n = 1, Tmp < 0.7*mD, Tmp = Tmp + sDrTable[[n]]; n++];
    VD[[3, i]] = (n - 1)*Prec;
    Tmp = 0;
    For[n = 1, Tmp < 0.7*mD, Tmp = Tmp + sDPTable[[n]]; n++];
    VP[[3, i]] = (n - 1)*Prec;
```

```

    Tmp = 0;
    For[n = 1, Tmp < 0.7*mD, Tmp = Tmp + sFlTable[[n]]; n++];
    VF[[3, i]] = (n - 1)*Prec;
    If[Mod[i, 500] == 0, Print[i]];
    i++;
]

```

Statistische Auswertung

```

AusT = Table[1, {i, 0, 6}, {j, 0, 4}];
AusT[[1, 1]] = "";
AusT[[1, 2]] = "Normalvert";
AusT[[1, 3]] = "Dreieck";
AusT[[1, 4]] = "Zwei-Peak";
AusT[[1, 5]] = "Flat";
AusT[[2, 1]] = "MW - EW";
AusT[[3, 1]] = "Var - EW";
AusT[[4, 1]] = "MW - Max";
AusT[[5, 1]] = "Var - Max";
AusT[[6, 1]] = "MW - 0.7-Quant";
AusT[[7, 1]] = "Var - 0.7-Quant";

AusT[[2, 2]] = Mean[VN[[1]]];
AusT[[4, 2]] = Mean[VN[[2]]];
AusT[[6, 2]] = Mean[VN[[3]]];
AusT[[3, 2]] = Variance[VN[[1]]];
AusT[[5, 2]] = Variance[VN[[2]]];
AusT[[7, 2]] = Variance[VN[[3]]];

AusT[[2, 3]] = Mean[VD[[1]]];
AusT[[4, 3]] = Mean[VD[[2]]];
AusT[[6, 3]] = Mean[VD[[3]]];
AusT[[3, 3]] = Variance[VD[[1]]];
AusT[[5, 3]] = Variance[VD[[2]]];
AusT[[7, 3]] = Variance[VD[[3]]];

AusT[[2, 4]] = Mean[VP[[1]]];
AusT[[4, 4]] = Mean[VP[[2]]];
AusT[[6, 4]] = Mean[VP[[3]]];

```

6. Anhang

```
AusT[[3, 4]] = Variance[VP[[1]]];  
AusT[[5, 4]] = Variance[VP[[2]]];  
AusT[[7, 4]] = Variance[VP[[3]]];
```

```
AusT[[2, 5]] = Mean[VF[[1]]];  
AusT[[4, 5]] = Mean[VF[[2]]];  
AusT[[6, 5]] = Mean[VF[[3]]];  
AusT[[3, 5]] = Variance[VF[[1]]];  
AusT[[5, 5]] = Variance[VF[[2]]];  
AusT[[7, 5]] = Variance[VF[[3]]];
```

```
Print[TableForm[AusT]];
```

D. Abbildungsverzeichnis

2.1	Entscheidungsproblem bei Sicherheit	19
2.2	Entscheidungsproblem bei Unsicherheit	21
2.3	Drei Nutzenzugehörigkeitsfunktionen	26
2.4	Bewertungsoperator für das Farbbeispiel	31
2.5	Qualitätsfunktion für verschiedene Einfüllzeiten	31
2.6	Qualität verschiedener Einfüllzeiten r	32
2.7	α -Schnitt	39
2.8	Kern	39
2.9	Spezifische Formen von ZGF	40
2.10	Klassischer Mittelwertoperator	41
2.11	Faltungsmittelwert	41
2.12	Zadeh-Plus-Operator für Zugehörigkeitsfunktionen	45
2.13	Gamma-Operator für Zugehörigkeitsfunktionen	45
2.14	Maximum-Likelihood-Prinzip	48
2.15	Schwerpunktmethode	49
2.16	$\epsilon (= 0.7)$ -Quantil	51
2.17	Max- α -Schnitt Prinzip	52
3.1	Zwei „Normalverteilungen“	60
4.1	Störung der Eingangsfunktionen für Bewertungsoperatoren	70
4.2	Ergebnisse für Faltungsoperator	71
4.3	Ergebnisse für Mittelwertoperator	72
4.4	Ergebnisse für Zadeh-Plus-Operator	73
4.5	Ergebnisse für Gamma-Operator	74
4.6	Deterministische Entropieerhöhung einer Variablen	75
4.7	Stochastische Entropieerhöhung einer Variablen	75
4.8	Deterministische Entropieerhöhung beider Variablen	76
4.9	Gamma-Operator bei Entropieerhöhung	78
4.10	Gamma-Operator bei unterschiedlichen Eingangserwartungswerten	79
4.11	Eingangsfunktionen zum Vergleich der Ordnungsfunktionale	82
4.12	i_1 überlagert von weißem Rauschen	83
4.13	Sensitivität des Maximumoperators	89

E. Verzeichnis der Definitionen

1	Entscheidungsmodell unter Sicherheit	18
2	Allgemeines Entscheidungsproblem	21
3	Parametrisierbarer Handlungsraum	22
4	Vollständiger Handlungsraum	22
5	Zustandsraum im Funktionalmodell	23
6	Vollständiger Zustandsraum	23
7	Informationsfunktion (IF)	24
8	Entscheidungsmodell als Funktional	25
9	Quasistetiges Entscheidungsmodell	27
10	Wahrscheinlichkeitsdichte	34
11	Fuzzy-Zugehörigkeitsfunktion (ZGF)	34
12	Träger	35
13	Beschränkte Zugehörigkeitsfunktion	35
14	Beschränkte Wahrscheinlichkeitsdichte	35
15	α -Schnitt	38
16	Kern	38
17	Normale Zugehörigkeitsfunktion	38
18	Einzelpunkt-Verteilung	39
19	Fuzzy-Zahlen und -Intervalle	39
20	Average-Maximum-Likelihood-Prinzip	47

21	Schwerpunktmethode ²	48
22	(μ, σ) -Prinzip	49
23	ϵ -Quantil	50
24	Max- α -Schnitt Prinzip	51
25	Min- α -Schnitt Prinzip	52
26	Hartley-Shannon-Entropie (HS-Entropie) [142]	56
27	Bernoulli-Prinzip/ Prinzip des fehlenden Grundes	58
28	Prinzip der maximalen Entropie	58

