

Fallbasierte Klassifikation in der Verfahrenstechnik am Beispiel der Rührorganauswahl

Diplomarbeit H II

von

Marion Brandner

Universität-GH Paderborn

Fachgebiet: Wissensbasierte Systeme
Prof. Dr. Kleine Büning

Betreuer: Benno Stein,
Andreas Brenke

Abgabetermin: 28.09.1994

Alle Urheberrechte liegen bei der Universität-GH Paderborn,
FB 17 - Fachgebiet Wissensbasierte Systeme

Inhaltsverzeichnis

1	Einführung	1
1.1	Ziel der Arbeit	1
1.2	Themenüberblick	1
2	Rührtechnische Grundlagen	2
2.1	Stoffeigenschaften	2
2.2	Grundaufgaben des Rührens	2
2.2.1	Homogenisieren	2
2.2.2	Suspendieren	3
2.2.3	Emulgieren	3
2.2.4	Begasen	3
2.2.5	Wärmeübergang	3
2.3	Rührwerke	4
2.3.1	Rührorgane	4
2.4	Leistungsbedarf	7
3	Wissenserwerb	8
3.1	Expertenwissen	8
3.2	Datensammlung	9
3.2.1	Inkonsistenz	10
3.2.2	Repräsentative Beispielmenge	12
3.3	Prinzipieller Aufbau der Lösungsansätze	12
4	Statistischer Ansatz	13
4.1	Beschreibung	13
4.1.1	Trainingsphase	14
4.1.2	Klassifikationsphase	15
4.2	Anwendungsergebnisse	16
5	Unscharfe Logik - Der Fuzzy Ansatz	22
5.1	Beschreibung	22
5.1.1	Linguistische Variablen und Zugehörigkeitsfunktionen	23
5.1.2	Fuzzy Regeln und Operatoren	23

5.1.3	Defuzzifizierung	24
5.2	Allgemeine Vorgehensweise	25
6	Generierung einer Regelmenge	26
6.1	Struktur der Regeln des Rührproblems	26
6.2	Methoden zur Regelerzeugung aus einer Datensammlung	27
6.2.1	Direkte Ableitung von Zusammenhängen	27
6.2.2	Der ID3 Algorithmus	28
6.3	Darstellung der Prämisse	29
6.4	Modellierungstiefe	32
6.4.1	Beschreibung	32
6.4.2	Gegenüberstellung	33
6.4.3	Verknüpfung der Rührcharakteristika	35
7	Verfahren zur Beurteilung der Regelgüte	42
7.1	Regelbewertung durch Verrechnung der Fuzzywerte	42
7.1.1	Übereinstimmungsgrad der Prämisse	43
7.1.2	Übereinstimmungsgrad der Konklusion	43
7.1.3	Verknüpfung von Prämisse und Konklusion	44
7.1.4	Beurteilung der Regel	44
7.1.5	Quantitative Bewertung	45
7.2	Gewichtung anhand neuronaler Netze	45
7.2.1	Allgemeine Beschreibung	46
7.2.2	Der Backpropagation Algorithmus	47
7.2.3	Verknüpfung von Regelmenge und neuronalen Netzen	49
7.2.4	Netzwerkstruktur zur Rührerauswahl	49
7.2.5	Einbeziehung der Fallsammlung in den Lernalgorithmus	50
7.2.6	Transformation der Verbindungsstärken zu Regelbewertungen	51
7.3	Vergleich der Gewichtungsverfahren	58
7.3.1	Übersicht	58
7.3.2	Testergebnisse	58
7.3.3	Größe der Regelmengen	59
7.3.4	Aufwand	60

7.3.5	Umgang mit Inkonsistenzen der Datensammlung	61
8	Regelbasierte Objektauswahl	62
8.1	Regelverarbeitung	62
8.1.1	Einbeziehen der Regelbewertung	62
8.1.2	Propagieren der Zahl begründender Fälle	64
8.2	Klassifikation	65
8.2.1	Auswahlkriterien	65
8.2.2	Vergleich von Defuzzifizierung und “best-of” Auswahl	68
9	Erweiterte Problemstellung	70
9.1	Eignung von neu entwickelten Rührern	70
9.2	Verallgemeinerte Fragestellungen	73
10	Zusammenfassung	74
A	Hinweise zur Implementation	79
A.1	Statistischer Ansatz	79
A.2	Regelbasierter Ansatz	80
A.2.1	Datenstrukturen	80
A.2.2	Funktionen	84
A.3	Benutzungshinweise	90
A.3.1	Darstellung der Datensammlung	90
A.3.2	Beispiel	91
A.3.3	Vorliegende Datensammlungen	92

1 Einführung

1.1 Ziel der Arbeit

Die speziellen Zusammenhänge und Problemstellungen einer Domäne erfordern unterschiedliche Repräsentationsformen und Lösungswege.

Ziel dieser Arbeit ist die Entwicklung und Beurteilung verschiedener Ansätze zur Darstellung und Verarbeitung von rührtechnischem Wissen, das den Aspekt der Rührerauswahl bei der Auslegung eines Rührwerks als zentrale Fragestellung untersucht.

Als Grundlage zur Lösung dieses verfahrenstechnischen Problems wird eine Sammlung von Rührwerksauslegungen benutzt, auf die alle betrachteten Auswahlmethoden aufbauen¹. Zusätzlich werden technische Zusammenhänge und Erfahrungen eines Ingenieurs in die Verarbeitung einbezogen, um eine der menschlichen Vorgehensweise ähnliche Abarbeitung zu gewährleisten.

Unter Berücksichtigung der gewonnenen Ergebnisse kann ein System bereitgestellt werden, das auch für andere fallbasierte Klassifikationsprobleme Anwendung finden kann.

1.2 Themenüberblick

Im folgenden werden rührtechnische Grundlagen erläutert und Gesichtspunkte des Wissenserwerbs behandelt. Die anschließenden Kapitel erörtern diverse fallbasierte Techniken zum Problem der Rührerauswahl, wobei die Verfahren zunächst beschrieben werden und danach die aus der Umsetzung am Rechner gewonnenen Ergebnisse untersucht werden.

Zuerst wird eine auf statistischen Methoden basierende Auswertung der gegebenen Rührwerksauslegungen durchgeführt. Zur Vermeidung der dort auftretenden Mängel werden alternative Ansätze betrachtet, deren Grundlage die unscharfe Logik (Fuzzy Logic) bildet. Diese Ansätze stützen sich auf eine Regelmenge, die aus der Fallsammlung generiert wird. Unterschiedliche Regelstrukturen und Verarbeitungsmethoden werden vorgestellt und diskutiert.

Im letzten Kapitel wird einerseits untersucht, inwiefern der regelbasierte Ansatz als Grundlage für andere Klassifikationsprobleme dienen kann. Andererseits werden weiterführende Fragestellungen erörtert, die unter Ausnutzung der gewonnenen Informationsmenge gelöst werden können.

¹Die hier verwendete Datensammlung wurde freundlicherweise von der Firma *Stelzer Rührtechnik GmbH*, Warburg zur Verfügung gestellt.

2 Rührtechnische Grundlagen

Zu der verfahrenstechnischen Grundoperation Rühren gehören alle Mischvorgänge, bei denen die flüssige Phase überwiegt.² Die Auslegung eines Rührwerks hängt zum einen von der Beschaffenheit der zu vermischenden Fluide ab und zum anderen von den geforderten Rühraufgaben.

2.1 Stoffeigenschaften

Die für die Konfiguration des Rührwerks relevanten Stoffwerte eines Rührmediums sind Viskosität und Dichte. Die Viskosität (Zähigkeit) kann als Maß für die innere Reibung gesehen werden und stellt ein wichtiges Kriterium bei der Beurteilung der zu lösenden Problemstellung dar. Die verschiedenen Rührorgane eignen sich jeweils nur für einen bestimmten Viskositätsbereich. Desweiteren trägt die Zähigkeit zusammen mit der Dichte zum Leistungsbedarf und damit zum erforderlichen Energieaufwand bei, der bei der Wahl des Antriebs und der Rührwelle berücksichtigt werden muß.

2.2 Grundaufgaben des Rührens

Die vielfältigen Aufgabenstellungen für technische Rührprozesse kann man systematisch beschreiben, indem man sie als Kombination von bestimmten Basisaufgaben auffaßt. Sollen mehrere Grundrühraufgaben gleichzeitig erfüllt werden, so ist zur Rührerauswahl eine Untersuchung der Rangfolge der Einzelaufgaben notwendig. Man unterscheidet folgende Grundaufgaben der Rührtechnik:

2.2.1 Homogenisieren

Homogenisieren bezeichnet den Erzeugungsprozeß bzw. das Aufrechterhalten einer einheitlichen Phase aus ineinander löslichen Flüssigkeiten. Diese Vermischung soll vor allem zum Konzentrations- oder Temperatúrausgleich der unterschiedlichen Komponenten führen. Als Anwendungsgebiet ist die Zusammenführung der Reaktionspartner einer chemischen Reaktion (Neutralisieren, Verdünnen, Temperieren usw.) zu nennen.

Ein wichtiges Kennzeichen von Homogenisierungsvorgängen ist die Mischzeit, die zum Erreichen einer definierten Mischgüte notwendig ist. Rührorgane besitzen spezifische Mischzeitcharakteristiken, die als Auswahlkriterium bei der Auslegung eines Rührwerks eine Rolle spielen.

²Detaillierte Informationen zum Thema *Rühren* können [Sti94] entnommen werden.

2.2.2 Suspendieren

Die gleichmäßige Verteilung eines dispersen Feststoffes in einer Flüssigkeit wird als Suspendieren bezeichnet. Beispielhaft ist das Lösen von Feststoffen und das Aufwirbeln zur Auswaschung von Feststoffen zu nennen. Es werden mehrere Suspendierzustände unterschieden, einerseits das In-Schwebe-Halten der Feststoffe bis zu einer bestimmten Behälterhöhe, andererseits wird die Verweildauer der Partikel am Boden betrachtet.

2.2.3 Emulgieren

Emulgieren bezeichnet die Vermischung von zwei nicht ineinander löslichen Flüssigkeiten. Dabei werden durch Zerteilen der einen Phase Tropfen gebildet (disperse Phase), die in der anderen Phase verteilt werden (kontinuierliche Phase). Dieser Vorgang dient dazu, eine möglichst große wirksame Oberfläche zu erzeugen, die einen Stoffaustausch beschleunigt, oder die Trennung der beiden Fluide zu verlangsamen oder zu unterbinden.

2.2.4 Begasen

Die Grundrühraufgabe Begasen ist eng mit Emulgierprozessen verbunden, jedoch handelt es sich hier um eine Feinverteilung von Gasblasen in einer Flüssigkeit. Ziel dieses Verfahrens ist wiederum die Erzeugung einer großen Austauschfläche, die einen Stoffübergang begünstigt.

Zur Realisierung eines Begasungssystems bieten sich mehrere Möglichkeiten an. Zum einen kann durch Fremdbegasung über eine Ringbrause oder Lanze das Gas in die Flüssigkeit eingeführt werden, alternativ dazu kann das Prinzip der Selbstansaugung über eine Trombe oder einen Hohlrührer angewandt werden (siehe [Sti94]).

2.2.5 Wärmeübergang

Ziel dieser Grundrühraufgabe ist eine kontrollierte Temperaturführung einer Flüssigkeit, die durch einen Heiz- oder Kühlmantel um den Rührbehälter oder durch Rohrschlangen in ihm ermöglicht wird. Aufgrund der geringen Wärmeübertragungsfläche im Verhältnis zum Behältervolumen soll mit Hilfe des Rührorgans stets das Produkt mit größtmöglichem Temperaturgradienten die Heiz- oder Kühlfläche anströmen, um so eine Übertragung größerer Wärmemengen pro Zeiteinheit zu gewährleisten.

2.3 Rührwerke

Ein Rührwerk besteht im wesentlichen aus den Komponenten *Rührorgan*, *Welle* und *Antrieb*. Ausschlaggebend für eine vorteilhafte Auslegung der Rühranlage ist der Rührorgantyp, der sich durch verschiedene Eigenschaften auszeichnet, die die Realisierung des zu lösenden Rührproblems ermöglichen...

Abb.2.1: Rührwerk mit Behälter

figure=ruehrwerk2.ps

Bezüglich der Anordnung der Rührwelle im Behälter kommt eine schräge, zentrierte oder exzentrische Lage in Frage, wobei in dieser Untersuchung allein zentrierte Einbauten berücksichtigt werden. Gleichmaßen spielen Koaxialrührwerke nur eine untergeordnete Rolle. Es handelt sich hierbei um eine Zusammenfassung zweier Rührwerke, wobei die beiden Rührwellen koaxial um eine gemeinsame geometrische Mittelachse angeordnet sind. Die Rührwellen können mit unterschiedlichen Rührorganen bestückt werden, die bestimmten Anforderungen genügen. Koaxialrührwerke gelten in dieser Arbeit allgemein als Sonderform.

2.3.1 Rührorgane

Im folgenden werden mehrere Rührorgantypen vorgestellt und deren Eignung im Hinblick auf bestimmte Aufgabenstellungen beschrieben. Neben klassischen Rührern werden auch Eigenentwicklungen der Firma Stelzer angeführt, die als Grundlage der Untersuchungen dienen (siehe [Ste93]).

Die Förderrichtung der Rührer stellt ein wichtiges Kennzeichen dar. Axial fördernde Rührer leiten das Rührgut zunächst in Richtung Behälterboden und dann entlang der Wand nach oben. Radiale Strömungsfelder verlaufen vom Rührer ausgehend horizontal zur Behälterwand. Eine kreisförmige Bewegung auf einer Ebene wird von tangential fördernden Rührern erzeugt, während Gegenstromrührer das Produkt im Zentrum nach unten und an der Behälterwand nach oben fördern.

Eine systematische Einteilung der Rührer stellt die Aufgliederung in Schnell- und Langsamläufer dar.

1. Schnellläufer

Rührorgane dieser Klasse zeichnen sich durch eine hohe Drehzahl und ein kleines Durchmesser Verhältnis Rührorgan zu Behälter (d_2/d_1) aus. Dadurch konzentriert sich das Einsatzgebiet von Schnellläufern auf niederviskose Produkte, damit eine auf den gesamten Behälterinhalt ausgedehnte Strömung erzielt werden kann.

- Propellerrührer

Dieser Rührer wird im turbulenten Strömungsbereich am häufigsten eingesetzt. Er besteht in der Regel aus drei Flügeln und zeichnet sich durch eine sehr strömungsgünstige Form aus. Der Propellerrührer erzeugt ein primär axiales Strömungsfeld und eignet sich daher gut für Suspendieraufgaben. Desweiteren findet der Rührer beim Homogenisieren und bei einfachen Wärmeübertragungen Verwendung.

- Scheibenrührer
Beim Scheibenrührer sind in der Regel sechs senkrecht stehende Rechteckblätter auf einer Kreisscheibe angeordnet und bewirken durch die kantige Form hohe Scherkräfte. Aus diesem Grund erweist sich der Rührer für Dispergierprobleme als vorteilhaft. Die radiale Förderichtung schränkt den weiteren Einsatzbereich auf die Kombination mit Wärmeübertragung oder Homogenisieren ein.

- Schrägblattrührer
Dieser Rührer besteht häufig aus sechs Blättern, die mit konstantem Winkel angestellt sind. Aufgrund eines höheren Leistungseintrags als beim Propellerrührer benötigt dieses Rührorgan eine geringere Umfanggeschwindigkeit und erzielt höhere Scherkräfte. Damit ergeben sich gute Voraussetzungen zur Lösung von Homogenisier- und Suspendieraufgaben. Durch die radialen Anteile im primär axialen Strömungsfeld kommt ebenfalls die Verwendung des Schrägblattrührers bei Wärmeübergang an der Behälterwand in Frage.

2. Langsamläufer

Im Gegensatz zu den oben beschriebenen Rührern besitzen Langsamläufer ein größeres Durchmesser Verhältnis d_2/d_1 , das bei wandgängigen Typen über 0.9 liegen kann, und rotieren mit geringerer Umfanggeschwindigkeit. Sie finden überwiegend Anwendung bei Rühraufgaben in Medien höherer Viskosität. Niederviskose Produkte können schonender aufgrund des großflächigen Energieeintrags verarbeitet werden.

- Kreuzbalkenrührer
Der Kreuzbalkenrührer ist aus zwei rechteckigen Rührarmen zusammengesetzt, die unter 45° angestellt sind. Häufig wird dieses Rührorgan mehrstufig konstruiert, wobei die einzelnen Stufen kreuzweise, d.h. um 90° versetzt angeordnet sind. Hauptrühraufgaben sind Homogenisieren, Suspendieren und Wärmeübergang. Wegen der axialen Förderrichtung ist das Durchmesser Verhältnis auf 0.7 bis 0.8 beschränkt.
- Trapezrührer
Im Gegensatz zum Kreuzbalkenrührer werden die Rührarme des Trapezrührers nach außen schmaler und sind flacher (24°) angeordnet, wodurch sich ein geringerer Leistungsbedarf ergibt.
- Sigmarührer
Diese Eigenentwicklung der Firma Stelzer entspricht dem Aufbau eines Kreuzbalkenrührers mit Umsetzung des Gegenstromprinzips. Diese Strömungsform wird durch inverse Anstellwinkel des wellennahen und des äußeren Teils der Rührblätter erreicht, so daß das Rührgut im Zentrum nach unten und wandnah nach oben gefördert wird. Damit ist eine äußerst gleichmäßige Verteilung der eingetragenen Energie gewährleistet. Desweiteren bietet sich der Einsatz von Sigmarührern mit großem Durchmesser Verhältnis in höherviskosen Medien an.

- Mehrstufen-Impuls-Gegenstrom-Rührer (MIG)
Bezüglich der Form und des Strömungsbildes besitzt der MIG-Rührer fast identische Eigenschaften des Sigmarrührers. Allerdings wird der MIG-Rührer in dieser Untersuchung ausschließlich in hochviskosen Fluiden eingesetzt, so daß eine Unterscheidung beider Rührformen sinnvoll erscheint.
- Alpharührer
Für Homogenisierungs- und Wärmetauschaufgaben im hochviskosen Bereich findet der von der Firma Stelzer entwickelte Alpharührer Verwendung. Dieses Rührorgan besitzt kreuzbalkenförmige Innenarme, deren Verlängerung an die Behälterwand angepaßte Außenblätter bilden und insgesamt eine Gegenstromförderung bewirken. Mit einem Durchmesser Verhältnis, das gewöhnlich über 0.9 liegt, stellt der Rührertyp einen Ersatz des klassischen Wendelrührers dar.
- Ankerrührer
Charakteristisch für die Form des Ankerrührers sind die beiden wandnahen, der Behälterform angeglichenen Rührarme, die sich vor allem zur Wärmeübertragung bei höherviskosen Produkten eignen. Aufgrund der tangentialen Förderrichtung fällt den übrigen Aufgabenstellungen eine untergeordnete Bedeutung zu, so daß sich eine Kombination mit anderen Rührern anbietet, die entsprechende Förderleistungen gewährleisten.

2.4 Leistungsbedarf

Jede Rührwerksauslegung erfordert eine Untersuchung des Leistungsbedarfs, um eine geeignete Wahl des Antriebs und der Rührwelle zu ermöglichen. Die Rührleistung P entspricht der Nettoleistung, die vom Rührer in die Flüssigkeit eingebracht wird. Sie berechnet sich aus der vom Motor aufzubringenden Bruttogleistung abzüglich der Reibungsverluste.

Die Leistung wird von der Dichte ρ , der Viskosität η , der Drehzahl n und dem Rührerdurchmesser d beeinflusst und berechnet sich wie folgt³:

$$P = Ne \cdot \rho \cdot n^3 \cdot d^5$$

Die Größe Ne bezeichnet die Newton-Zahl, die anhand der Gleichung $Ne = f(Re)$ bestimmt werden kann und den Strömungswiderstand kennzeichnet. Re steht für die Rührer-

³siehe [Sti94], S. 207

Reynolds-Zahl, für die gilt⁴:

$$Re = \frac{n \cdot d^2 \cdot \rho}{\eta}$$

Für jeden Rührer beschreibt die sogenannte *Leistungscharakteristik* die Abhängigkeit von Ne und Re . Die Kurvenverläufe zeichnen sich durch drei Strömungsbereiche aus: die laminare Strömung, die turbulente Strömung und der Übergangsbereich dazwischen. Der Einfluß der Zähigkeit auf die Leistung nimmt vom laminaren zum turbulenten Bereich hin ab.

Um bei der Rührwerksauslegung die verschiedenen Rührorgane vergleichen zu können, erweist sich die Betrachtung der spezifischen Leistung, die sich aus Nettoleistung bezogen auf das Füllvolumen ergibt, als vorteilhaft.

3 Wissenserverb

Grundlegend für die Entwicklung eines wissensbasierten Systems sind die Informationsquellen, aus denen Domänenwissen zur Akquisition und Verarbeitung entnommen wird.

Der Wissenserverb zu dieser rührtechnischen Anwendung basiert auf dem Erfahrungswissen eines Experten und auf einer Sammlung von Rührwerksauslegungen.

3.1 Expertenwissen

Bisher entwickelte Systeme zur Konfigurierung eines Rührbehälters mit Rührwerk unterstützen den Verfahrenstechniker bei der Auslegung der Rührwelle, des Antriebs und des Behälters, während die Wahl des Rührorgans dem Ingenieur zukommt. Die Entscheidung für einen bestimmten Rührertyp fällt der Experte, indem er Kriterien untersucht, die die Eigenschaften der Rührer und die zu lösende Aufgabenstellung in Zusammenhang bringen. Hier spielen Erfahrungswissen und Heuristiken eine große Rolle.

Es ist sinnvoll, das Expertenwissen für eine rechnergestützte Lösung des Problems der Rührerauswahl zu nutzen und die Darstellung und Verarbeitung dieser Informationen entsprechend anzupassen. Einen geeigneten Ansatz bietet das Modell der unscharfen Logik, das als zentrale Methode in dieser Arbeit untersucht wird.

Insgesamt konnten folgende Aspekte des Wissenserverbs vom Ingenieur bereitgestellt werden:

- Für jeden der in Abschnitt 2.3.1 vorgestellten Rührorgane wird eine Charakterisierung abgelegt, die die wesentlichen Eigenschaften des Rührertyps erfaßt. Ansatzweise wurden im vorangegangenen Kapitel bereits einzelne Rührermerkmale aufgeführt. Die Menge der relevanten Merkmale setzt sich wie folgt zusammen:

⁴siehe [Sti94], S. 208

- *Grundrühraufgaben*, für die der Rührer geeignet ist,
- *Viskositätsbereich*, in dem der Rührer eingesetzt werden kann,
- *Drehzahlbereich*,
- *Durchmesser Verhältnis*,
- *Förderrichtung*,
- *Form*,
- *Stufenzahl*.

Daß diese Eigenschaften nahezu vollständig sind, beruht auf dem Erfahrungswissen des Experten und bildet eine weitere Information des Ingenieurs.

- Um die in Kapitel 5 verwendeten unscharfe Mengen modellieren und verarbeiten zu können, ist die Angabe von *Merkmalsabstufungen* und *Zugehörigkeitsfunktionen* für bestimmte Parameter notwendig. Diese Definitionen werden vom Experten bereitgestellt. Eine detaillierte Beschreibung der unscharfen Logik erfolgt in Kapitel 5.
- Verschiedene heuristische Kriterien werden in das Auswahlverfahren eingebettet. Beispielhaft ist die Vernachlässigung der Grundrühraufgaben *Homogenisieren* und *Wärmeübertragung* zu nennen, wodurch die Hervorhebung der übrigen geforderten Grundaufgaben ermöglicht wird.

Im folgenden wird explizit darauf verwiesen, an welcher Stelle Erfahrungswerte in die Verarbeitung einfließen.

3.2 Datensammlung

Aufgrund der Komplexität des Erfahrungswissens eines Experten ist eine vollständige Wiedergabe und verbale Formulierung des Wissens nicht möglich. Es bietet sich jedoch an, die vom Ingenieur entwickelten Lösungen zu betrachten und daraus Informationen herauszufiltern.

In der hier untersuchten Fragestellung einer geeigneten Rührerauswahl steht eine Datensammlung von 188 Rührwerksauslegungen zur Verfügung, die die Basis für die nachfolgenden, fallbasierten Ansätze bildet. Folgende Parameter beschreiben ein Fallbeispiel:

- Grundrühraufgaben
- Stoffeigenschaften (Viskosität, Dichte)
- Behälterabmessungen (Höhe und Durchmesser)
- Füllvolumen
- Rührorgan

- Rührerdurchmesser
- Stufenzahl
- Drehzahl
- Leistung
- Wellendurchmesser
- Anzahl der Stromstörer

Diese Rührwerksauslegungen beinhalten implizit eine Reihe von Informationen über die Vorgehensweise des Verfahrenstechnikers beim Entwurf der Anlagen. Die in dieser Arbeit untersuchten Methoden haben das Ziel, versteckte Zusammenhänge und Regeln aus der Fallsammlung aufzufinden, darzustellen und verarbeiten.

3.2.1 Inkonsistenz

Betrachtet man die in Kapitel 2.3.1 beschriebenen Rührorgane in Bezug auf den Einsatzbereich, so fällt auf, daß für viele Aufgabenstellungen mehrere Rührer in Frage kommen. Die Wahl eines bestimmten Rührorgans trifft der Ingenieur, indem er neben den Parametern *Aufgabe* und *Viskosität* weitere Anforderungen untersucht, beispielsweise eine schonende Verarbeitung des Rührguts. Außerdem müssen Kundenwünsche berücksichtigt werden, die die Auslegung beeinflussen. Daß sich mehrere Rührer für eine Problemstellung eignen, spiegelt sich in der vorliegenden Datensammlung wider. Hier liegt keine eindeutige Zuordnung von Rührern bezüglich einer Aufgabenstellung vor. Es handelt sich daher um eine inkonsistente Menge von Beispielen. Tabelle 3.1 gibt an, welche Rührer für bestimmte Aufgabenstellungen gewählt worden sind.

Tab.3.1: Wahl der Rührorgane in der Datensammlung.
 Grundrühraufgaben sind durch den Anfangsbuchstaben gekennzeichnet.

Grundaufgaben					Viskositätsbereich				
H	S	E	W	B	sehr niedrig	niedrig	mittel	hoch	sehr hoch
×					<u>Propeller</u> Schrägblatt Sigma Trapez Scheibe Kreuz MIG Sonderform	<u>Propeller</u> Schrägblatt Sigma Trapez Scheibe Sonderform	<u>Propeller</u> Schrägblatt <u>Sigma</u> Trapez Kreuz Anker	<u>Sigma</u> <u>Trapez</u> Kreuz Anker Alpha	<u>Anker</u> <u>Alpha</u>
×			×		<u>Propeller</u> <u>Schrägblatt</u> Scheibe Sigma Anker	<u>Propeller</u> Sigma	Schrägblatt <u>Sigma</u> Scheibe	<u>Kreuz</u> Anker MIG	Sigma Kreuz Anker <u>MIG</u> Sonderform
×	×		×			<u>Propeller</u> Schrägblatt <u>Sigma</u> Kreuz	Schrägblatt <u>Sigma</u> Alpha Sonderform	<u>Propeller</u> Schrägblatt <u>Sigma</u> <u>Kreuz</u>	
×	×				<u>Propeller</u> Sigma Trapez	<u>Propeller</u> <u>Schrägblatt</u> Trapez	Schrägblatt <u>Sigma</u>		
	×				<u>Propeller</u> Kreuz	<u>Propeller</u> Sigma Sonderform	<u>Propeller</u> Scheibe	<u>Propeller</u>	
×			×	×	<u>Sonderform</u> Scheibe	<u>Sonderform</u>			
			×				<u>Sigma</u>	<u>Sigma</u>	
×		×			<u>Sigma</u>	<u>Sigma</u>			
×		×	×		<u>Sonderform</u>			<u>Sigma</u> Schrägblatt	
		×			<u>Scheibe</u>			<u>Sigma</u>	<u>Sigma</u>
×		×	×		<u>Propeller</u> Sigma			<u>Kreuz</u>	
×	×	×			<u>Sigma</u>				
×		×	×			<u>Propeller</u>			
×				×	<u>Scheibe</u>				

Für die Lösungsverfahren stellt sich die Aufgabe, sinnvoll mit diesen Inkonsistenzen umzugehen. Einzelheiten dazu werden in den nachfolgenden Kapiteln behandelt.

3.2.2 Repräsentative Beispielmenge

Zur Beurteilung der verschiedenen Lösungsansätze ist es notwendig, anhand einer Menge von bekannten Rührwerksauslegungen Testreihen durchzuführen. Die Häufigkeit der Lösungsübereinstimmung gibt Aufschluß über die Qualität eines Verfahrens und ermöglicht den Vergleich unterschiedlicher Ansätze.

Als Testgrundlage kommt eine Teilmenge der vorliegenden Datensammlung in Frage, die konsistent ist und alle in der Fallsammlung auftretenden Problemstellungen repräsentiert. Der Experte hat hierzu eine Auswahl von 36 Beispielen getroffen, die in Tabelle 3.1 unterstrichen dargestellt sind.

Entfernt man nun diese Teilmenge bei der Verarbeitung aus der Datensammlung, um eine unabhängige Testgrundlage zu erhalten, so bedeutet dies für die verbleibende Datensammlung eine derart starke Einschränkung, daß keine ausreichende Grundlage zur Wissensakquisition mehr gegeben ist. Beispielsweise liegt dann zu den Grundrühraufgaben *Emulgieren* und *Begasen* allein ein Fallbeispiel in der reduzierten Fallsammlung vor. Aus diesem Grunde ist die Entfernung der Testbeispiele aus der Datensammlung nicht zweckmäßig.

Da jedoch mit den nachstehenden Verfahren stets ein Informationsverlust verbunden ist, ist es möglich, die Testmenge in der Verarbeitungsmenge zu belassen.

3.3 Prinzipieller Aufbau der Lösungsansätze

Das Problem der Rührerauswahl besteht darin, für eine vorgegebene *Rühraufgabe* eines Rührguts bestimmter *Viskosität* eine geeignete Zuordnung eines *Rührorgans* zu treffen.

Zur Lösung dieses Problems wurden zwei Herangehensweisen ausgearbeitet.

A. Direkte Klassifikation

Die naheliegendste Formulierung stellt die Beschreibung des *direkten Zusammenhangs* von Eingabegrößen und der gesuchten Größe dar. Das bedeutet, daß in jedem Verfahren versucht wird, durch Ausnutzung der vorhandenen Informationen einen Zusammenhang von Rühraufgabe und Viskosität mit einem Rührorgan herzustellen.

figure=direct.ps

Abb.3.1: Direkter Klassifikationszusammenhang

B. Klassifikation über Rührermerkmale

Idee: Die rechnergestützte Verarbeitung soll ähnlich zu der Vorgehensweise eines Ingenieurs aufgebaut werden. Dazu ist ein Ansatz notwendig, der auf einer *tieferen Modellierung* basiert.

Vorgehen eines Ingenieurs: Der Experte untersucht jede rührtechnische Aufgabenstellung hinsichtlich der erforderlichen Eigenschaften, denen ein geeignetes Rührorgan entsprechen sollte⁵. Diesbezüglich kann derjenige Rührer selektiert werden, dessen Merkmale den größten Ähnlichkeitsgrad mit den ermittelten Anforderungen aufweisen.

Übertragung auf den Lösungsprozeß: Zunächst werden die notwendigen Rühreigenschaften aus der Aufgabenstellung ermittelt. Durch Ausnutzung des Zusammenhangs von Rührorgan und zugehörigen Merkmalen kann derjenige Rührertyp ausgewählt werden, der die größte Ähnlichkeit mit den Anforderungen aufweist.

figure=tiefmod.ps

Abb.3.2: Tiefere Modellierung

4 Statistischer Ansatz

Die naheliegendste Möglichkeit, die vorhandenen Daten auszuwerten und zur Lösung des Problems der Rührerauswahl zu verwenden, stellt eine statistische Untersuchung der Datensammlung dar. Im folgenden wird dieser Ansatz zur Klassifikation vorgestellt und anhand der erzielten Ergebnisse der Implementation beurteilt.

4.1 Beschreibung

Allgemein läßt sich das Klassifikationsproblem wie folgt beschreiben (vergleiche [YoCa74] und [Gr69]):

Gegeben ist ein N -dimensionaler Merkmalvektor x , der anhand einer Entscheidungsfunktion $\delta(x)$ einer Klasse C zugeordnet wird.

Die Entscheidungsfunktion hat zur Aufgabe, den Musterraum Ω_x , der die Menge aller möglichen Werte von x enthält, derart aufzuteilen, daß disjunkte Entscheidungsräume für jede Klasse entstehen, um so eine Zuordnung mit minimaler Fehlklassifikation zu ermöglichen.

Übertragen auf die Fragestellung der Rührerauswahl bedeutet dies, daß der Merkmalvektor bestehend aus Rühraufgabe und Viskosität einem bestimmten Rührorgan zugeordnet

⁵Anforderungen, die ein Rührer abhängig von den Stoffeigenschaften erfüllen sollte, wurden in Kapitel 2.3.1 dokumentiert.

wird. Hierzu ist es notwendig, unter Zuhilfenahme der Datensammlung eine geeignete Regel zu finden, die den zweidimensionalen Musterraum in Entscheidungsräume für jeden Rührer aufteilt.

Jeder vorliegende Fall ist charakterisiert durch den Merkmalvektor x und die Zuordnung zu einem Rührorgan C . Dieser Zusammenhang muß ausgenutzt werden, um anhand der Verteilungsstruktur von x neue Muster klassifizieren zu können. Da die Verteilungsfunktion des Merkmalvektors nicht bekannt ist, ergeben sich für die statistische Klassifikation zwei Schritte:

1. Die Trainingsphase, in der die Dichtefunktion anhand der Beispielsammlung bestimmt wird.
2. Die Klassifikationsphase, in der eine neue Aufgabenstellung in Form eines Merkmalvektors einer Klasse zugeordnet wird.

4.1.1 Trainingsphase

In diesem Teil wird für jede Klasse von Rührern eine Dichtefunktion erstellt, die die Zugehörigkeit der Merkmalvektoren zu dieser Klasse beschreibt. Bei der hier vorliegenden Problemstellung kann man keine allgemeine Verteilungsfunktion voraussetzen, wie beispielsweise eine Normalverteilung, in der die Parameter Mittelwert und Varianz zu bestimmen sind, da das Merkmal Rühraufgabe einen diskreten Wertebereich besitzt. Daher wird die Dichtefunktion anhand der Häufigkeitsverteilung ermittelt, indem eine Einteilung des zweidimensionalen Musterraums in Sektoren erfolgt entsprechend der beiden Merkmale:

1. Die Menge der Rühraufgaben, die alle Kombinationen der fünf Grundrühraufgaben beinhaltet und somit einen diskreten Wertebereich von 32 Elementen darstellt.
2. Die Viskosität, die einen stetigen Wertebereich besitzt, und zur Erstellung der Dichtefunktionen in fünf Intervalle (“sehr niedrig”, “niedrig”, “mittel”, “hoch”, “sehr hoch”) unterteilt wird⁶.

Für jede Klasse C_k sei n die Anzahl aller Fallbeispiele aus C_k und n_t die Zahl der Beispiele mit Merkmalvektor x , der auf den Sektor t des Musterraums Ω abgebildet wird und zur Klasse C_k gehört. Dann ergibt sich für alle x die Dichtefunktion f :

$$f(x) = \frac{n_t}{n}$$

⁶Es ist zu beachten, daß diese Aufteilung eine Vergrößerung der Informationen darstellt, und damit ein Informationsverlust verbunden ist.

4.1.2 Klassifikationsphase

Um eine eindeutige Zuordnung zu einer Klasse gewährleisten zu können, bietet sich eine Grenzwertklassifikation an. Ziel der Grenzwertklassifikation ist die Auffindung einer Entscheidungsfunktion ρ , die den Musterraum Ω_x in K disjunkte Entscheidungsräume $\Omega_1, \dots, \Omega_K$ aufteilt und optimal bezüglich einer Fehlklassifikation ist.

In dieser Arbeit wurde dabei das Bayes-Kriterium zugrunde gelegt, welches das Vorliegen der a-priori-Wahrscheinlichkeiten p_k für jede Klasse C_k , $k \in \{1, \dots, K\}$ voraussetzt (siehe unten). Seien

$$H_k \equiv \text{“}x \text{ gehört zur Klasse } C_k, k \in \{1, \dots, K\}\text{”}$$

Hypothesen und

$$f_k(x) = f(x|x \text{ gehört zu } C_k, k \in \{1, \dots, K\})$$

Dichtefunktionen.

Sei zunächst $K = 2$. Damit ergibt sich genau dann ein Fehler, wenn H_2 akzeptiert wird, obwohl H_1 gilt und umgekehrt. Die bedingten Wahrscheinlichkeiten beider Fehler ergeben sich zu

$$\begin{aligned} \alpha &= \int_{\Omega_2} f_1(x) dx && \text{und} \\ \beta &= \int_{\Omega_1} f_2(x) dx \end{aligned}$$

Die Fehlerwahrscheinlichkeit ρ hängt von der a-priori-Wahrscheinlichkeit p_i , $i \in \{1, 2\}$, und von α und β ab:

$$\begin{aligned} \rho &= \alpha p_1 + \beta p_2 \\ &= p_1 \int_{\Omega_2} f_1(x) dx + p_2 \int_{\Omega_1} f_2(x) dx \end{aligned}$$

Die Minimierung dieses Fehlers liefert eine optimale Einteilung in disjunkte Entscheidungsräume. Sei

$$\Lambda(x) = \frac{f_2(x)}{f_1(x)}$$

das Wahrscheinlichkeitsverhältnis, welches bei der Verwendung des Bayes-Kriteriums mit dem Grenzwert

$$A = \frac{p_1}{p_2}$$

verglichen wird (siehe [Gr69]). Dann wird der Fehler ρ genau dann minimal, wenn Ω_1 aus allen x mit $\Lambda(x) \leq A$ und Ω_2 aus allen x mit $\Lambda(x) > A$ besteht [YoCa74]. Bei beliebigem K folgt daraus für die Fehlerwahrscheinlichkeit:

$$\rho = \sum_{k=1}^K \int_{\Omega_k} \left[\sum_{l \neq k} p_l \cdot f_l(x) \right] dx$$

Hier ergibt sich die Minimierung des Fehlers durch folgende Einteilung des Musterraums Ω_x :

x liegt genau dann in Ω_j , wenn für alle k ungleich j gilt: $p_j \cdot f_j(x) \geq p_k \cdot f_k(x)$ [YoCa74]. Übertragen auf die vorliegende Problemstellung der Rührerauswahl wird also das Maximum der Produkte aus der a-priori-Wahrscheinlichkeit und dem Dichtefunktionswert jeder Klasse bestimmt. Damit wird eine Zuordnung zu der entsprechenden Klasse möglich.

4.2 Anwendungsergebnisse

A. Direkte Klassifikation

Bei der Umsetzung des statistischen Ansatzes am Rechner wurde zuerst der auf Seite 12 beschriebene direkte Klassifikationszusammenhang formuliert, der die Eingabegrößen *Rühraufgabe* und *Viskosität* unmittelbar auf ein bestimmtes *Rührorgan* abbildet.

\Rightarrow Merkmalvektor $x = (\text{Viskosität}, \text{Rühraufgabe})$

Menge der Klassen $C_k = \text{Menge der Rührorgane}$

Unübliche Rühraufgaben: Dabei stellt sich zunächst heraus, daß bei den möglichen 32 Rühraufgaben, die sich aus der Kombination der Grundrühraufgaben ergeben, allein 14 tatsächlich in der Fallsammlung auftreten. Diese Besonderheit ist dadurch zu erklären, daß bestimmte Kombinationen unübliche Aufgabenstellungen verkörpern. Für die Anwendung bedeutet das, daß bei Auftreten solcher Kombinationen als Eingabe die Werte der Dichtefunktionen Null sind und daher keine Aussagen gemacht werden können.

Teilung der Dichtefunktion: Es zeigt sich, daß die Häufigkeiten in den Funktionen, die aus den Aufgabenstellungen der Datensammlung gebildet werden, nur gering ausgeprägt sind. Abbildung 4.1 verdeutlicht dies am Beispiel des Propellerrührers. Obwohl dieser Rührertyp insgesamt am häufigsten in der Fallsammlung eingesetzt wird, entstehen über dem Merkmalvektor (*Aufgabe*, *Viskosität*) nur sehr geringe Häufigkeitswerte. Für alle anderen Rührerklassen entstehen noch wesentlich mehr Sektionen mit einer Häufigkeit von Null.

Abb.4.1: Häufigkeitsverteilung der Klasse *Propellerrührer*

Die Urache liegt in der Größe des Merkmalraums, der mit 160 bzw. 70 Elementen⁷ gegenüber 10 Klassen und einer Sammlung von 188 Fällen zu groß ist.

Lösung: Zur Verbesserung des Ansatzes wird die Dichtefunktion nach den Komponenten des Merkmalvektors x aufgesplittet, so daß für alle Klassen C_k jeweils zwei eindimensionale Funktionen $f_{k,vis}$ und $f_{k,auf}$ entstehen:

$$f_{k,vis}(x) = f(x|x \in \text{Viskosität}, x \text{ gehört zu } C_k, k \in \{1, \dots, K\})$$

$$f_{k,auf}(x) = f(x|x \in \text{Rühraufgabe}, x \text{ gehört zu } C_k, k \in \{1, \dots, K\})$$

Diese Aufteilung führt zu stärkeren Häufigkeitsverteilungen, die in den Abbildungen 4.2 und 4.3 dargestellt sind. Bei der Klassifikation werden diese zwei Dichtefunktionswerte mit dem Produkt- oder Maximumoperator verrechnet.

⁷Bei Zugrundelegung von 32 Aufgaben- und 5 Viskositätsintervallen ergeben sich 160 Elemente, während 70 Elemente unter Berücksichtigung der vorhandenen 14 Aufgaben- und 5 Viskositätsbereiche auftreten.

Abb.4.2: Häufigkeitsverteilung der Komponente *Aufgabe* zur Klasse *Propellerrührer*

Abb.4.3: Häufigkeitsverteilung der Komponente *Viskosität* zur Klasse *Propellerrührer*

Test der reduzierten Beispielmenge: Um eine Einschätzung in Bezug auf die Klassifikationsgenauigkeit zu erhalten, wurde die in Kapitel 3.2.2 angeführte Datenmenge getestet. Der statistische Ansatz lieferte dabei in 75% der Fälle eine übereinstimmende Zuordnung. Es ist zu berücksichtigen, daß mit der Diskretisierung der Merkmalkomponente Viskosität ein Informationsverlust verbunden ist.

Test der vollständigen Fallsammlung: Legt man die gesamte Datensammlung für einen Testlauf zugrunde, so ergibt sich eine Fehlklassifikation von 48%. Die Ursache hierfür liegt in der Mehrdeutigkeit der Auslegungsbeispiele, so daß häufig Abweichungen entstehen.

Betrachtet man Aufgabenstellungen, zu denen Beispiele mit verschiedenen Rührerzuordnungen existieren, so fällt auf, daß der statistische Ansatz denjenigen Rührer selektiert, der am häufigsten zugeordnet worden ist.

Beispiel: Zur Homogenisierung eines sehr niederviskosen Produktes wird der Propellerührer klassifiziert. Innerhalb der Datensammlung wird dieser Rührertyp in 19 von 28 Fällen gewählt, wogegen die sieben anderen Rührer ein- bis sechsmal ausgesucht werden.

Hier wird deutlich, wie das Verfahren eine inkonsistente Datenmenge behandelt: Es wird das am häufigsten auftretende Objekt klassifiziert, wobei alle anderen unberücksichtigt bleiben.

Einfluß der a-priori-Wahrscheinlichkeit: Betrachtet man die einzelnen Klassifikationen, so fällt auf, daß von den möglichen zehn Rührertypen zwei (Trapez- und Alphanührer) in keinem der Fälle klassifiziert werden und fünf Typen nur bei ein bzw. zwei festen Aufgabenstellungen, obwohl diese Rührer real vielseitiger eingesetzt werden. Alle übrigen Zuordnungen gelten den beiden Rührern, die die am häufigsten in der Datensammlung auftreten, Sigma- und Propellerrührer.

Begründung: Ursache hierfür ist das starke Gewicht der a-priori-Wahrscheinlichkeit p_j einer Klasse j in der Klassifikationsphase. Die a-priori-Wahrscheinlichkeit fließt neben dem Dichtefunktionswert $f_j(x)$ als Faktor in die Entscheidungsfunktion

$$x \text{ gehört zu Klasse } j \Leftrightarrow p_j \cdot f_j(x) \geq p_k \cdot f_k(x) \text{ für alle } k \text{ ungleich } j$$

ein.

Legt man die vollständige Datensammlung zugrunde, so wird der Propellerrührer, der eine a-priori-Wahrscheinlichkeit von 30% besitzt, in 54% der Fälle selektiert und der Sigmarührer mit einem tatsächlichen Auftreten von 24% wird in 28% der Fälle klassifiziert.

Eindeutige Klassifikation: Nachteilig erweist sich ferner die eindeutige Festlegung auf eine Rührerklasse, die keine Alternativlösungen erlaubt. Die Angabe mehrerer Lösungsvorschläge ist bei der hier betrachteten Fragestellung allerdings wünschenswert, da sich verschiedene Rührer für dieselbe Aufgabe eignen können und allein unterschiedliche Gewichtungen bezüglich der Eigenschaften auftreten. Da das hier entwickelte System einen Ingenieur bei der Rührwerksauslegung unterstützen soll, ist die Angabe von Alternativlösungen unerlässlich.

Vermeidung des Problems: Versucht man mit dem statistischen Ansatz mehrere Lösungen zu erhalten, so scheint statt der Wahl des Maximums aller Produkte aus Dichtefunktion und a-priori-Wahrscheinlichkeit die Angabe von zwei oder mehr Objekten, die hier die größten Werte aufweisen, sinnvoll.

Angewandt auf die vollständige Datensammlung ergibt die Umsetzung dieses modifizierten Ansatzes, daß ein Drittel der sonst fehlerhaft klassifizierten Rührer mit der zweitbesten Lösung übereinstimmen. Es handelt sich hier vor allem um Trapez- und Schrägblattrührer, deren a-priori-Wahrscheinlichkeiten nach Sigma- und Propellerrührer die zweithöchsten sind, wodurch wiederum das Gewicht der a-priori-Wahrscheinlichkeit deutlich wird.

Fehlende Erklärungskomponente: Für ein wissensbasiertes System, das Entscheidungen unterstützen soll, ist eine Begründung des vorgeschlagenen Ergebnisses angebracht, um den Lösungsweg nachvollziehbar zu machen. Dadurch bietet sich dem Anwender die Möglichkeit, das Ergebnis zu übernehmen oder es entsprechend der Erklärung zu modifizieren.

Der statistische Ansatz erlaubt keinen derartigen Nachweis, da die Auswahlkriterien allein auf Dichtefunktionen und Wahrscheinlichkeiten beruhen. Die Anführung dieser Werte gibt wenig Aufschluß über die Verlässlichkeit der Lösung.

B. Klassifikation über Rührermerkmale

Zur Verbesserung des Ansatzes bietet sich die auf Seite 12 beschriebene, erweiterte Betrachtungsweise an, die zunächst die Rührereigenschaften untersucht, um danach eine Auswahl zu treffen.

Bemerkungen zur Vorgehensweise:

- Klassifikation der Rührermerkmale
Zur tieferen Modellierung auf der Basis des statistischen Ansatzes müssen zunächst die Rührereigenschaften bestimmt werden. Dies kann anhand der Klassifikation jedes Merkmals geschehen analog zur oben dargestellten, direkten Klassifikation:
Klassifikation aller Rührermerkmale m :
 \Rightarrow Merkmalsvektor $x = (\text{Viskosität}, \text{Rühraufgabe})$
Menge der Klassen $C_k = \text{Menge aller Ausprägungen von } m$
- Parameter mit stetigem Wertebereich
Die vom Experten gelieferten, charakteristischen Eigenschaften eines Rührers besitzen zum Teil einen stetigen Wertebereich, so daß zur Klassifikation eine Einteilung dieses Bereichs in Klassen notwendig ist. Konkret handelt es sich um die Parameter *Drehzahl* und *Durchmesser Verhältnis*, deren Partitionierung in fünf Klassen vom Experten vorgegeben wurde.
Die übrigen relevanten Rührermerkmale besitzen einen diskreten Wertebereich.
- Abbildung der Merkmale auf den Rührertyp
Die Bestimmung eines Rührorgans kann nun anhand einer vom Experten gelieferten *Ähnlichkeitsfunktion* erfolgen. Bei dieser Funktion handelt es sich um eine Li-

nearkombination von Übereinstimmungswerten von ermittelten und tatsächlichen Rührermerkmalen. Dadurch resultiert ein Ähnlichkeitsgrad für jeden vorhandenen Rührer, der angibt, inwieweit dieser Rührer den berechneten Eigenschaften entspricht. Dieser Ähnlichkeitsgrad ermöglicht die Wahl eines passenden Rührers.

Eine nähere Beschreibung der Ähnlichkeitsfunktion wird in Kapitel 6.4.3 *Verknüpfung der Rührcharakteristika* angegeben.

Resultate:

- Die Testreihe ergab wie bei der direkten Formulierung eine Trefferquote von 75%.
- Auffällig ist wiederum die bevorzugte Klassifizierung von Merkmalwerten mit hoher a-priori-Wahrscheinlichkeit. Beispielweise wird zur Homogenisierung eines hochviskosen Produkts mit Wärmetausch ein Rührer mit Gegenstromprinzip gewählt, obwohl beide Dichtefunktionen⁸ eine tangentiale Förderrichtung favorisieren. Ausschlaggebend ist jedoch die a-priori-Wahrscheinlichkeit, die für Gegenstromförderung einen höheren Wert aufweist.

Insgesamt wird dadurch abermals die Wahl von häufig auftretenden Rührorganen begünstigt.

- Analog zur obigen Darstellungsweise bilden die erhaltenen Rührereigenschaften eindeutige Zuordnungen zu einer Klasse, ohne Alternativmöglichkeiten zu offerieren oder eine Gewichtung der Klassenzugehörigkeit anzugeben, die die Wahl eines weiteren Ergebnisses ermöglichen könnte.
- Vorteilhaft gegenüber der Beschreibung des direkten Zusammenhangs erweist sich die sich Ähnlichkeitsfunktion, die die klassifizierten Rührermerkmale mit den definierten verbindet und damit für jedes Rührorgan einen Ähnlichkeitsgrad liefert. Dadurch wird die Angabe von Alternativlösungen möglich, da verschiedene Rührer den berechneten Eigenschaften völlig oder teilweise entsprechen. Somit stellt diese Formulierung eine Verbesserung der einfachen Betrachtungsweise dar.
- Bedingt durch die Ähnlichkeitsfunktion ergibt sich im Vergleich zu der direkten Klassifikation eine gleichmäßigere Konzentrationsverteilung der klassifizierten Rührer bei den Testläufen: Jeder Rührer wird bei mindestens einer Aufgabenstellung ausgewählt, wobei sieben Rührertypen vergleichbare Häufigkeiten aufweisen.
- Ebenso bietet dieses Verfahren ansatzweise eine Erklärungsmöglichkeit für den Benutzer, da die ermittelten und definierten Rührparameter des selektierten Rührers angeführt werden können.

⁸siehe 4.2 *Teilung der Dichtefunktion*.

5 Unscharfe Logik - Der Fuzzy Ansatz

Um die negativen Aspekte des statistischen Ansatzes, wie keine Alternativlösungen, fehlende Transparenz und Begünstigung häufig auftretender Rührer, zu beseitigen, wird ein andersartiger Ansatz zur Problemlösung untersucht.

Die Aufdeckung des Lösungsweges zur Benutzerunterstützung wird durch eine regelbasierte Wissensrepräsentation ermöglicht. Die dabei zu verarbeitenden Regeln sollten aus der gegebenen Fallsammlung generiert oder vom Experten vorgegeben werden. Es ist sinnvoll, bei der Regelformulierung eine Darstellung zu wählen, die der menschlichen Ausdrucksweise nahekommt. Dadurch wird eine direkte Umsetzung der Erfahrungswerte eines Ingenieurs in das Verarbeitungsmodell gewährleistet. Hierzu bietet der Ansatz der unscharfen Logik eine passende Grundlage. Bei der Wissensverarbeitung kommen unterschiedliche Methoden in Frage, die die zur Verfügung stehenden Informationen miteinbeziehen (vergleiche [Za84], [vA191] und [Tr90]).

5.1 Beschreibung

Im Gegensatz zur zweiwertigen Logik, die nur Aussagen zuläßt, die angeben, ob eine Variable einen bestimmten Wert besitzt oder nicht, besteht der Grundgedanke der Fuzzy Logik darin, daß bei jeder Aussage der Grad der Zugehörigkeit der Variablen zu diesem Wert berücksichtigt wird. Der Grad kann als ein Maß für die Glaubwürdigkeit der Aussage gesehen werden und liegt im Bereich von 0 und 1. Diese Betrachtungsweise unterstützt die Formulierung einer eher der menschlichen Ausdrucksweise angepaßten Darstellung von Aussagen, da bestimmte Werte in Grenzbereiche zweier Werteklassen fallen und keine klare Abgrenzung möglich ist.

Beispiel: Die Idee kann anhand des Parameters Körpergröße veranschaulicht werden.

Will man zwischen großen und kleinen Körpergrößen differenzieren, so können Aussagen zu den Größen 210 cm und 140 cm eindeutig getroffen werden. Es stellt sich aber die Frage, ob eine Person der Größe 175 cm als groß oder klein einzustufen ist. Legt man eine Grenze bei 170 cm fest, von der an Personen in den Bereich groß eingeordnet werden, so würde eine 169 cm große Person klein eingestuft, was nicht sinnvoll ist.

Zur Vermeidung dieses Problems ist die Verwendung der unscharfen Logik hilfreich, die graduelle Bereichsübergänge durch Angabe von Zugehörigkeitswerten ermöglicht. Im vorliegenden Beispiel könnte demnach die Körpergröße 170 cm als groß mit Zugehörigkeitswert 0.5 und als klein mit Grad 0.4 eingestuft werden.

5.1.1 Linguistische Variablen und Zugehörigkeitsfunktionen

Zweiwertige Variablen der klassischen Logik finden ihr Pendant in der Fuzzy Logic in *linguistischen Variablen*, deren Wertebereich eine Menge von Termen bildet. Die linguistische Variable repräsentiert einen Parameter mit stetigem Wertebereich, wobei ein Zahlenwert des Parameters auf Variablensterme mit Zugehörigkeitsgrad abgebildet wird unter Verwendung einer vorgegebenen *Zugehörigkeitsfunktion*.

figure=zuehfkt.ps

Abb.5.1: Zugehörigkeitsfunktion der Variablen *Durchmesser Verhältnis*

In Abbildung 5.1 ist die Zugehörigkeitsfunktion der Variablen *Durchmesser Verhältnis* dargestellt. Die Transformation des Parameterwertes *Durchmesser Verhältnis* = 0.6 führt beispielsweise zu den unscharfen Aussagen (*Durchmesser Verhältnis* = “mittel”; 0.6) und (*Durchmesser Verhältnis* = “groß”; 0.4).

Der Zugehörigkeitsgrad ist nicht als Wahrscheinlichkeit zu deuten, da der Grad die Unschärfe einer Aussage widerspiegelt und keine Verallgemeinerung einer Häufung von Ereignissen angibt.

Bei der Frage der Rührerauswahl erzeugen die Parameter *Viskosität*, *Drehzahl* und *Durchmesser Verhältnis* linguistische Variablen. Die Zugehörigkeitsfunktionen werden als Dreieck- und Trapezlinienzüge dargestellt, die vom Experten vorgegeben sind, womit sein Verständnis von Bereichsübergängen abgebildet wird.

5.1.2 Fuzzy Regeln und Operatoren

Bei der Formulierung von Regeln, in denen linguistische Variablen auftreten, muß die Verknüpfbarkeit der Ausdrücke analog zur zweiwertigen Logik bestimmt werden.

In den nachstehenden Definitionen sei $\mu_M(x)$ der Zugehörigkeitsgrad eines Variablenwertes x einer linguistischen Variablen M .

- Operatoren einer UND-Verknüpfung:

1. $\mu_M(A \wedge B) = \min\{\mu_M(A), \mu_M(B)\}$ Minimum-Operator
2. $\mu_M(A \wedge B) = \mu_M(A) * \mu_M(B)$ Produkt-Operator
3. $\mu_M(A \wedge B) = \max\{0, \mu_M(A) + \mu_M(B) - 1\}$

- Operatoren einer ODER-Verknüpfung:

1. $\mu_M(A \vee B) = \max\{\mu_M(A), \mu_M(B)\}$ Maximum-Operator
2. $\mu_M(A \vee B) = \mu_M(A) + \mu_M(B) - \mu_M(A) * \mu_M(B)$
3. $\mu_M(A \vee B) = \min\{1, \mu_M(A) + \mu_M(B)\}$

- NEGATIONS-Operator:

$$\mu_M(\neg A) = 1 - \mu_M(A)$$

Daß diese Operatoren sinnvoll sind, zeigt die Einsetzung der Wahrheitswerte 0 und 1 in die Relationen, weil dadurch eine exakte Übereinstimmung mit den Verknüpfungen der klassischen Logik erzielt wird.

5.1.3 Defuzzifizierung

Die Invertierung der Abbildung eines Parameterwertes auf einen linguistischen Term mit Zugehörigkeitswert wird als *Defuzzifizierung* bezeichnet. Die am häufigsten angewendete Technik der Rückübersetzung stellt die *Flächenschwerpunktmethod*e dar:

Gegeben sei eine Menge von Term-Zugehörigkeitswert-Verbindungen einer linguistischen Variablen M . Ausgehend von diesen Informationen soll der Wert innerhalb des stetigen Wertebereichs des Parameters M bestimmt werden.

Dazu wird für jeden Term die Fläche unterhalb des zugehörigen Linienzuges betrachtet, die von der Horizontalen in der Höhe des entsprechenden Zugehörigkeitswertes bis zur x-Achse reicht. Die so entstehenden Flächen werden zu einer Gesamtfläche zusammengefaßt. Die Defuzzifizierung erfolgt nun mit Hilfe des Flächenschwerpunktes (x_g, y_g) der so gebildeten Gesamtfläche:

Der gesuchte Wert des Parameters M ist gleich der x-Koordinate x_g dieses Punktes.

Der Flächenschwerpunkt einer aus n Teilflächen zusammengesetzten Fläche wird mit Hilfe der Flächeninhalte f_i und Flächenschwerpunkte (x_{g_i}, y_{g_i}) aller Teilflächen i wie folgt berechnet (vergleiche [BrSe87]):

$$x_g = \frac{\sum_{i=1}^n x_{g_i} \cdot f_i}{\sum_{i=1}^n f_i}$$

$$y_g = \frac{\sum_{i=1}^n y_{g_i} \cdot f_i}{\sum_{i=1}^n f_i}$$

figure=defuzzy.ps

Abb.5.2: Defuzzifizierungsfläche der Variablen *Durchmesser Verhältnis*

Abbildung 5.2 veranschaulicht die relevante Fläche für die Term-Zugehörigkeitswert-Verbindungen

- (Durchmesser Verhältnis = "sehr klein" ; 0.4)
- (Durchmesser Verhältnis = "klein" ; 0.7)
- (Durchmesser Verhältnis = "mittel" ; 0.5)
- (Durchmesser Verhältnis = "groß" ; 0.3)
- (Durchmesser Verhältnis = "sehr groß" ; 0)

und den daraus resultierenden Flächenschwerpunkt (0.43 0.25).

Ein Durchmesser Verhältnis von 0.43 entspricht wiederum den Aussagen:

- (Durchmesser Verhältnis = "klein" ; 0.29)
- (Durchmesser Verhältnis = "mittel" ; 0.71).

5.2 Allgemeine Vorgehensweise

Wissenserwerb und Verarbeitung auf der Basis von unscharfer Logik kann auf unterschiedliche Weise gestaltet werden. Bei der vorliegenden Fragestellung der Rührerauswahl wurden diverse Ansätze am Rechner umgesetzt und untersucht.

Allgemein setzt sich der Lösungsweg aus folgenden Schritten zusammen:

1. Aus der zur Verfügung stehenden Fallsammlung wird Wissen in Form von Regeln akquiriert, wobei die *erzeugte Regelmenge* als Grundlage für die nachfolgende Verarbeitung dient.

In Kapitel 6 werden zwei Ansätze zur Regelgenerierung vorgestellt. Außerdem werden verschiedene Regelstrukturen verglichen, wobei sowohl die direkte Klassifikation als auch die tiefere Modellierung die Grundlage bildet.

2. Um eine aussagekräftige Regelmenge zu erhalten, die möglichst kompakt ist und in der keine ungeeigneten Regeln beinhaltet sind, wird mit Hilfe der Fallsammlung und der Zugehörigkeitswerte eine Einschätzung für die *Gültigkeit jeder Regel* ermittelt. Auf diese Weise kann eine Selektion derjenigen Regeln erfolgen, die hohe Gültigkeitswerte aufweisen und insgesamt eine gehaltvolle Grundlage für die Problemlösung darstellen.

Es werden in Kapitel 7 zwei Verfahren vorgestellt, die zu einer Bewertung der Regeln führen. Die resultierenden Gültigkeitswerte (Konfidenzen) liefern eine Einschätzung in Bezug auf die Aussagekraft der Regeln und fließen in die nachfolgende Regelverarbeitung mit ein.

3. Die *Verarbeitung* basierend auf der so generierten Regelmenge liefert für eine gegebene Problemstellung geeignete Kandidaten aus der Menge der Rührorgane.

Die Ausführungen in Kapitel 8 beschreiben die Abarbeitung der Menge bewerteter Regeln. Die Vorgehensweise setzt sich aus folgenden Schritten zusammen:

- (a) Für eine gegebene Aufgabenstellung werden alle Regeln hinsichtlich der Übereinstimmung mit der Vorbedingung überprüft.
(\Rightarrow Gültigkeitswert der Prämisse)
- (b) Zur Bestimmung der Gültigkeit der Konklusion wird der Gültigkeitswert der Prämisse mit der entsprechenden Regelbewertung verrechnet.

- (c) Die Regelpropagierung liefert auf diese Weise unterschiedlich gewichtete Rührorgane als Resultat, so daß eine geeignete Auswahl möglich wird.

Für jeden Verarbeitungsteil bieten sich verschiedene Lösungsmöglichkeiten an, die in den Kapiteln 6 bis 8 beschrieben und diskutiert werden. Dabei wird stets ein Vergleich der Methoden hinsichtlich der durch Testreihen erzielten Resultate angeführt. Für die Testläufe werden zum Teil Verarbeitungsschritte vorausgesetzt, die erst später ausführlich behandelt werden.

Die hier angegebene Auflistung der drei Komponenten des Lösungsweges soll einen Überblick über den Gesamtablauf geben, damit die Ergebnisse der Testreihen an jeder Stelle nachvollziehbar sind.

6 Generierung einer Regelmenge

Bei einem auf Fuzzy Logic basierenden Ansatz zur Lösung des Rührproblems stellt sich die Frage, wie eine aussagekräftige Regelmenge gewonnen werden kann. Hierzu bietet sich die Generierung von Regeln basierend auf der zur Verfügung stehenden Fallsammlung an, in der implizit rührtechnische Zusammenhänge dargestellt sind.

6.1 Struktur der Regeln des Rührproblems

Bei der Untersuchung einer geeigneten Struktur der Regeln hinsichtlich der Aufgabenstellung, stellte sich folgendes heraus:

Setzt man eine Vorwärtsinferenz voraus, so sollten die Vorbedingungen der Regeln die Parameter der Aufgabenstellung repräsentieren: Viskosität des Rührguts und Rühraufgabe. Bei dem Attribut Rühraufgabe handelt es sich in der Regel um eine Kombination der fünf Grundaufgaben, so daß beispielsweise eine Regelstruktur der Form

“IF (Aufgabe=Homogen.) AND (Aufgabe=Suspend.) AND (...) THEN ...”

geeignet scheint. Allerdings werden sich damit sehr spezielle Regelausdrücke ergeben, die wenig Allgemeingültigkeit besitzen, da nur eine kleine Zahl von Fallbeispielen so formulierten Vorbedingungen genügt. Insgesamt wird dadurch das Entstehen einer sehr großen Regelmenge begünstigt. Ziel ist jedoch die Gewinnung einer kompakten Regelmenge, die eine hohe Aussagekraft besitzt.

Lösung: Splitten der Regelprämisse

Es fällt auf, daß Schlußfolgerungen, die aus Kombinationen von Rühraufgaben abgeleitet werden, für jedes Element der Aufgabenkombination Gültigkeit besitzen. Damit kann der Zusammenhang von Rühraufgabe und Parameter aufgeteilt werden in Zusammenhänge, die jede enthaltene Grundrühraufgabe mit dem Parameter verbinden.

Für die Regelformulierung ergibt sich daher zusammen mit der Eingabegröße Zähigkeit folgende Struktur:

“IF (Aufgabe = a) AND (Viskosität = v) THEN ...”

Obwohl bei dieser Regelformulierung Informationen in Bezug auf die Verwendung der Rührorgane bei kombinierten Aufgaben verloren gehen, kann die Vorteilhaftigkeit dieses Aufbaus in Kapitel 6.3 anhand von Testergebnissen nachgewiesen werden.

6.2 Methoden zur Regelerzeugung aus einer Datensammlung

6.2.1 Direkte Ableitung von Zusammenhängen

Die naheliegendste Möglichkeit, Regeln aus einer Fallsammlung zu generieren, besteht in der direkten Ausformulierung der durch jeden Fall dargelegten Zusammenhänge. Die Vorgehensweise kann dem nachfolgenden Algorithmus entnommen werden. Man erzielt auf diese Weise eine Verallgemeinerung der Beispiele, die eine Übertragung auf neue Fälle erlaubt, auch wenn die Fallsammlung unvollständig und teilweise inkonsistent ist. Die Mängel können durch die in Kapitel 7 dargelegten Verfahren zur Regelbewertung und durch entsprechende Heuristiken bei den in Kapitel 8 beschriebenen Auswahltechniken aufgehoben bzw. vermindert werden.

Algorithmus Regelgenerierung

Für alle Fallbeispiele i

Sei p ein Parameterwert, der für i gültig ist

Sei Aufgabe_i die Menge aller Grundrühraufgaben aus i

Sei Viskosität_i die Menge aller Viskositätsbereiche aus i ⁹

Für alle a aus Aufgabe_i

Für alle v aus Viskosität_i

Trage die Regel IF $\text{Aufgabe}=a$ AND $\text{Viskosität}=v$ THEN $\text{Parameter}=p$
in die Regelmenge ein, falls sie noch nicht enthalten ist

Beispiel: Sei eine direkte Regelformulierung nach Kapitel 3.3 Unterpunkt 3.3 vorausgesetzt, wobei das Fallbeispiel bei einer kombinierten Homogenisier- und Suspenderaufgabe eines niederviskosen Produktes einen Schrägblattrührer einsetzt. Dabei erzeugt der Algorithmus folgende Regeln:

“IF $\text{Aufgabe}=\text{Homogenis.}$ AND $\text{Viskosität}=\text{“niedrig”}$ THEN $\text{Rührer}=\text{Schrägblatt}$ ”

“IF $\text{Aufgabe}=\text{Suspend.}$ AND $\text{Viskosität}=\text{“niedrig”}$ THEN $\text{Rührer}=\text{Schrägblatt}$ ”

⁹Ein fester Parameterwert kann in mehrere linguistische Terme übertragen werden.

Auf diese Weise wird jeder in den Fallbeispielen dargestellte Zusammenhang als allgemeine Regel formuliert, wobei stets die gleiche Struktur zugrunde gelegt wird, die eine zweistellige, mit AND verknüpfte Prämisse und eine einstellige Konklusion aufweist.

Anmerkung: Ist eine Regelformulierung gesucht, die in der Prämisse kombinierte Aufgabenstellungen voraussetzt, so ist im Algorithmus die Schleife “Für alle a aus Aufgabe _{i} ” zu entfernen und alle Elemente aus Aufgabe _{i} mit AND verknüpft in den Prämissen aufzunehmen.

6.2.2 Der ID3 Algorithmus

Eine weitere Methode zur Regelerzeugung stellt der klassische *ID3 Algorithmus von Quinlan* dar. Dieses Verfahren generiert eine Regelmenge aus klassifizierten Fallbeispielen, die durch verschiedene, relevante Attribute gekennzeichnet sind. Dabei wird ein Entscheidungsbaum aufgebaut derart, daß Knoten die Attribute, Kanten die Attributausprägungen und Blätter Entscheidungsklassen darstellen. Die Attribute und Ausprägungen entlang eines Weges, der von der Wurzel bis zu einem Blatt reicht, können dann in Regelform transformiert werden. Ziel des Verfahrens ist der Aufbau eines Entscheidungsbaums mit möglichst kleiner externer Pfadlänge, da dies zu allgemeineren Regeln mit kurzer Prämisse führt (siehe [KB91]).

Im Vergleich zum Algorithmus *Regelgenerierung* sprechen folgende Gründe gegen die Verwendung des ID3 Algorithmus bei der hier betrachteten Problemstellung:

1. Die Voraussetzung einer konsistenten Objektmenge, die gleiche Attributkombinationen verschiedener Beispiele in dieselbe Klasse einordnet, ist nicht gegeben, da in der vorliegenden Fallsammlung bei gleichen Aufgabenstellungen unterschiedliche Rührorgane selektiert werden. Damit ist die Konstruktion eines Entscheidungsbaumes mit homogenen Blättern nicht gewährleistet.
2. Zum Aufbau eines Entscheidungsbaumes wird als erforderliche Heuristik für die Attributauswahl in der Regel der Informationsgehalt aller Attribute untersucht. Diese Größe kann im vorliegenden Fall nicht bestimmt werden, wenn eine Aufspaltung der Rühraufgabe in Grundrühraufgaben vorausgesetzt wird, weil dadurch a-priori-Wahrscheinlichkeiten verfälscht werden. Hier ist die Summe aller Fälle, die Einzelaufgabe und Klasse verbindet, größer als die Zahl der tatsächlich auftretenden Fälle, die in diese Klasse fallen.

Somit muß entweder eine andersartige Heuristik verwendet werden oder es müssen die kombinierten Rühraufgaben in den Regeln betrachtet werden. Die zweite Möglichkeit führt zu einer starken Verzweigung des Entscheidungsbaumes und damit zu sehr speziellen Regeln, die der Forderung möglichst großer Allgemeingültigkeit widersprechen.

3. Bei der Frage nach den Vorteilen des ID3 Ansatzes gegenüber der direkten Formulierung im Algorithmus **Regelgenerierung** stellt sich heraus, daß die beabsichtigte Verkürzung der Regeln und Verringerung der Regelmenge nur unwesentlich sein kann. Die Verkürzung einer Prämisse kann höchstens um einen einzigen Ausdruck erfolgen, da von vornherein maximal zwei Attribute verknüpft werden müssen. Außerdem bietet sich die Möglichkeit einer einstelligen Prämisse in keinem der in Frage kommenden Fälle¹⁰.
4. Da nicht alle Ausprägungen der Attribute nominal skaliert sind und es sich teilweise um stetige Wertebereiche handelt, muß eine Aufteilung dieser Wertebereiche in Intervalle erfolgen, um eine endliche Zahl von Attributwerten zu erzielen. Es bietet sich eine der statistischen Einteilung ähnliche Bereichsaufteilung an, die beispielsweise den Viskositätsbereich in fünf Intervalle gliedert. Hierbei geht jedoch die Grundidee der Fuzzy Logic verloren, die eine scharfe Abgrenzung zweier Bereiche vermeiden will.

Legt man daher eine Zuordnung von linguistischen Termen gemäß der Zugehörigkeitsfunktionen zugrunde, so bleibt der Aspekt der allmählichen Bereichsübergänge erhalten. Der ID3 Algorithmus muß jedoch modifiziert werden, um einerseits Zugehörigkeitswerte einbinden zu können und andererseits mit Mehrfachwerten eines Attributs umgehen zu können. In [WeZi91] wird ein Ansatz vorgestellt, der den ID3 Algorithmus hinsichtlich der Verwendung von unscharf formuliertem Wissen untersucht. Zugehörigkeitswerte werden dort entsprechend der Verknüpfungsfunktionen verrechnet, die bereits in 5.1.2 vorgestellt worden sind. Die Frage der Behandlung von Mehrfachwerten eines Attributs wird bezüglich der Bestimmung des Informationsgehaltes wie folgt gelöst:

Eine Berechnung der a priori Wahrscheinlichkeit ist nicht möglich, wie bereits am Beispiel der Rühraufgaben verdeutlicht wurde. Stattdessen wird in [WeZi91] unter Ausnutzung der Zugehörigkeitswerte die relative Kardinalität p_i betrachtet:

$$p_i = \frac{\text{Summe der Zugehörigkeitswerte zur } i\text{-ten Ausprägung}}{\text{Anzahl der Beispiele}}$$

Für linguistische Variablen bietet sich somit eine geeignete Möglichkeit der Einbindung in den ID3 Algorithmus, während für die Aufspaltung der kombinierten Rühraufgaben kein vergleichbarer Ansatz erkennbar ist.

Aufgrund der aufgeführten Nachteile wurde in dieser Arbeit auf eine nähere Betrachtung des ID3 Algorithmus verzichtet.

6.3 Darstellung der Prämisse

In den obigen Ausführungen wurde bereits auf die verschiedenen Möglichkeiten der Formulierung von Regelprämissen hingewiesen:

¹⁰Es wird jedes Rührermerkmal und jeder Rührer als Schlußfolgerung zugrunde gelegt.

1. Die Verknüpfung von mehreren Grundrühraufgaben, die zusammengesetzt eine Aufgabenkombination darstellen und mit der Viskosität durch den UND-Operator verbunden sind (kombinierte Formulierung).
2. Die Betrachtung von genau einer Grundrühraufgabe, die mit einem Viskositätswert verknüpft ist (gesplittete Formulierung).

Die auf Seite 26 geäußerte Vermutung, die kombinierte Formulierung der Prämissen sei nachteilhaft gegenüber der gesplitteten Darstellung, soll im folgenden durch konkrete Ergebnisse gestützt werden.

Vergleich der Größe der Regelmengen: Bei der Untersuchung der anhand des Algorithmus *Regelgenerierung* erzeugten Regelmengen werden zwei Arten der Schlußfolgerung unterschieden, die bereits in Kapitel 3.3 vorgestellt wurden.

A. Direkte Klassifikation:

Beim direkten Schließen auf die Rührorgane ergaben sich ähnlich große Regelmengen: 101 Regeln für die gesplitteten Prämissen und 108 Regeln bei kombinierter Formulierung. Allerdings ist zu bemerken, daß die Zahl der in den Vorbedingungen enthaltenen Ausdrücke bei der aufgeteilten Version gleichbleibend zwei ist, während die kombinierte Form auch längere Verknüpfungen bis zu vier Termen zuläßt. Der Vergleich der Gesamtzahl der Rühraufgaben in den Prämissen ergibt ein Verhältnis von 190 Termen bei zusammengeführten Regeln gegenüber 101 Ausdrücken bei fester Regelform.

B. Klassifikation über Rührermerkmale:

Deutlicher zeigt sich die Vorteilhaftigkeit der gesplitteten Struktur, wenn als Regelkonklusion die Rührermerkmale verwendet werden. In diesem Fall kann ein Vergleich von 353 aufgegliederten Regeln und 484 kombinierten Regeln angestellt werden, wobei die Prämissen der Kombinationsregeln insgesamt aus 837 Grundaufgaben gebildet werden, wogegen die gesplitteten Regeln 353 Terme ergeben.

Vergleich der begründenden Fälle: Die Aussagekraft einer Regel wird gefestigt durch die Zahl der Fallbeispiele, die die Regel unterstützen, denn je mehr Fälle den in der Regel ausgedrückten Zusammenhang enthalten, desto mehr spricht für die Gültigkeit der Regel.

Bei der Betrachtung dieses Kriteriums wird analog zu der obigen Auswertung differenziert:

A. Direkte Klassifikation:

Werden als Schlußfolgerung Rührorgane gewählt, so liegen die durchschnittlichen Zahlen der erklärenden Fallbeispiele zwischen eins und acht bei den gesplitteten

Prämissen und zwischen eins und fünf bei den kombinierten, wenn die Regelmenge nach den einzelnen Rührern aufgegliedert wird. Der Gesamtmittelwert liegt bei 3,4 Beispielen bei aufgeteilter Formulierung und nur bei 2,3 bei den verknüpften Aufgaben. Diese Werte sind bei beiden Strukturen recht gering. Das Mißverhältnis zeigt sich klarer bei der Untersuchung der indirekten Abbildung.

B. Klassifikation über Rührermerkmale:

Wird eine Regelformulierung betrachtet, in der die Rührereigenschaften die Konklusion bilden, so fällt der Unterschied von durchschnittlich zwei bis fünfzehn begründenden Fallbeispielen bei gesplitteten Regeln gegenüber nur ein bis sechs Beispielen bei der verknüpften Form gravierender aus. Der Gesamtdurchschnitt liegt bei 8,1 Beispielen im Vergleich zu 3,7 Fällen. Desweiteren wird der Maximalwert von 48 unterstützenden Beispielen bei geteilter Prämissendarstellung mit einer Zahl von 31 im kombinierten Fall kaum erreicht. Dies bestätigt die Vorteilhaftigkeit der festen Regelstruktur mit jeweils einer Grundrühraufgabe.

Eignung bei der Übertragung auf neue Fälle: Neben einer geringeren Zahl von Regeln, die von mehr Fallbeispielen gestützt werden, erweist sich bei der gesplitteten Regelform der Aspekt der erweiterten Verwendbarkeit vorteilhaft. In Kapitel 4.2 wurde bereits erläutert, daß die Fallsammlung nur einen Teil aller möglichen Rühraufgaben repräsentiert, obwohl andere Aufgabenstellungen denkbar wären. Soll beim Einsatz der hier vorgestellten Ansätze eine bisher noch nicht betrachtete Fragestellung gelöst werden, so ist die aufgegliederte Regelmenge unproblematisch verwendbar.

Diese universelle Nutzbarkeit ist bei einer fallbasierten Generierung von Regeln mit kombinierter Aufgabenstellung nicht möglich, da für bestimmte Aufgaben keine anwendbaren Regeln dieser Form verfügbar sind. Beispielsweise existiert in der vorliegenden Fallsammlung keine Rührwerksauslegung, die allein *Begasen* zur Aufgabe hat. Daher erzeugt der Algorithmus *Regelgenerierung* keine Regel mit diesbezüglicher Prämisse, so daß Angaben bei einem derartigen Klassifikationsproblem nicht möglich sind, während die gleiche Fragestellung bei aufgeteilter Regelform durchaus lösbar wäre.

Testergebnisse: Legt man die auf Seite 12 angeführte Beispielsammlung für Testläufe zugrunde, so ergibt sich eine Trefferquote von 78% bei direkt formulierten Regeln mit gesplitteter Prämisse gegenüber einer 72%-igen Übereinstimmung bei der kombinierten Regelformulierung¹¹.

Auffällig ist, daß bei der kombinierten Regelmenge allein Fragestellungen mit zusammengesetzter Rühraufgabe zu Fehlern führen. Im Gegensatz zur gesplitteten Regelformulierung fließen keine Informationen von Fallbeispielen ein, die die Aufgabe nur zum Teil erfüllen. Angesichts der inkonsistenten Datenmenge ist es jedoch vorteilhaft, so viele

¹¹Die Regelmengen wurden nach der in Abschnitt 7.1 dargestellten Methode gewichtet. Die Verarbeitung erfolgte nach dem in Kapitel 8 beschriebenen Verfahren.

Fälle wie möglich zu berücksichtigen, um Entscheidungen zu festigen. Die gesplitteten Form gewährleistet den partiellen Einfluß von Beispielen, so daß hinsichtlich der zusammengesetzten Aufgabenstellungen hier deutlich bessere Ergebnisse erzielt werden (Fehlklassifikation bei einer von 22 Fragestellungen).

Die Fehlerquote von 22% ergibt sich in diesem Fall durch Beispiele, die aus nur einer Grundrühraufgabe bestehen. Die Ergebnisse basieren auf *allen* Fällen, die diese Rühraufgabe beinhalten. Daher spiegelt das Resultat die Gesamtheit dieser Beispiele wider, so daß in Bezug auf die geforderten Auslegungen abweichende Zuordnungen entstehen können. Demgegenüber liefert die kombinierte Regelformulierung bei einzelnen Grundrühraufgaben sehr gute Ergebnisse (keine Fehlklassifizierung bei 14 Fragestellungen). Ausschlaggebend ist, daß nur Beispiele, die genau diese Aufgabe erfüllen, in die Verarbeitung einfließen, und so die Nähe zu der Fallsammlung gewahrt werden kann. An dieser Stelle wird der mit der Aufteilung der Rühraufgaben verbundene Informationsverlust in den gesplitteten Regeln deutlich.

Fazit: Aufgrund der angeführten Argumente wird für die anschließenden Untersuchungen allein die gesplittete Formulierung der Prämissen zugrunde gelegt.

6.4 Modellierungstiefe

6.4.1 Beschreibung

Im folgenden wird die regelbasierte Umsetzung der beiden auf Seite 12 vorgestellten Lösungsmodelle untersucht.

- A. Die Formulierung des *direkten Zusammenhangs* der Eingabegrößen und der gesuchten Rührorgane liefert die Regelstruktur:

“IF Aufgabe = a AND Viskosität = v THEN Rührer = r”

- B. Die *erweiterte Modellierung*, die Zwischenergebnisse in Form von Rühreigenschaften betrachtet, und anhand dieser Resultate ein Rührorgan selektiert, formt zunächst Regeln der Struktur:

“IF Aufgabe = a AND Viskosität = v THEN Merkmal = m”

Bei den Merkmalen handelt es sich um diejenigen Rühreigenschaften, die eine nahezu vollständige Charakterisierung von Rührorganen ermöglichen. Diese Merkmalmenge wurde vom Experten vorgegeben und ist in Kapitel 3.1 *Expertenwissen* aufgezeigt worden.

Der zweite Schritt besteht in der Abbildung der Merkmale auf die Rührorgane. Hier kann einerseits eine weitere Regelmenge verwendet werden, die den Zusammenhang

Merkmal → *Rührer* beschreibt. Alternativ dazu kann der Gedanke aus dem statistischen Ansatz aufgegriffen werden: Die Verwendung einer Ähnlichkeitsfunktion, die aus einer Gegenüberstellung von tatsächlichen und berechneten Rührermerkmalen einen Ähnlichkeitswert für jeden Rührer bestimmt, so daß eine Klassifikation ermöglicht wird.

6.4.2 Gegenüberstellung

Vergleicht man beide Darstellungsweisen, so sprechen mehrere Gesichtspunkte für die Klassifikation über Rührermerkmale:

Grundgedanke: Die Idee der tieferen Modellierung entsteht aus der Nachbildung des Vorgehens eines Ingenieurs bei der Lösung des Rührproblems.

Der Experte betrachtet die in der Aufgabe geforderten Größen Rühraufgabe und Viskosität nicht nur bezüglich einer direkten Verbindung zu den Rührorganen, sondern leitet aus der Aufgabenstellung Eigenschaften ab, die ein passendes Rührorgan aufweisen sollte.

Beispiel: Die Wahl eines kantig geformten Rührers ist günstig für Begasungsvorgänge. Für Suspendieraufgaben eignen sich oft Rührer mit axialer Förderrichtung. Die Vermischung hochviskoser Produkte kann von Rührern mit großem Durchmesser-verhältnis ausgeführt werden.

Unter Berücksichtigung derartiger Heuristiken wählt der Ingenieur einen Rührer aus, der die größte Übereinstimmung mit den so ermittelten Eigenschaften aufweist. Demzufolge bietet sich beim regelbasierten Ansatz eine analoge Regelformulierung an.

Testergebnisse: Im vorangegangenen Abschnitt wurde bereits aufgezeigt, daß für die direkte Regelformulierung eine Trefferquote von 78% gegenüber einer Fehlklassifizierung von 22% ergibt. Verwendet man die Regelmenge, die Schlußfolgerungen zu den Rührermerkmalen zieht, zusammen mit der Ähnlichkeitsfunktion, so werden 81% der Fälle richtig klassifiziert. Neben dieser Verbesserung des Testergebnisses erweist sich außerdem die korrekte Klassifikation seltener Rührer, wie beispielsweise Alpha- und Ankerrührer, positiv gegenüber der direkten Klassifikation.

Erklärungskomponente: Neben der Lösung des Problems der Rührerauswahl wird die Intention verfolgt, Erklärungen zur Lösungsfindung zu machen, damit der Lösungsweg für den Benutzer transparent wird. Die Begründung von Resultaten gewährleistet eine Einschätzung in Bezug auf die Akzeptanz des Ergebnisses und stellt eine wichtige Komponente bei der Umsetzung der Resultate dar. Dem Benutzer bietet sich die Möglichkeit, einen Lösungsvorschlag zu verwerfen oder zu modifizieren, falls Erklärungen unbefriedigend sind.

Der erweiterte Ansatz stellt hierzu günstige Voraussetzungen bereit, wodurch folgende Informationen geliefert werden können:

1. Angabe aller anwendbaren Regeln, die auf Rührermerkmale schließen
2. Anführung des Grads der Übereinstimmung der Aufgabenstellung mit den Regelprämissen
3. Auflistung der so erhaltenen Rührermerkmale
4. Darstellung der Verknüpfungsmethode der Merkmale
5. Ausgabe des klassifizierten Rührers

Im Gegensatz dazu ermöglicht die direkte Formulierung allein die Angabe der Regel, die die beste Einschätzung bezüglich Glaubwürdigkeit¹² und Übereinstimmungsgrad mit der Aufgabenstellung aufweist.

Alternativlösungen: Die Zweckmäßigkeit mehrerer Lösungsvorschläge wurde bereits beim statistischen Ansatz in Kapitel 4.2 deutlich gemacht. Die tiefere Modellierung des regelbasierten Ansatzes bietet hier vorteilhaftere Möglichkeiten als die unmittelbare Darstellung.

Alternativlösungen bilden diejenigen Parameter, die nach der ermittelten Lösung die besten Werte bezüglich der Glaubwürdigkeit der zugehörigen Regeln und der Übereinstimmung der Prämisse mit der Aufgabenstellung aufzeigen.

Können bei der direkten Regelformulierung allein andere Rührorgane ausgegeben werden, so erlaubt die erweiterte Form außerdem Alternativangaben zu einzelnen Merkmalen. Damit kann der Benutzer vorgeschlagene Rührermerkmale abändern, wodurch sich eine zweite Möglichkeit ergibt, alternative Rührorgane zu finden.

Zusätzliche Informationen: Neben der Angabe eines geeigneten Rührers zu einem gegebenen Rührproblem stellt das tiefere Modell zusätzliche Informationen zur Realisierung des Rührwerks bereit. So werden Aussagen bezüglich der *Drehzahl* und des *Durchmesserhältnisses* gemacht, so daß der Ingenieur bei der Konstruktion der Gesamtanlage unterstützt wird.

Flexiblere Verarbeitung: Die Anwendung des erweiterten Ansatzes ermöglicht die Ausnutzung einer größeren Informationsmenge, wodurch der Wissensgehalt bei der Verarbeitung steigt. Beispielsweise hat der Ingenieur einen hohen Einfluß auf die Art der Verknüpfung der Merkmale. Er kann bestimmte Eigenschaften hervorheben und so spezifische Auswahlkriterien schaffen (siehe Abschnitt 6.4.3). Damit sind verbesserte Ergebnisse zu erwarten.

¹²Ansätze zur Ermittlung von Regelgewichtungen werden in Kapitel 7 vorgestellt.

Einbindung neuer Rührer: Die aus der Fallsammlung gewonnen Regeln stellen Verallgemeinerungen dar, die generelle Zusammenhänge von Problemstellung und Rührermerkmal bzw. Rührorgan beschreiben. Leitet man aus einer Aufgabe zunächst eine Liste von Rührermerkmalen ab, so bietet sich folgende Erweiterung an:

Neu spezifizierte Rührer, die nicht in der Fallsammlung erscheinen, können in die Rührersammlung aufgenommen werden. Durch den Vergleich von ermittelten und echten Rührermerkmalen kann für bestimmte Aufgabenstellungen ein neu definierter Rührer selektiert werden, falls dieser den größten Ähnlichkeitsgrad aufweist. Ein Vorhandensein in der Fallsammlung ist nicht notwendig.

Es ist zu beachten, daß dieser Aspekt im Zusammenhang mit direkt formulierten Regeln ungeeignet ist. Um hier neu spezifizierte Rührorgane einzubauen, ist die Generierung von Regeln notwendig, die diese Rührer als Schlußfolgerung enthalten. Das ist nur dann möglich, wenn die Beispielsammlung entsprechende Fälle beinhaltet.

Fazit: Die tiefere Modellierung der Rührzusammenhänge erweist sich in vieler Hinsicht vorteilhaft gegenüber der direkten Darstellung. Eine leicht nachzuvollziehende, überschaubare Regelmenge und gute Resultate sprechen jedoch für die unmittelbare Regelformulierung. Daher werden bei den anschließenden Betrachtungen beide Formen untersucht.

6.4.3 Verknüpfung der Rührcharakteristika

Im folgenden werden zwei unterschiedliche Methoden vorgestellt, die eine Menge von Rührermerkmalen auf ein Rührorgan abbilden und bei der erweiterten Modellbildung eingesetzt werden.

1. Ähnlichkeitsfunktion

Idee: Bei der Abbildung der gefolgerten Rührermerkmale auf ein Rührorgan soll derjenige Rührer bestimmt werden, der die größte Übereinstimmung mit den Merkmalen aufweist. Daher bietet es sich an, für jeden Rührer zu berechnen, inwieweit dieser den geforderten Eigenschaften entspricht.

Dazu ist es notwendig, alle in Frage kommenden Rührorgane zu charakterisieren, damit ein Vergleich von tatsächlichen und ermittelten Merkmalen möglich wird. Der Experte legt also für alle in der Fallsammlung auftretenden Rührer eine Spezifikation anhand der nachstehenden Merkmale ab:

- *Grundrühraufgaben*, für die der Rührer geeignet ist
 $\subset \{ \text{Homogenisieren, Suspendieren, Wärmeübergang, Emulgieren, Begasen} \}$
- *Viskositätsbereich*, in dem der Rührer eingesetzt werden kann
 $\subset \{ \text{„sehr niedrig“, „niedrig“, „mittel“, „hoch“, „sehr hoch“} \}$

- *Drehzahlbereich*
 $\subset \{ \text{“sehr niedrig”, “niedrig”, “mittel”, “groß”, “sehr groß”} \}$
- *Durchmesser Verhältnis*
 $\subset \{ \text{“sehr niedrig”, “niedrig”, “mittel”, “groß”, “sehr groß”} \}$
- *Förderrichtung*
 $\subset \{ \text{axial, radial, tangential, gegenstrom} \}$
- *Form*
 $\subset \{ \text{sehr kantig, kantig, mittel, strömungsgünstig, sehr strömungsgünstig} \}$
- *Stufenzahl*
 $\subset \{ \text{einstufig, mehrstufig} \}$

Beispiel: Für jede Eigenschaft m des *Scheibenrührers* wird folgender Wertebereich m_{def} definiert:

Viskositätsbereich	“sehr niedrig”, “niedrig”
Rühraufgaben	Emulgieren, Begasen, Wärmeübergang
Drehzahl	“mittel”, “groß”
Durchmesser Verhältnis	“klein”, “mittel”
Förderrichtung	radial
Form	sehr kantig
Stufenzahl	einstufig

Aus einer zu lösenden Problemstellung wird zunächst durch Verarbeitung der Regelmenge für jedes Merkmal m ein Wert m_{comp} ermittelt, der eine bestimmte Gültigkeit g_{comp} besitzt¹³. Beispielsweise ergibt sich für das Merkmal Drehzahl der Wert “groß” mit einer Gültigkeit von 0.7.

Um zu bestimmen, welches Rührorgan den propagierten Merkmalen am ähnlichsten ist, bietet sich die Verwendung einer *Ähnlichkeitsfunktion* an (siehe auch [KB92] *Fallbasierte Diagnose*). Diese Funktion ermittelt für jeden Rührer einen Ähnlichkeitswert a_r . Dazu wird ein Ähnlichkeitsvektor v_{equ} gebildet, der den Grad der Übereinstimmung g_{comp} jedes Rührermerkmals mit den geforderten Merkmalen enthält. Um die Bedeutung der einzelnen Merkmale in Bezug auf die Gesamtstruktur unterschiedlich zu gewichten, wird v_{equ} mit einem Gewichtsvektor v_{weight} multipliziert. Die Größe des Einflusses eines Merkmals wird vom Ingenieur vorgegeben. Für die

¹³Bei Einsatz der Ähnlichkeitsfunktion im *statistischen Ansatz* muß folgendes beachtet werden: Jedes Merkmal wird hier anhand von Dichtefunktionen klassifiziert. Daher wird bei diesem Ansatz als Gültigkeitswert stets 1.0 verwendet.

anschließenden Untersuchungen wurden folgende Gewichtungen zugrunde gelegt:

$$\begin{pmatrix} \text{Viskosität} \\ \text{Aufgaben} \\ \text{Drehzahl} \\ \text{Durchm.Verh.} \\ \text{Förderrichtung} \\ \text{Form} \\ \text{Stufenzahl} \end{pmatrix} = \begin{pmatrix} 1.0 \\ 0.75 \\ 0.75 \\ 0.75 \\ 0.5 \\ 1.0 \\ 0.25 \end{pmatrix}$$

Der nachstehende Algorithmus zeigt die Vorgehensweise zur Berechnung der Ähnlichkeitswerte auf:

Algorithmus Ähnlichkeitsgrad

Für alle Rührorgane r

Sei v_{equ} ein Vektor der Dimension #Merkmale,
der mit 0 initialisiert ist

Sei v_{weight} ein Vektor der Dimension #Merkmale, der an der i -ten
Position die Gewichtung des Merkmals i enthält

Für alle Merkmale m

Sei m_{def} die Menge definierter Werte des Merkmals m von r

Sei m_{comp} der ermittelte Wert des Merkmals m

Sei g_{comp} die Gültigkeit für den ermittelten Merkmalwert

Falls $m_{comp} \in m_{def}$

Trage g_{comp} in die m -te Vektorposition von v_{equ} ein

Sei der Ähnlichkeitswert a_r gleich dem Skalarprodukt

aus v_{equ} und v_{weight}

Gebe $\text{MAX}\{a_r \mid r \text{ Rührorgan}\}$ als Ergebnis aus

Sonderstellung des Parameters Rühraufgabe:

Zur Bestimmung der Komponenten des Ähnlichkeitsvektors v_{equ} wird für jedes Merkmal der Gültigkeitswert verwendet, der aus der Regelmenge abgeleitet wurde. Für die Merkmale Viskosität und Rühraufgabe gilt dies jedoch nicht, da diese Parameter durch die Aufgabenstellung vorgegeben sind, so daß die Aufgabenstellung als Grundlage für den Übereinstimmungswert dient. Die angegebene Viskosität wird umgeformt in Term-Zugehörigkeitswert-Verbindungen, so daß der Zugehörigkeitswert im Algorithmus als Gültigkeits- bzw. Ähnlichkeitswert aufgefaßt werden kann. Bei der geforderten Rühraufgabe handelt es sich nicht um eine linguistische Variable, die zudem aus mehreren Grundrühraufgaben zusammengesetzt sein kann. Hinsichtlich der Ermittlung eines Ähnlichkeitswertes der Menge geforderter und ausführbarer Grundrühraufgaben wurden zwei Methoden untersucht:

- (a) **Idee:** Für jede in der Aufgabenstellung geforderte Grundrühraufgabe wird geprüft, ob das betrachtete Rührorgan für diese Aufgabe geeignet ist. Falls die

Rühraufgabe in der Rührerspezifikation enthalten ist, so wird der Ähnlichkeitswert anteilmäßig erhöht (initial = 0), sonst bleibt der Wert unverändert.

Beispiel: Der *Scheibenrührer* ist für die Aufgaben *Emulgieren*, *Begasen* und *Wärmeübergang* einsetzbar. Liegt als zu lösende Aufgabenstellung *Suspendieren* und *Emulgieren* vor, so ergibt sich ein Gültigkeitswert von 0.5, da der Rührer zum *Emulgieren* verwendbar ist (+ 0.5), jedoch nicht zum *Suspendieren* (+ 0).

- (b) Bei der obigen Methode zur Ermittlung des Ähnlichkeitswertes zum Parameter Rühraufgabe fällt auf, daß häufig geringe Ähnlichkeitswerte für Rührer entstehen, obwohl diese in der Fallsammlung zur Lösung der Aufgabe eingesetzt werden.

Bei den Charakterisierungen der Rührorgane wird stets eine eindeutige Aussage in Bezug auf den Einsatzbereich der Rührer gemacht. Das bedeutet, daß eine Rühraufgabe entweder völlig von einem Rührer gelöst werden oder gar nicht. Dies entspricht jedoch nicht der Realität: Da beispielsweise alle Rührer das Rührgut auf gewisse Weise vermischen, sind sie mehr oder weniger zum Homogenisieren geeignet. Dies "mehr oder weniger" sollte in die Spezifikation der einzelnen Rührorgane einfließen.

Die Charakterisierungen der Rührer werden derart vom Experten abgeändert, daß für jede Rühraufgabe ein Beiwert im Bereich [0, 1] hinzugefügt wird, der ein Maß für die Verwendbarkeit darstellt.

Beispiel: Für den *Scheibenrührer* wird nun folgende Definition abgelegt: (Homogenisieren 0.75) (Suspendieren 0.2) (Wärmeübergang 0.75) (Emulgieren 1.0) (Begasen 1.0)

Diese Beiwerte dienen nun bei der Berechnung des Ähnlichkeitswertes als Grundlage: Für jede geforderte Rühraufgabe wird untersucht, inwieweit ein Rührer für diese Aufgabe geeignet ist, d.h. der entsprechende Beiwert wird gemerkt. Bei zusammengesetzten Aufgabenstellungen wird zur Verknüpfung dieser Werte zu einem Ähnlichkeitswert das arithmetische Mittel gebildet.

Beispiel: Sind die Grundrühraufgaben *Suspendieren* und *Emulgieren* in einer Aufgabenstellung gefordert, so ergibt sich für den *Scheibenrührer* ein Ähnlichkeitswert von $(0.2 + 1.0) / 2 = 0.6$

Obwohl detailliertere Informationen bei Methode (b) verwendet werden, ergeben sich gegenüber (a) keine verbesserten Resultate. Es werden jeweils die gleichen Rührer vorgeschlagen, wobei allein die zugehörigen Gesamtähnlichkeitswerte variieren. Die Ursache ist darin zu finden, daß das Merkmal *Rühraufgabe* nur einen von sieben Summanden bei der Bestimmung des Gesamtähnlichkeitswertes darstellt.

2. Regelmenge

Die erweiterte Modellierung deckt anhand der Datensammlung zunächst Zusammenhänge auf, die Eingangsgrößen und Rührermerkmale in Beziehung setzen, und

bildet diese in Form von Regeln ab. Gleichmaßen kann der Zusammenhang von Rührermerkmalen und Rührorgan abgeleitet werden, um die Klassifikation eines Rührers zu ermöglichen. Der in Kapitel 6.2.1 vorgestellte Algorithmus zur Regelgenerierung eignet sich in modifizierter Form zur Erzeugung von Regeln der Bauart:

“IF Merkmal_i = m THEN Rührer = r”

Die Argumentation, die zu der Wahl dieser Regelstruktur führt, verläuft analog zu den Überlegungen bezüglich der Aufgliederung der Rühraufgaben in einzelne Grundrühraufgaben:

Die Menge der relevanten Rührermerkmale setzt sich zusammen aus Drehzahl, Durchmesser Verhältnis, Föderrichtung, Form und Anzahl der Stufen. Da ein Fallbeispiel die Verbindung aller Merkmale mit einem Rührorgan beschreibt, ist folgende Regelstruktur denkbar:

“IF Merkmal₁ = m₁ AND ... AND Merkmal₅ = m₅ THEN Rührer = r”

Die nachstehenden Gründe zeigen jedoch, daß diese Regelformulierung wenig Sinn macht. Es wird kaum Allgemeingültigkeit erzielt, da die Prämissen sehr spezielle Aussagen repräsentieren. Insgesamt können 1000 unterschiedliche Vorbedingungen erzeugt werden, so daß eine nicht überschaubare Regelmenge entstehen kann. Ergänzend ist eine geringe Zahl von begründenden Beispielen zu erwarten, so daß die Aussagekraft der Regeln schwach ausprägt ist. Problemstellungen, die neue Charakterisierungen hervorrufen, können nicht berücksichtigt werden.

Eine Aufteilung der Rührermerkmale erweist sich also als vorteilhaft. Zusammen mit der anfangs erzeugten Regelmenge ergibt sich ein zweistufiges Regelmodell, mit dessen Hilfe das Klassifikationsproblem der Rührerauswahl gelöst werden kann.

Vergleich der Ansätze: Stellt man die beiden aufgeführten Techniken zur Merkmalsverknüpfung nebeneinander, so fallen folgende Unterschiede auf:

- Der erste Ansatz erlaubt eine Klassifikation anhand einer Ähnlichkeitsfunktion, die einen direkten Vergleich des ermittelten Merkmalswerts mit der realen Rührereigenschaft herstellt. Dabei wird eine Auswahl desjenigen Rührers gewährleistet, der mit den berechneten Merkmalen die größte Übereinstimmung erzielt.

Im Gegensatz dazu wird die Lösung beim zweiten Verfahren mittels einer weiteren, aus der Fallsammlung gewonnenen Regelmenge gefunden.

Die Hauptunterschied liegt also in der *Nutzung unterschiedlicher Informationsquellen*. Während die Ähnlichkeitsfunktion einen Bezug zu den tatsächlichen Rührermerkmalen herstellt, extrahiert das regelbasierte Vorgehen diese Zusammenhänge aus der Fallsammlung.

- Vergleicht man die Testergebnisse der repräsentativen Datenmenge, so werden durch Einbindung der Ähnlichkeitsfunktion bessere Resultate als mit der erweiterten Regelmenge erzielt. Im vorangegangenen Abschnitt wurde bereits angeführt, daß die Ähnlichkeitsfunktion eine Trefferquote von 81% aufweist. Dagegen führt das zweistufige Regelmodell nur zu einer Übereinstimmung von 69%. Dieses Ergebnis verdeutlicht die Nachteiligkeit dieses Verfahrens.

Betrachtet man die Häufigkeiten der selektierten Rührer, so wird die eingeschränkte Auswahl bei der erweiterten Regelmenge sichtbar. Sämtliche Klassifikationen beziehen sich auf nur vier Rührorgane: Sigma- (21×), Propeller- (9×), Schrägblatt- (4×) und Trapezrührer (2×). Die übrigen Rührer bleiben selbst bei den Alternativvorschlägen unberücksichtigt.

Im Gegensatz dazu werden mit Hilfe der Ähnlichkeitsfunktion alle Rührertypen klassifiziert. Die Häufigkeiten liegen im Bereich von eins bis elf: Schrägblatt- (11×), Sigma- (10×), Propeller- (7×), Kreuzbalken- (6×), MIG (5×), Trapez- (4×), Scheiben- (3×), Alpha- (1×) und Ankerrührer als Alternativrührer.¹⁴ Insgesamt führt die Verwendung einer Ähnlichkeitsfunktion zu günstigeren Ergebnissen, die die Vielfalt der in Frage kommenden Objekte widerspiegeln.

- Vorteilhaft erweist sich ferner die Verwendung der Ähnlichkeitsfunktion hinsichtlich einer individuellen Beurteilung der Merkmale durch den Anwender. Hier liefert der Gewichtungsvektor zusammen mit dem Ähnlichkeitsvektor eine Einschätzung für die Eignung eines Rührers. Der Gewichtungsvektor ermöglicht die Hervorhebung bestimmter Merkmale und beeinflusst so entscheidend die Bewertung der Rührorgane. Durch eine entsprechende Anpassung dieses Vektors wird erreicht, daß jeder Benutzer unterschiedliche Schwerpunkte bei der Klassifikation setzen kann. Die Verknüpfung durch Regeln eröffnet diese Möglichkeit nicht.
- Ein weiterer Ausbau der Ähnlichkeitsfunktion besteht darin, die Rührerdefinitionen modifizierbar zu gestalten. Damit kann jeder Anwender die einzelnen Rührer nach seinem Verständnis charakterisieren. Dies stellt eine weitere Methode dar, unterschiedliche Betrachtungsweisen in die Verarbeitung einfließen zu lassen.
- Die Erzeugung einer aufbauenden Regelmenge zeigt sich dahingehend positiv, daß die Nähe zu den implizit in der Fallsammlung enthaltenen Zusammenhängen gewahrt bleibt. Die Lösung wird unabhängiger von Expertenwissen, da der Bezug von Rührermerkmalen und Rührorgan nicht von einem Ingenieur hergestellt wird, sondern anhand der konkreten Auslegungsbeispiele gewonnen wird. Bei den Merkmalen *Form*, *Förderrichtung* und *Stufenzahl* handelt es sich um eindeutige Zusammenhänge. Es können jedoch abweichende Aussagen zu den Merkmalen *Drehzahl*, *Durchmesser Verhältnis*, *Viskositätsbereich* und *Rühraufgabe* hinsichtlich des Einsatzgebietes eines Rührers gemacht werden. Dies unterliegt der Beurteilung des

¹⁴Manche Aufgabenstellungen rufen gleich starke Bewertungen für zwei Rührer hervor, so daß die Summe der Häufigkeiten hier größer ist als die Anzahl der Beispiele.

Experten und kann vor allem in Grenzbereichen zu unterschiedlichen Ergebnissen führen. Beispielsweise kann in einer Rührwerksauslegung ein Propellerrührer für ein hochviskoses Produkt gewählt werden, während ein anderer Ingenieur für diesen Viskositätsbereich stets Langsamläufer einsetzen würde.

Durch die Erzeugung von Regeln aus einer Datensammlung werden diejenigen Prinzipien formuliert, nach denen der Ingenieur vorgeht, der diese Rührwerke ausgelegt hat. Es sind keine zusätzlichen Informationen eines Experten notwendig, die jeden Rührer explizit charakterisieren.

- Gleichermaßen kann man bei der Verwendung einer Ähnlichkeitsfunktion die Intention verfolgen, eine möglichst große Informationsmenge aus der Fallsammlung zu gewinnen, um unabhängiger von Expertenwissen zu werden.

Statt der vom Ingenieur vorgegebenen Rührerbeschreibungen bietet sich eine Aufdeckung aller Charakterisierungen aus der Datensammlung an. Der Wissenserwerb kann hier durch Ausnutzung der statistischen Zusammenhänge erfolgen. Für jeden Rührer werden dazu die Häufigkeiten der einzelnen Merkmalwerte gemäß der Fallbeispiele aufgelistet, wodurch die gesuchte Charakterisierung möglich wird.

Beispiel: Der Sigmarührer wird bei der vorliegenden Fallsammlung bezüglich des Drehzahlbereichs folgendermaßen eingesetzt:

Drehzahl "sehr klein": 23 Fälle

Drehzahl "klein": 41 Fälle

Drehzahl "mittel": 23 Fälle

Drehzahl "groß": 4 Fälle

Drehzahl "sehr groß": 0 Fälle

Für eine sinnvolle Beschreibung muß eine geeignete Häufigkeitsgrenze gefunden werden, die selten auftretende und untypische Merkmalswerte unberücksichtigt läßt, aber erkennt, wenn es sich um übliche und festzuhaltende Zusammenhänge handelt. Sowohl ein zu breiter Merkmalraum, als auch ein eindeutiger Merkmalwert erweist sich nachteilig.

Ergebnis: Die Untersuchung der vorliegenden Fallsammlung hat ergeben, daß eine sinnvolle Grenze $1/6$ der Maximalhäufigkeit darstellt.

Das obige Beispiel würde danach den Drehzahlbereich eines Sigmarührers mit "sehr klein", "klein" und "mittel" festlegen. Das sind alle Terme, die durch mindestens 7 Fallbeispiele gestützt werden.

Fazit: Hinsichtlich der deutlichen Vorteile der Ähnlichkeitsfunktion wird die Verwendung der aufbauenden Regelmenge bei den folgenden Betrachtungen vernachlässigt.

7 Verfahren zur Beurteilung der Regelgüte

Im folgenden werden zwei unterschiedliche Ansätze vorgestellt, die eine Einschätzung der Aussagekraft jeder zur Verfügung stehenden Regel wiedergeben¹⁵.

Zur Motivation können nachstehende Begründungen angeführt werden:

- Das oben beschriebene Verfahren zur Regelgenerierung basiert auf einer Datensammlung, die unvollständig ist und Inkonsistenzen beinhaltet. Folglich werden diese Mängel auf die Regelmenge übertragen. Mit Hilfe von Bewertungen soll die Glaubwürdigkeit abgebildet werden, um eine Aussage bezüglich der Regelgüte zu ermöglichen.
- Ferner ist die Menge der erzeugten Regeln sehr umfangreich, da alle in den Fallbeispielen enthaltenen Zusammenhänge zu Regeln transformiert werden¹⁶. Ziel des regelbasierten Ansatzes ist jedoch der Aufbau einer möglichst kompakten Regelmenge, die alle für den Problembereich wesentlichen Zusammenhänge repräsentiert. Die Beurteilung der Regelgüte erlaubt die Erstellung eines überschaubaren und gleichzeitig aussagekräftigen Modells, das eine einfachere Verarbeitung gewährleistet.
- Sind für eine Aufgabenstellung mehrere Regeln anwendbar, so können gegensätzliche Schlußfolgerungen gezogen werden. Aus diesem Grund müssen Kriterien gefunden werden, die eine geeignete Regelanwendung sicherstellen.

Die Gültigkeit der Konklusion ist eng mit der Anwendbarkeit der Prämisse verbunden. Daneben wird die Aussagekraft einer Schlußfolgerung von der Glaubwürdigkeit der Regel selbst beeinflusst, so daß es Sinn macht, eine Regelbewertung zu bestimmen.

7.1 Regelbewertung durch Verrechnung der Fuzzywerte

Der Algorithmus zur Generierung von Regeln betrachtet die vorliegenden Fallbeispiele hinsichtlich der enthaltenen Zusammenhänge. Dabei werden zu den linguistischen Variablen allein die Terme berücksichtigt, während die entsprechenden Zugehörigkeitswerte nicht in die Verarbeitung einfließen. Diese Werte können jedoch zur Ermittlung der Aussagekraft einer Regel ausgenutzt werden.

Zusätzlich bietet sich die Einbeziehung der Anzahl der unterstützenden Beispiele an. Alle in den Regeln formulierten Zusammenhänge werden in unterschiedlich vielen Fallbeispielen repräsentiert. Da eine Verbindung zwischen der Zahl der unterstützenden Fälle und der Glaubwürdigkeit einer Regel besteht, ergibt sich ein weiterer Ausgangspunkt zur Regelbewertung.

¹⁵Dieser Teil des regelbasierten Ansatzes wurde bereits auf Seite 25 als Schritt 2 erwähnt.

¹⁶Näheres zur Größe der erzeugten Mengen findet man in Kapitel 6.3.

Nachfolgend wird mit Hilfe der vorliegenden Fallsammlung versucht, anhand der Parameter *Zugehörigkeitswert* und *Anzahl begründender Fälle* eine Rechenvorschrift herzuleiten, die in geeigneter Weise ein Glaubwürdigkeitsmaß zu einer Regel bestimmt.

7.1.1 Übereinstimmungsgrad der Prämisse

Zur Einschätzung der Aussagekraft einer Regel wird zunächst für jedes Fallbeispiel die Gültigkeit der Prämisse getestet. Ist die Vorbedingung erfüllt, so kann der Grad der Übereinstimmung nach den auf Seite 23 dargestellten Operatoren für die linguistischen Variablen ermittelt werden. Für Variablen mit einem diskreten Wertebereich existieren jedoch keine Zugehörigkeitswerte.

Lösung: Es wird die Vereinbarung getroffen, daß den Termen der Variablen mit diskretem Wertebereich Binärwerte zugeordnet werden. Dies entspricht den Zugehörigkeitswerten der linguistischen Variablen. Eine gedachte Zugehörigkeitsfunktion würde aus Rechtecklinienzügen der Höhe eins bestehen.

Beispiel: Die Variable Grundrühraufgabe besitzt einen diskreten Wertebereich von fünf Elementen. Wird in einem Beispiel die Aufgabe *Homogenisieren mit Wärmeübergang* beschrieben, so ergeben sich die Term-Zugehörigkeitswert-Verbindungen (Aufgabe=Homogenisieren; 1), (Aufgabe=Wärmeübergang; 1).

Die Vorbedingungen der hier erzeugten Regeln setzen sich aus einer UND-Verknüpfung zusammen, so daß die durch ein Fallbeispiel gegebenen Zugehörigkeitswerte mit dem Minimum-Operator verrechnet werden können.

7.1.2 Übereinstimmungsgrad der Konklusion

Trifft die Prämisse einer Regel für ein Beispiel zu, so wird die Gültigkeit der Konklusion überprüft. Es ist zu beachten, daß die hier verwendete Regelstruktur nur einstellige Nachbedingungen zuläßt. Enthält das Fallbeispiel den Konklusionszusammenhang, so wird folgendermaßen ein Übereinstimmungsgrad ermittelt.

Aufgrund der Mehrwertigkeit von linguistischen Variablen müssen außer dem zutreffenden Term auch die gleichzeitig nicht gültigen Terme berücksichtigt werden.

Idee: Es bietet sich an, die Zugehörigkeitswerte dieser Terme als Negativwerte in die Berechnung einfließen zu lassen, da dies eine Regelverletzung darstellt.

Die Summe von Zugehörigkeitswert des gültigen Terms und Durchschnitt der negierten Zugehörigkeitswerte aller nicht geltenden Terme stellt eine sinnvolle Rechenvorschrift zur Ermittlung des Übereinstimmungsgrads der Konklusion dar. Es wird eine gleichmäßige Gewichtung von gültigen und ungültigen Termen erzielt, da der Mittelwert aller Negativwerte Verwendung findet. Würde die Summe der Negativwerte mit dem positiven

Zugehörigkeitswert verrechnet, so würden die ungültigen Terme ein zu starkes Gewicht erhalten. Der Übereinstimmungsgrad liegt somit im Bereich von $(-1, 1]$

Beispiel: Sei eine Rührwerksauslegung gegeben, in der für die Variable *Drehzahl* die Terme “klein” zu 0.6, “mittel” zu 0.8 und “groß” zu 0.4 enthalten sind. Für eine Regel der Form “IF ... THEN Drehzahl = mittel” würde folgendes berechnet:
 $\mu(\text{“mittel”}) + \text{MITTELWERT}(-\mu(\text{“klein”}), -\mu(\text{“groß”})) = 0.8 + \frac{-0.6-0.4}{2} = 0.3$

Entspricht das Beispiel nicht der betrachteten Konklusion, so bleibt dieser Fall für die Regelbewertung unberücksichtigt.

7.1.3 Verknüpfung von Prämisse und Konklusion

Wird der in einer Regel formulierte Zusammenhang durch ein Fallbeispiel gestützt, so gibt sowohl die Vorbedingung als auch die Schlußfolgerung Aufschluß über die Qualität der Regel selbst. Somit bietet sich eine Verrechnung der oben ermittelten Übereinstimmungswerte an, so daß beide Werte gleich stark berücksichtigt werden. Als Rechenvorschrift kommt der Mittelwert oder das Produkt in Frage.

7.1.4 Beurteilung der Regel

Wählt man als Gültigkeitswert einer Regel bezogen auf *ein* Fallbeispiel das obige Verrechnungsschema, so kann aufbauend eine Gesamteinschätzung anhand der Gültigkeitswerte *aller* Fälle gefunden werden. Dazu ist eine Verknüpfungsoperation notwendig, die alle Rechenwerte derart erfaßt, daß eine allgemeine Aussage möglich wird. Eine einfache und zweckmäßige Funktion stellt das arithmetische Mittel dar, da jeder Einzelrechenwert mit gleichem Gewicht in die Gesamtbewertung einfließt.

Zusammenfassend kann folgende Vorgehensweise zur Ermittlung der Aussagekraft einer Regel aufgezeichnet werden:

Algorithmus Regelbewertung

Gegeben sei eine Regel der Form

‘‘IF prem₁ = p₁ AND ... AND prem_i = p_i THEN concl = c’’

Für alle Fallbeispiele n

Falls die Prämisse gültig ist

bestimme Übereinstimmungsgrad der Prämisse =: w-p_n

Falls die Konklusion gültig ist

bestimme Übereinstimmungsgrad der Konklusion =: w-c_n

Verrechne w-p_n mit w-c_n (=: w_n)

sonst sei w_n = NIL

Bestimme arithmetisches Mittel aller w_n ≠ NIL, n ∈ Menge der Fallbeispiele.

⇒ Regelbewertung

Die so bestimmten Gültigkeitswerte liegen im Bereich von $(-1, 1]$, da der Übereinstimmungsgrad der Konklusion in dieses Intervall fällt und der der Prämisse im Bereich $(0, 1]$ liegt. Die Verknüpfungsoperatoren zur Regelbewertung stellen sicher, daß dieser Bereich nicht überschritten wird.

Große Werte nahe eins stehen für eine hohe Gültigkeit, während negative Werte eine Ablehnung der Regel ausdrücken.

7.1.5 Quantitative Bewertung

Die oben beschriebene Vorgehensweise betrachtet allein Zugehörigkeitswerte, um eine Regeleinschätzung zu ermitteln. Dabei wird nicht deutlich, wieviele Fälle in die Berechnung einbezogen worden sind. Die Zahl der Fallbeispiele, die einen Regelzusammenhang unterstützen, gibt aber zusätzlich Aufschluß über die Aussagekraft einer Regel.

Beispiel: Für die Regel

“IF Aufgabe=Homog. AND Viskosität=sehr niedrig THEN Durchm.verh.=klein”
wird nach dem obigen Verfahren eine Bewertung von 0.6 ermittelt, wobei 40 Beispiele diesen Zusammenhang enthalten.

Für die Regel

“IF Aufgabe=Homog. AND Visko.=sehr niedrig THEN Durchm.verh.=sehr groß”
konnte ebenfalls eine Bewertung von 0.6 bestimmt werden, allerdings auf der Grundlage von nur drei Fällen.

Die erste Regel besitzt offensichtlich eine höhere Gültigkeit, obwohl die Bewertung genauso groß ist wie die der zweiten Regel. Erkennbar wird die höhere Gültigkeit allein an der Zahl der Beispiele, die auf die Regel zutreffen.

Somit ist es vorteilhaft, neben der Regelbeurteilung anhand von Übereinstimmungswerten auch die Zahl der unterstützenden Beispiele zu berücksichtigen. Insgesamt wird zu jeder Regel die Bewertung und die Anzahl der damit verbundenen Fälle festgehalten. Inwieweit diese Zahlen in die Klassifikationphase einfließen, wird in Kapitel 8 dargelegt.

7.2 Gewichtung anhand neuronaler Netze

Die im vorangegangenen Abschnitt vorgestellte Methode zur Verrechnung der Zugehörigkeitswerte scheint einleuchtend. Trotzdem ist es offen, inwieweit diese Rechenvorschrift eine geeignete Wiedergabe der Aussagekraft von Regeln darstellt. Aus diesem Grunde wird im folgenden ein alternativer Ansatz untersucht, der auf einer bekannten Vorgehensweise basiert.

Im obigen Verfahren wird die vorliegende Datensammlung genau einmal betrachtet, um alle Gültigkeitswerte zu ermitteln. Ein Ansatz, der sukzessive Gewichtungen ermittelt und in der Lage ist, aus Beispielen zu lernen, besteht in der Verwendung neuronaler Netze.

Inwiefern eine Einbindung dieser Methode in die hier vorliegende Problemstellung möglich ist, wird im Anschluß an die Beschreibung der Idee erläutert.

7.2.1 Allgemeine Beschreibung

Das Vorbild für neuronale Netze in der Informatik findet sich in der biologischen Sichtweise des Gehirns. Durch die netzwerkartige Struktur der Nervenverbände wird eine hochgradig parallele Informationsverarbeitung im Gehirn ermöglicht. Lebewesen besitzen die Fähigkeit, den Organismus und das Verhalten veränderten Bedingungen anzupassen. Sie können aus Erfahrungen lernen und das Gelernte auf neue, ähnliche Situationen anwenden.

Die Orientierung an der Funktionsweise des Gehirns soll in der technischen Informationsverarbeitung dazu dienen, aus Beispielen zu lernen, Beispiele zu verallgemeinern und Muster zu erkennen. Dazu werden Modelle konstruiert, die an die physiologische Gehirnstruktur angepaßt sind.

Aufbau neuronaler Netze:

figure=neuronet.ps

Abb.7.1: Neuronales Netz mit zwei verdeckten Schichten

Ein neuronales Netz setzt sich aus einer Eingabeschicht, einer Ausgabeschicht und eventuell aus mehreren verdeckten Schichten zusammen, wie Abbildung 7.1 veranschaulicht.

Eine Schicht ist aufgebaut aus einer Reihe von Basiseinheiten (mindestens einer), sogenannten Neuronen oder Prozessorelementen (siehe Abbildung 7.2). Ein Neuron kann mehrere Eingänge (Synapsen) aufweisen und besitzt einen Ausgangszustand (Aktivität).

figure=one-neuron.ps

Abb.7.2: Neuron der Schicht s

Die Verbindungen der Prozessorelemente (Netzwerkanten) sind gewichtet und dienen zur Weiterleitung von Informationen. Ziel der Lernalgorithmen ist die Einstellung der Kantengewichte (Verbindungsstärken) anhand von Beispielen, d.h. das Netzwerk wird trainiert.

Die Verarbeitung erfolgt im allgemeinen schichtenweise von unten nach oben. Dazu werden die Eingänge der Eingabeschicht mit Informationen von außen belegt. Die Aktivität eines Neurons ergibt sich aus der Verarbeitung der entsprechenden Eingangswerte und wird entlang der Kanten an die darüberliegende Schicht als Eingangsinformation weitergeleitet ([NN91] und [SchHaGa90]).

7.2.2 Der Backpropagation Algorithmus

Bei dem Backpropagation Algorithmus handelt es sich um einen assoziativen Lernalgorithmus. Das bedeutet, daß ein neuronales Netz trainiert wird, indem eine Menge von Beispielen mit bekannten Ausgangswerten Verwendung findet. Die Eingabeinformationen werden durch das vollständige Netzwerk bis zur Ausgabeschicht geleitet, wo der Fehler bezüglich des wahren Ausgangsvektors bestimmt wird. Das Backpropagation Netzwerk zeichnet sich dadurch aus, daß für diesen Fehler alle Prozessorelemente und Verbindungen verantwortlich gemacht werden. Aus diesem Grunde wird der Fehler in umgekehrter Richtung entlang der Verbindungen bis zur Eingabeschicht propagiert. Dabei erfolgt eine Modifikation der Kantengewichte in Abhängigkeit des Fehlers ([NN91]).

Notation: Zur formalen Beschreibung der Vorgehensweise wird folgende Notation vereinbart:

$x_j^{[s]}$	aktueller Ausgangszustand des j-ten Neurons der Schicht s
o_j	aktueller Ausgangszustand des j-ten Neurons der Ausgabeschicht
d_j	durch das Beispiel gegebener tatsächlicher Ausgangszustand des j-ten Neurons
\underline{d}	Ausgabevektor, der sich aus allen d_j zusammensetzt
\underline{i}	durch das Beispiel gegebener Eingabevektor, dessen Komponenten als Eingang der Neuronen der untersten Schicht dienen
$w_{ji}^{[s]}$	Gewicht der Verbindung des i-ten Neurons der Schicht (s-1) mit dem j-ten Neuron der Schicht s
$e_j^{[s]}$	Fehler des j-ten Neurons der Schicht s
$I_j^{[s]}$	Propagierungsfunktion des j-ten Neurons der Schicht s; häufig Summe der gewichteten Eingangswerte des j-ten Neurons ¹⁷

Vorwärts gerichtete Propagierung: Bei der Propagierung der Eingabewerte zur Ausgabeschicht wird die Aktivität eines Prozessorelements anhand der Eingangsinformationen wie folgt bestimmt:

$$\begin{aligned} x_j^{[s]} &= f \left(\sum_i (w_{ij}^{[s]} \cdot x_i^{[s-1]}) \right) \\ &= f (I_j^{[s]}) \end{aligned}$$

Bei der Funktion f handelt es sich um eine Transferfunktion, die das Resultat der Propagierungsfunktion auf den Ausgangszustand abbildet und differenzierbar sein muß¹⁸. Üblicherweise wird die Sigmoidfunktion als Transferfunktion eingesetzt, die in Abbildung

¹⁷Andere Propagierungsfunktionen werden in [SchHaGa90], Seite 49 f, aufgezeigt.

¹⁸Siehe *Fehlerermittlung*.

7.3 dargestellt ist¹⁹.

$$f(x) = (1.0 + e^{-x})^{-1} \quad \text{und} \quad f'(x) = f(x) \cdot (1.0 - f(x))$$

figure=sigmoid.ps

Abb.7.3: Sigmoidfunktion

Die Sigmoidfunktion hat die Eigenschaft, jeden Eingabewert auf das Intervall (0, 1) abzubilden, wobei große positive bzw. negative Eingabewerte zu Aktivitäten nahe 1 bzw. 0 führen. Durch die Steilheit der Kurve im Bereich von $x=0$ werden Eingabewerte in diesem Bereich stärker auseinandergezogen und daher besser getrennt.

Fehlerermittlung: Sind die Aktivitäten o_j der Neuronen der Ausgabeschicht berechnet worden, so kann der globale Fehler E anhand des Ausgabevektors \underline{d} bestimmt werden:

$$E = 0.5 \cdot \sum_k ((d_k - o_k)^2)$$

Da alle Neuronen für diesen Fehler verantwortlich gemacht werden, ergibt sich der lokale Fehler eines Neurons zu:

$$e_j^{[s]} = -\frac{\delta E}{\delta I_j^{[s]}}$$

Für die Prozessorelemente der Ausgabeschicht folgt damit:

$$e_k^{[output]} = (d_k - o_k) \cdot f'(I_k^{[output]})$$

Dieser Fehler wird rückwärts durch das Netzwerk weitergeleitet, so daß der lokale Fehler der Elemente der verdeckten Schichten wie folgt berechnet wird:

$$e_j^{[s]} = f'(I_j^{[s]}) \cdot \sum_k (e_k^{[s+1]} \cdot w_{kj}^{[s+1]})$$

Minimierung des globalen Fehlers: Ziel des Lernprozesses ist die Minimierung des globalen Fehlers durch Modifikation der Kantengewichte. Die Gewichte können unter Zuhilfenahme der lokalen Fehler der Prozessorelemente angepaßt werden anhand der Delta-Lernregel²⁰ mit *lcoef* als Lernrate:

$$\begin{aligned} \Delta w_{ij}^{[s]} &= -lcoef \cdot (\delta E / \delta w_{ij}^{[s]}) \\ &= lcoef \cdot e_j^{[s]} \cdot x_i^{[s-1]} \end{aligned}$$

Der nachstehende Backpropagation Algorithmus beschreibt die Vorgehensweise einer Trainingsphase. Dieser Ablauf sollte solange wiederholt werden, bis die Kantengewichte so eingestellt sind, daß (nahezu) eine Konvergenz erreicht ist. Es ist zu beachten, daß die Reihenfolge der Lernbeispiele Einfluß auf das Lernverhalten hat und daher in jeder Trainingsphase diese Reihenfolge geändert werden sollte.

¹⁹Andere Transferfunktionen findet man in [SchHaGa90], Seite 50 f.

²⁰Dies basiert auf der biologisch motivierten Hypothese von Hebb, daß eine Verbindung gestärkt wird, wenn beide zugehörigen Neuronen gleichzeitig aktiv sind. Siehe [SchHaGa90].

Algorithmus Backward Propagation

Sei ein Netzwerk aus $n+1$ Schichten $\{0, \dots, n\}$ aufgebaut.

Für jedes Lernbeispiel c

Sei \underline{i} der Eingabevektor von c und

\underline{d} der Ausgabevektor von c

Aktiviere die Neuronen in Schicht 0 gemäß \underline{i}

Für alle Schichten $s = 1, \dots, n$

Bestimme die Aktivität aller Neuronen in s

Bestimme für alle Neuronen in Schicht n den lokalen Fehler

Für alle Schichten $s = n, \dots, 1$

Modifiziere die Kantengewichte $w_{ij}^{[s]}$

Bestimme den lokalen Fehler $e_{[s-1]}$

7.2.3 Verknüpfung von Regelmengen und neuronalen Netzen

Die Einbindung von neuronalen Netzen in den regelbasierten Ansatz kann in Zusammenhang mit der Gültigkeitsbestimmung von Regeln erfolgen. Dazu muß eine Abbildung der Regelmengen auf eine Netzstruktur gefunden werden.

Durch die Verbindung zweier Prozessorelemente in einem neuronalen Netz wird die Verbindung der zugehörigen Neuronen beschrieben, wobei das Kantengewicht angibt, wie stark dieser Zusammenhang ausgeprägt ist. Hier findet sich eine Analogie zu den Regeln:

- Eine Regel beschreibt den gerichteten Zusammenhang von Bedingung und Folgerung. Die Verbindungen im Netzwerk sind ebenfalls gerichtet.
- Eine Regel besitzt eine bestimmte Gültigkeit. Im neuronalen Netz entspricht dies der Verbindungsstärke einer Kante.
- Die vorhandene Beispielsammlung stellt jeweils die Aufgabenstellung und die zugehörige Lösung bereit. Damit sind die Voraussetzungen für den Backtracking Algorithmus (überwachtes Lernen) geschaffen.

Eine Regel der Form “IF premise THEN conclusion” wird daher abgebildet auf eine Netzkante, die vom Neuron “premise” zum Neuron “conclusion” gerichtet ist. Die Verbindungsstärke gibt dann Aufschluß über die Gültigkeit der Regel.

7.2.4 Netzwerkstruktur zur Rührerauswahl

Bei der Transformation der generierten Regelmengen, die sich auf das Problem der Rührerauswahl beziehen, werden die Vorbedingungen abgebildet auf Neuronen der Eingabeschicht. Die Menge der möglichen Vorbedingungen setzt sich zusammen aus allen Ver-

knüpfungen (Grundrühraufgabe \times Viskositätsbereich). Damit ergeben sich 25 Prozessorelemente für die Eingabeschicht des Netzwerkes.

Betrachtet man die in Kapitel 6 erzeugte Regelmenge, die den direkten Zusammenhang von Aufgabenstellung und Rührorgan formuliert, so verkörpern die Neuronen der Ausgabeschicht die möglichen zehn Rührorgane. Da der direkte Zusammenhang dargestellt wird, entstehen keine verdeckten Schichten und das Netz hat folgendes Aussehen:

figure=direct-net.ps

Abb.7.4: Neuronales Netz zur direkten Klassifikation

Die tiefere Modellierung berücksichtigt zusätzlich die Rührermerkmale. Die entsprechenden Regeln schließen zunächst von der Aufgabenstellung auf eine Rührereigenschaft und anschließend auf ein Rührorgan, falls keine Ähnlichkeitsfunktion verwendet wird. Die in den Regeln auftretenden Merkmale sind nicht miteinander verknüpft und werden einzeln betrachtet. So ergibt sich eine Netzwerkschicht, die die Menge ($Wert_k$ des Merkmals[i]) für alle Werte k der relevanten Rührermerkmale i repräsentiert. Das Netzwerk besitzt folgende Struktur:

figure=3-layer-net.ps

Abb.7.5: Dreischichtiges Netzwerk

Soll das Zusammenwirken von Merkmalregeln²¹ und Ähnlichkeitsfunktion untersucht werden, so bildet die Menge der Rührermerkmale in dem zugehörigen Netzwerk die Ausgabeschicht:

figure=merkmal-net.ps

Abb.7.6: Netzwerk zur Merkmalbestimmung

Da die Datensammlung Angaben zu den Rührermerkmalen zur Verfügung stellt, ist die Voraussetzung zum überwachten Lernen durch den Backtracking Algorithmus gegeben.

7.2.5 Einbeziehung der Fallsammlung in den Lernalgorithmus

Grundlage für das Training der oben aufgezeigten Netzwerke bildet die vorliegende Fallsammlung. Dazu muß für jedes Beispiel der Eingabevektor \underline{i} und der Ausgabevektor \underline{d} bekannt sein. Die Komponente k des Eingabevektors \underline{i} dient als Eingangsinformation für das k -te Neuron der Eingabeschicht. Konkret handelt es sich bei dem Rührproblem um den durch das Fallbeispiel formulierten Zusammenhang von Rühraufgabe und Viskositätsbereich. Analog zu der in Abschnitt 7.1.1 beschriebenen Bestimmung des Übereinstimmungsgrads von Prämissen berechnen sich die Komponenten des Eingangsvektors \underline{i} für das neuronale Netzwerk.

²¹ "IF Aufgabe=a AND Viskosität=v THEN Merkmal[i]=m"

Ebenso kann der Ausgabevektor \underline{d} für jedes Fallbeispiel bestimmt werden. Einzelheiten dazu wurden bereits in Abschnitt 7.1.2 *Übereinstimmungsgrad der Konklusion* aufgezeigt.

Beispiel: Seien folgende Parameter durch ein Fallbeispiel vorgegeben:

- Aufgabe = Homogenisieren, Suspendieren
- Viskosität = 50 mPas \equiv (“sehr niedrig” 0.3), (“niedrig” 0.7)
- Rührorgan = Schrägblatt

Daraus resultiert für die Eingänge der Neuronen der Eingabeschicht:

- (Homogenisieren \times sehr niedrig) \leftarrow 0.3
- (Homogenisieren \times niedrig) \leftarrow 0.7
- (Suspendieren \times sehr niedrig) \leftarrow 0.3
- (Suspendieren \times niedrig) \leftarrow 0.7

An den Eingängen der übrigen Neuronen liegt 0 an.

Der Ausgabevektor \underline{d} enthält in der Komponente, die für den Schrägblattrührer steht, eine 1, die restlichen Einträge sind 0.

7.2.6 Transformation der Verbindungsstärken zu Regelbewertungen

Ziel der Einbindung des neuronalen Ansatzes in das Regelmodell ist die Ermittlung von Gültigkeitswerten, die Aufschluß über die Aussagekraft der Regeln geben. Durch die Gleichwertigkeit von Regel und zugehöriger Netzkante besteht ein Zusammenhang von graduell berechnetem Kantengewicht und der Regelbewertung.

Bei der Einstellung der Netzkantengewichte ergeben sich große positive Werte für die Verbindungen, die von vielen Lernbeispielen unterstützt werden und damit eine hohe Konfidenz aufweisen. Demgegenüber entstehen negative Werte, wenn häufig Beispiele auftreten, die entlang dieser Verbindung zu einem Fehler führen.

Für die Transformation der Verbindungsstärken zu Regelbewertungen muß ein geeigneter Gültigkeitsbereich zu den Regeln gefunden werden. Der in Kapitel 7.1 beschriebene Ansatz zur Bestimmung der Aussagekraft von Regeln liefert Bewertungen im Bereich $(-1, 1]$. Werte von 1 stehen hier für eine hohe Akzeptanz, während negative Werte nahe -1 die Regel verwerfen.

In ähnlicher Weise werden die in [KB92] vorgestellten Konfidenzfaktoren (Certainty Factors) aufgefaßt, die wie folgt beschrieben werden können (vergleiche [KB92]).

Einem Zusammenhang der Form “Hypothese H_i gilt, falls die Annahme E gilt” werden zwei Maße $\in [0, 1]$ zugeordnet, die den Grad der Gültigkeit bzw. der Ablehnung der Aussage angeben (im Gegensatz zur klassischen Logik, in der eine Aussage nur wahr oder falsch sein kann). Aus diesem Glaubwürdigkeits- und Fragwürdigkeitsgrad wird nun ein Certainty Factor $CF(H_i|E) \in [-1, 1]$

durch Differenzbildung bestimmt. Er ist ein Maß für die Zustimmung einer Hypothese H_i unter der Voraussetzung, daß die Annahme E gilt.

Bezogen auf Regeln der Form “IF E THEN H ” stellt der Konfidenzfaktor $CF(H|E)$ ein Maß für die Gültigkeit der Regel dar. Positive Werte geben eine Zustimmung und negative Werte eine Ablehnung des Regelzusammenhangs an.

Die Verwendung des Bewertungsmaßstabes $[-1, 1]$ kann somit als Grundlage für die generierte Regelmenge dienen, da auch eine geeignete Regelverarbeitung auf der Basis der Certainty Factors möglich ist (siehe Kapitel 8). Die Verbindungsstärken des hier betrachteten neuronalen Ansatzes liegen außerhalb des Intervalls $[-1, 1]$. Zur Anpassung an die obige Darstellung bieten sich zwei Methoden an.

1. Normierung der Gewichte

Zum einen kommt eine Normierung aller Kantengewichte in Frage, so daß der Maximalwert auf 1 abgebildet wird und das Minimum auf -1. Die restlichen Gewichte werden verhältnismäßig angeglichen.

figure=normalize.ps

Abb.7.7: Normalisierung

2. Obere und untere Schranken

Eine andere Möglichkeit stellt die Verwendung fester Schranken dar. Bei dieser Methode werden alle Verbindungsstärken, die größer als eine obere Schranke sind, auf den Wert 1 gesetzt, und alle Gewichte, die kleiner als eine untere Schranke sind, werden auf den Wert -1 abgebildet. Die Gewichte innerhalb der Schranken werden wie oben verhältnismäßig angepaßt.

figure=bounding.ps

Abb.7.8: Feste Schranken *low* und *up*

Vergleicht man beide Methoden, so stellt sich folgendes heraus:

- Durch die Normierung der Gewichte wird allein der Zusammenhang mit maximalem Kantengewicht mit 1.0 bewertet, d.h. die zugehörige Regel wird akzeptiert. Da alle übrigen Gewichte entsprechend angepaßt werden, wird hier die Regelgültigkeit herabgestuft, obwohl es sich um Kantengewichte handeln kann, die deutlich größer als eins sind.
- Anhand fester Schranken ergeben sich eine Reihe von Regeln, die akzeptiert bzw. verworfen werden. Hier zeichnet die Möglichkeit ab, die Regelmenge auf diejenigen Zusammenhänge zu reduzieren, die eine große Gültigkeit aufweisen.

- Die Regelbewertungen liegen durch die Normierung sehr dicht beieinander, da sich alle Werte im eingegrenzten Intervall $[-1, 1]$ bewegen. Die Differenzierung zweier Werte wird damit undeutlicher als vor der Normierung.
- Bei Verwendung fester Schranken wird lediglich der Abstand der Gewichte innerhalb der Schranken angeglichen. Alle Werte außerhalb dieses Bereichs werden eindeutig 1 und -1 zugeordnet, so daß hier klare Aussagen zur Gültigkeit vorliegen.

Bevor beide Methoden hinsichtlich der erzielten Testergebnisse gegenübergestellt werden, müssen geeignete Schranken zu der zweiten Methoden gefunden werden.

Überlegungen zum Gültigkeitsbereich: Zur Auffindung geeigneter Schranken wird die Menge aller ermittelten Kantengewichte untersucht. In Tabelle 7.1 sind die Gewichte der Netzwerkkanten dargestellt, die den *direkten* Zusammenhang von Aufgabe und Rührorgan repräsentieren. Leere Felder geben an, daß keine Kante und damit keine entsprechende Regel existiert. 10.4 10.5 Es fällt auf, daß ein großer Teil (71%) der Verbindungen ein negatives Gewicht aufweist, sogar 48% kleiner als -1.

Begründung: Die Ursache hierfür liegt vor allem in der Mehrdeutigkeit der Auslegungsbeispiele. Ferner kann die Aufspaltung der Rühraufgaben in Grundrühraufgaben dafür verantwortlich gemacht werden. In beiden Fällen ergibt sich bei Anwendung des Backpropagation Algorithmus für die einzelnen Ausgabeknoten immer bei denjenigen Fallbeispielen ein Fehler, die einen andersartigen Rührer zuordnen.

Beispiel: Zur Homogenisierung eines sehr niederviskosen Produktes wird in 19 von 28 Fällen ein Propellerrührer gewählt. Daneben existieren jedoch Auslegungen für sieben weitere Rührorgane, so daß in diesen Fällen ein Fehler am Ausgang des für den Propeller stehenden Neurons entsteht. Damit wird das entsprechende Kantengewicht herabgesetzt.

Gleiches gilt für die Verbindungsstärken zu den anderen möglichen Rührern.

Die Wahl der Schranken -1.0 und 1.0 wird somit zu ungünstigen Ergebnissen führen, da viele glaubwürdig einzuschätzende Regeln eine negative bzw. ungültige Bewertung erhalten würden. Beispielsweise existieren zur Homogenisierung insgesamt nur zwei positive Kantengewichte. Für diese Rühraufgabe kann jedoch eine Vielzahl von Rührern eingesetzt werden, so daß wesentlich mehr gültige Regeln erzeugt werden müssen.

Konkret ergibt sich für die Schranken -1.0 und 1.0 bei der *direkten Klassifikation* eine Trefferquote von 67%, wenn man die repräsentative Beispielmenge für einen Testlauf zugrunde legt.

Folgerung: Zur Abbildung negativer Kantengewichte auf positive Regelbewertungen, muß der Betrag der unteren Schranke größer als der Betrag der oberen Schranke sein.

Zur Einstellung günstiger Schranken wurden Testläufe mit unterschiedlichen Werten durchgeführt. Ausschlaggebend für eine gute Lösung ist neben der *Trefferquote* die *Größe der Regelmenge* und die *Anzahl der ausgegebenen Alternativlösungen*. Bei letzterem gelten Zahlen von maximal drei Vorschlägen als akzeptabel, wobei zusätzlich höchstens 2.0 Angaben im Durchschnitt gemacht werden sollten. Größere Zahlen setzen die Qualität eines Ergebnisses herab und gelten als inakzeptabel.

Die Zahl der entstehenden gültigen Regeln gibt Aufschluß über den Verarbeitungsaufwand. Als vorteilhaft gilt eine möglichst kompakte und gleichzeitig aussagekräftige Regelmenge.

Tab.7.2: Testergebnisse mit verschiedenen unteren Schranken

untere Schranke	obere Schranke	Trefferquote	‡ Lösungen	‡ gültiger Regeln
-1.5	1.0	64%	akzeptabel	67
-1.7	1.0	69%	akzeptabel	72
-1.8 bis -2.2	1.0	72%	akzeptabel	77
-2.3	1.0	64%	akzeptabel	80
-2.5	1.0	69%	akzeptabel	83
-2.8 bis -3.0	1.0	66%	akzeptabel	91
-4.0	1.0	75%	akzeptabel	102
-5.0	1.0	77%	akzeptabel	111
kleiner -6.0	1.0	$\geq 77\%$	inakzeptabel	> 114

In Tabelle 7.2 sind die Resultate zu unterschiedlichen unteren Schranken aufgeführt. Es fällt auf, daß die Zahl der Alternativlösungen mit der Verringerung der unteren Schranke zunimmt: Während bei einer unteren Grenze von -1.0 im Durchschnitt noch 1.4 Rührer vorgeschlagen werden, so sind es bei einer Grenze von -2.0 bereits 1.67 Rührer und beim Grenzwert -5.0 sogar 1.83 Rührer. Untere Schranken, die kleiner als -5.0 sind, führen zu Durchschnittszahlen größer 2.0 bei mehr als drei Lösungsvorschlägen. Sie gelten daher trotz verbesserter Trefferquote als inakzeptabel. Diese Testläufe zeigen für die Werte um -2.0 und für -5.0 als untere Schranke günstige Ergebnisse.

Offensichtlich nimmt die Zahl gültiger Regeln mit Verkleinerung der unteren Schranke zu. Vergleicht man die Werte -2.0 und -5.0 so stehen 77 (48 positive und 29 negative) Regeln 111 (72 positive und 39 negative) Regeln gegenüber. Da hier 30% mehr Regeln entstehen, wird anschließend als untere Schranke der Wert -2.0 betrachtet.

Welcher Wert erweist sich als obere Schranke vorteilhaft?

Die anfängliche Wahl des oberen Grenzwertes 1.0 wird durch die in Tabelle 7.3 aufgeführten Testergebnisse bestätigt.

Gegenüberstellung: Nun kann ein Vergleich zum *Normierungsverfahren* in Bezug auf die Testergebnisse angestellt werden: Durch die Normierung der Kantengewichte wurde bei der direkten Klassifikation eine Trefferquote von nur 61% erzielt gegenüber 72% durch

Tab.7.3: Testergebnisse mit verschiedenen oberen Schranken

untere Schranke	obere Schranke	Treffer	‡ Lösungen	‡ Regeln mit Bewertung 1.0
-2.0	0.8	58%	akzeptabel	25
-2.0	0.9	66%	akzeptabel	24
-2.0	1.0	72%	akzeptabel	23
-2.0	1.1	66%	akzeptabel	22
-2.0	1.3	52%	akzeptabel	20

die festen Schranken -2.0 und 1.0. Außerdem existiert bei der Normierung allein eine Regel, die mit -1.0 bewertet wird, so daß die Zahl gültiger Regeln (=116) sich fast gar nicht von der Ausgangsregelmenge unterscheidet.

Fazit: Zur Abbildung der Kantengewichte auf die Regelbewertungen werden bei den nachfolgenden Betrachtungen vor allem die festen Schranken -2.0 und 1.0 verwendet.

Modifikationsmöglichkeiten: Es können folgende Abwandlungen zur Einbindung des neuronalen Ansatzes betrachtet werden, um eine Verbesserung der Ergebnisse zu erzielen.

1. Bei der Konstruktion des neuronalen Netzes kann man die generierte Regelmenge unberücksichtigt lassen und ein *vollständiges Netz* zugrunde legen, bei dem jedes Prozessorelement einer Schicht mit allen Neuronen der darüber liegenden Schicht verbunden ist. Dies kommt einer vollständigen Regelmenge gleich, die anhand des Lernalgorithmus auf die gültigen Regeln beschränkt werden kann.

Tab.7.4: Testergebnisse des vollständigen Netzwerks

untere Schranke	obere Schranke	Trefferquote	‡ Lösungen	‡ gültiger Regeln
-2.0	1.0	72%	akzeptabel	168
-5.0	1.0	81%	akzeptabel	225

Die Tabelle 7.4 zeigt die erzielten Resultate auf der Basis der repräsentativen Datenmenge auf. Die Testergebnisse haben sich hierbei nur geringfügig verbessert, während die Größe der Regelmenge sich mehr als verdoppelt hat. Es ist anzumerken, daß fast alle Regeln, die nicht durch den Generierungsalgorithmus erzeugt werden, hier eine negative Bewertung erhalten. Das ist einleuchtend, da diese Zusammenhänge nicht durch die Fallsammlung gestützt werden.

Fazit: Da die Testergebnisse nur unwesentlich verbessert werden und dagegen eine sehr große Regelmenge entsteht, wird der Ansatz auf der Basis eines vollständigen Netzwerkes nicht weiter berücksichtigt.

2. Der Backpropagation Algorithmus kann dahingehend modifiziert werden, als daß man die Kantengewichte *kumulativ* verändert. Das bedeutet, daß die Kantenge-

wichte nicht nach jedem Fallbeispiel angepaßt werden, sondern erst nach einer bestimmten Zahl von Lernschritten. Die einzelnen Gewichtsänderungen Δw_{ij} werden gespeichert, aufsummiert und erst am Ende (beispielsweise nach einem kompletten Durchlauf aller Lernbeispiele) zum Kantengewicht w_{ij} addiert. Durch die kumulative Veränderung der Gewichte wird erreicht, daß bei der Anpassung der Gewichte tatsächlich der Gesamtfehler betrachtet und minimiert wird. Damit ist die Reihenfolge der Beispiele beim Lernprozeß unerheblich (siehe [SchHaGa90]).

Tabelle 7.5 gibt die erzielten Testergebnisse an. Die kumulative Veränderung der Kantengewichte führt zu verbesserten Ergebnissen für die unteren Schranken, die kleiner als -2.0 sind. Daher wird diese Modifikation im folgenden beibehalten.

Tab.7.5: Testergebnisse der kumulativ veränderten Kantengewichte

untere Schranke	obere Schranke	Treffer	\emptyset # Lösungen	# gültiger Regeln
-1.0	1.0	64%	1.6	59
-2.0	1.0	72%	1.72	81
-3.0	1.0	72%	1.88	90
-4.0	1.0	78%	1.92	105
-5.0	1.0	83%	1.92	111

- Bei neuronalen Netzen mit verdeckten Schichten ist es vorteilhaft, die Kantengewichte mit unterschiedlichen Werten zu initialisieren, um die entstehenden Symmetrien in den verdeckten Schichten zu vermeiden. Dadurch kann die Konvergenz eines Netzwerkes verhindert werden. Aus diesem Grunde bietet sich eine Initialisierung der Verbindungsstärken mit Zufallswerten im Intervall $[-0.1, 0.1]$ an, wenn ein Netzwerk mit verdeckten Schichten vorliegt.
- Es sollte untersucht werden, inwieweit eine Einbindung der *Zahl begründender Fallbeispiele* einer Regel in Frage kommt, wie es in dem in Abschnitt 7.1 beschriebenen Verfahren gehandhabt wird.

An dieser Stelle seien diesbezüglich nur die damit erzielten Testergebnisse der direkten Klassifikation in Tabelle 7.6 angeführt, die sich in ähnlicher Weise bei der Modellierung über Rührermerkmale ergeben. Es wird eine Verschlechterung der Resultate um ca. zehn Prozentpunkte bewirkt. Daher ist die Einbeziehung der Zahl unterstützender Fallbeispiele für den Ansatz der neuronalen Netze nicht sinnvoll. Diesbezüglich findet man Erläuterungen im nachfolgenden Kapitel 7.3 unter *Aufwand*.

Testergebnisse zur tieferen Modellierung: Alle in den letzten beiden Abschnitten angeführten Ergebnisse beziehen sich auf den *direkten* Klassifikationszusammenhang. Legt man die *Abbildung über Rührermerkmale* zugrunde, werden die unter 1. und 2. stehenden Ergebnisse erzielt:

Tab.7.6: Testergebnisse anhand der Einbindung der Zahl begründender Fälle

untere Schranke	obere Schranke	Trefferquote	∅ # Lösungen	# gültiger Regeln
-2.0	1.0	61%	1.1	81
-3.0	1.0	64%	1.3	80
-4.0	1.0	69%	1.44	105
-5.0	1.0	72%	1.47	111

1. Klassifikation anhand des *Netzwerks zur Merkmalbestimmung* mit nachfolgender Verwendung der *Ähnlichkeitsfunktion*:

Werden die Netzkantengewichte mittels fester Schranken -2.0 und 1.0 abgebildet, so erzielt man für die repräsentative Datenmenge eine Trefferquote von 78%.

Die Schranken -5.0 und 1.0 sowie das Normierungsverfahren führt in 75% der Fälle zu korrekt zugeordneten Rührorganen.

2. Legt man das *dreischichtige Netzwerk* zugrunde, so entsteht ein zweistufiges Regelmodell. Die diesbezüglichen Testläufe führen zu einer hohen Fehlerquote von über 40%.

7.3 Vergleich der Gewichtungsverfahren

In diesem Abschnitt wird ein Vergleich der Gewichtungsverfahren aus den Kapiteln 7.1 *Gewichtung durch Verrechnung der Fuzzywerte* und 7.2 *Gewichtung anhand neuronaler Netze* angestellt. Dazu werden verschiedene Kriterien untersucht.

7.3.1 Übersicht

Viele Eigenschaften der beiden Ansätze zur Regelbewertung wurden bereits in den zwei vorangegangenen Abschnitten angeführt. Zur Gegenüberstellung werden diese Ergebnisse nochmal tabellarisch in Tabelle 7.7 dargestellt. Dabei wird zwischen *direkter* Klassifikation und der Klassifikation über Rührparameter (mit *Ähnlichkeitsfunktion* oder *zweistufigem* Regelmodell) unterschieden.

7.3.2 Testergebnisse

Vergleicht man die aufgelisteten Trefferquoten, so zeigt das erste Verfahren *Verrechnung der Fuzzywerte* Vorteile gegenüber dem neuronalen Ansatz. Nur bei der direkten Klassifikation unter Verwendung der Schranken -5.0 und 1.0 liegen die Testergebnisse um

²⁴Es wird ein Netzwerk zugrunde gelegt, dessen Kanten der generierten Regelmenge entsprechen, wobei die Kantengewichte kumulativ verändert werden.

²⁵Die Trefferquote von 86% gegenüber der auf Seite 33 angegebenen Quote von 81% basiert auf einer modifizierten Klassifikationsmethode, die auf Seite 70 erläutert wird.

Tab.7.7: Übersicht der Testergebnisse

Klassifik.	Verrechnung der Fuzzywerte			Neuronales Netz ²⁴			
	Treffer	Ø #Lösg.	#Regeln	Schranken	Treffer	Ø #Lösg.	#Regeln
direkt	78%	1.77	117	-2.0; 1.0	72%	1.72	81
				-5.0; 1.0	83%	1.92	111
				normiert	64%	1.35	114
Ähnl.Fkt.	86% ²⁵	1.97	347	-2.0; 1.0	78%	1.88	294
				-5.0; 1.0	75%	1.88	348
				normiert	75%	1.75	354
zweistufig	69%	5.54	457	-2.0; 1.0	52%	1.5	187
				-5.0; 1.0	61%	1.7	325
				normiert	52%	1.67	469

fünf Prozentpunkte höher. In allen anderen Fällen ist der neuronale Ansatz dem ersten unterlegen.

Es ergeben sich bei neuronalen Netzen zwar kürzere Durchschnittszahlen von angegebenen Lösungsvorschlägen, allerdings auf Kosten der Trefferquote.

7.3.3 Größe der Regelmengen

Der Aspekt, eine möglichst kleine Regelmenge zu erzeugen, wird in geeigneter Weise durch Einbindung neuronaler Netze umgesetzt. Bei der Abbildung der Netzkantengewichte auf Regelbewertungen anhand fester Schranken wird ein großer Teil der Regeln als ungeeignet eingestuft, so daß diese Regeln bei der weiteren Verarbeitung nicht mehr berücksichtigt werden.

Die Reduzierung der Regelmenge wird durch den Ansatz der Verrechnung der Fuzzywerte nicht gewährleistet. Dies Verfahren erzeugt größtenteils positive Regelbewertungen, so daß die anfangs generierte Regelmenge vollständig erhalten bleibt, solange alle Regeln mit einer Gültigkeit größer als -1.0 zugelassen sind. Genau hier liegt ein Ansatzpunkt, um eine Regelreduzierung bei diesem Verfahren zu erreichen.

Reduzierung der Anzahl gültiger Regeln: Zur Auffindung einer verkleinerten und gleichzeitig aussagekräftigen Regelmenge bezüglich des Verfahrens der Verrechnung der Fuzzywerte kann die Einführung einer unteren Schranke dienen, so daß alle Regeln mit einem Gültigkeitswert kleiner als diese Schranke nicht bei der Verarbeitung berücksichtigt werden. Ähnlich wie bei dem neuronalen Ansatz muß durch Testläufe eine solche Schranke bestimmt werden.

Es stellt sich heraus, daß für die direkte und die tiefere Modellierung unterschiedliche Schranken geeignet sind. Die Ursache liegt darin, daß alle Regeln, die den direkten Zusammenhang formulieren, eine Bewertung größer als 0.5 besitzen, während die Regeln, die

auf Rührermerkmale schließen auch negative Gültigkeitswerte aufweisen. In Tabelle 7.8 sind die Ergebnisse unterschiedlicher Grenzwerte aufgeführt. Für die direkte Formulierung ergibt sich sogar eine Erhöhung der Trefferquote, wenn man eine untere Schranke im Bereich von 0.6 bis 0.73 wählt.

Aufgrund der aufgelisteten Resultate ist es vorteilhaft, zur direkten Klassifikation alle Regeln zuzulassen, die anhand der Bewertungsmethode *Verrechnung der Fuzzywerte* Gültigkeitswerte größer als 0.73 aufweisen. Die Klassifikation über Rührermerkmale bietet sich auf der Grundlage aller Regeln an, die eine höhere Bewertung als 0.4 besitzen.

Durch Verwendung dieser Schranken wird eine wesentliche Verkleinerung der Regelmenge erzielt.

Tab.7.8: Testergebnisse für verschiedene untere Schranken

Art der Modellierung	untere Schranke	Trefferquote	Anzahl gültiger Regeln
direkt	-1	78%	117
	0.6	81%	115
	0.7	81%	92
	0.8	41%	64
	0.76	69%	81
	0.73	81%	89
über Merkmale	-1	81%	347
	0.3	81%	277
	0.4	81%	245

7.3.4 Aufwand

Betrachtet man die beiden Gewichtungsverfahren hinsichtlich des erforderlichen Aufwandes, so stellt sich folgendes heraus: Das zuerst vorgestellte Verfahren der Verrechnung der Fuzzywerte legt die Gültigkeit aller Regeln mittels einer einmaligen Betrachtung der vorhandenen Datensammlung fest. Anhand eines Durchlaufs werden gleichzeitig die Regelpbewertungen und die Zahl begründender Beispiele bestimmt.

Im Gegensatz dazu liegt der Aufwand, die Regelpbewertungen durch neuronale Netze zu ermitteln, erheblich höher. Zur Einstellung geeigneter Kantengewichte muß der Backpropagation Algorithmus vielfach angewendet werden, damit annähernd eine Konvergenz erreicht wird.

Ist eine passende Einstellung der Gewichte entstanden, so ist dort bereits implizit die Zahl begründender Fälle enthalten. Im Gegensatz zum ersten Verfahren werden nämlich bei dem neuronalen Netz *alle* propagierten Knoten auf der Ausgabeschicht hinsichtlich der Übereinstimmung mit dem Lernbeispiel überprüft. Übertragen auf eine Regelmenge entspricht dies einer Einflußnahme aller Beispiele auf eine Konklusion, für die die Prämisse gültig ist. Damit fließt in die graduell ermittelte Regelpbewertung die Anzahl der Regeln

ein, für die Prämisse *und* Konklusion gültig ist.

Beispiel: Das auf Seite 45 angeführte Beispiel wird hier wieder aufgegriffen:

Die Methode der Verrechnung der Fuzzywerte liefert folgende Einschätzung zur Gültigkeit zweier Regeln:

“IF Aufgabe=Homog. AND Viskosität=sehr niedrig THEN Durchm.verh.=klein”
zu 0.6 mit 40 begründenden Fallbeispielen;

“IF Aufgabe=Homog. AND Viskosität=sehr niedrig THEN Durchm.verh.=sehr groß” zu 0.6 mit 3 begründenden Fallbeispielen.

Das Verfahren zur Bestimmung dieser Gültigkeitswerte berücksichtigt nur diejenigen Fallbeispiele, für die sowohl Vorbedingung als auch Schlußfolgerung gelten. Darauf bezieht sich auch die Zahl der unterstützenden Fallbeispiele. Diese Zahl dient der Beurteilung der ermittelten Gültigkeitswerte im Verhältnis zu anderen Regeln.

Die Einbindung neuronaler Netze liefert dagegen nachstehende Kantengewichte:

“IF Aufgabe=Homog. AND Viskosität=sehr niedrig THEN Durchm.verh.=klein”
mit einem Kantengewicht von 0.71;

“IF Aufgabe=Homog. AND Viskosität=sehr niedrig THEN Durchm.verh.=sehr groß” mit einem Kantengewicht von -4.44.

Das hohe negative Gewicht zu der zweiten Regel entsteht dadurch, daß viele Beispiele zwar der Prämisse genügen, jedoch der überwiegende Teil nicht mit der Konklusion übereinstimmt, so daß häufig ein Fehler entsteht, der das Kantengewicht herabsetzt.

7.3.5 Umgang mit Inkonsistenzen der Datensammlung

Es wurde bereits erwähnt, daß bedingt durch die inkonsistente Datensammlung die Kanten der neuronalen Netze häufig mit negativen Werten eingestellt werden, obwohl die entsprechenden Regeln als gültig einzuschätzen sind. Bei der Transformation der Kanten zu Regeln, können die negativen Gewichte dadurch kompensiert werden, daß der Betrag der unteren Schranke kleiner als der Betrag der oberen Schranke gewählt wird, um eine Verlagerung in den positiven Bereich zu bewirken.

Bei dem Ansatz der Verrechnung der Fuzzywerte wirken sich gegensätzliche Auslegungsbeispiele nicht auf den Gültigkeitsgrad aus, da nur diejenigen Beispiele in die Bewertung einer Regel einfließen, die sowohl mit der Vorbedingung als auch mit der Schlußfolgerung übereinstimmen. Die Inkonsistenz wird dagegen in den Zahlen begründender Fallbeispiele festgehalten: Besitzen zwei Regeln die gleichen Prämissen, aber verschiedene Konklusionen, so können die Regeln durchaus ähnlich hohe Gültigkeitswerte aufweisen, obwohl sie im Widerspruch zueinander stehen. Statt dessen gibt die jeweilige Zahl begründender Fallbeispiele an, wie die in den Regeln formulierten Aussagen zahlenmäßig unterstützt werden. Näheres zur Verarbeitung der erzeugten Regelmengen wird im anschließenden Kapitel 8 dargestellt.

Fazit: Da sowohl die Testergebnisse als auch der Aufwand zur Bestimmung der Gültigkeitswerte anhand der Verrechnung von Fuzzywerten vorteilhaft gegenüber der Einbindung neuronaler Netze sind, wird diese Methode für die vorliegende Problemstellung als Grundlage genommen. Allein die Größe der entstehenden Regelmenge liegt im Verhältnis zu neuronalen Netzen grundsätzlich höher. Jedoch kann eine Reduzierung der Regelmenge durch die Einführung unterer Schranken (siehe Seite 59) erreicht werden, so daß die Zahl der gültigen Regeln nahezu gleich groß ist.

Somit zeigt sich insgesamt, daß das Bewertungsverfahren der Verrechnung von Fuzzywerten aus Kapitel 7.1 gegenüber der Einbindung neuronaler Netze zu favorisi

8 Regelbasierte Objektauswahl

Nachdem Untersuchungen zu geeigneten Regelstrukturen und zu Verfahren zur Regelbewertung durchgeführt wurden, werden nachfolgend Überlegungen hinsichtlich der Regelanwendung angestellt²².

8.1 Regelverarbeitung

Die erzeugte Regelmenge zeichnet sich dadurch aus, daß jede Regel eine bestimmte Glaubwürdigkeit besitzt und Regeln mit gleicher Vorbedingung existieren, die zu widersprüchlichen Schlußfolgerungen führen. Es muß also eine Vorgehensweise gefunden werden, die eine passende Regelauswahl und damit eine sinnvolle Klassifikation gewährleistet.

8.1.1 Einbeziehen der Regelbewertung

Die Regeln sind so aufgebaut, daß eine vorwärtsverkettende Verarbeitung angebracht ist. Um nun eine bestimmte Aufgabenstellung zu lösen, muß für jede vorliegende Regel der Grad ermittelt werden, zu dem die Vorbedingung erfüllt ist (siehe Kapitel 7.1.1 *Übereinstimmungsgrad der Prämisse*). Neben dieser Einflußgröße muß zur geeigneten Regelauswahl die Bewertung der Regel selbst berücksichtigt werden. Setzt man diese beiden Größen ins Verhältnis, so resultiert ein Gültigkeitswert für die Regelkonklusion. Hier wird eine zu *Certainty Factors* ähnliche Vorgehensweise erkennbar. Dieser Ansatz verwendet Bewertungen des Wahrheitsgehaltes von Fakten und Regeln und ermöglicht die Einschätzung des Gültigkeitsgrads der Konklusion. In [?] wird darauf hingewiesen, daß Fuzzy Logic eine geeignete Basis für Berechnungen mit Certainty Factors darstellt, indem Zugehörigkeitswerte verknüpft werden. Danach ergibt sich als Gültigkeitswert für die Schlußfolgerung einer Regel das Produkt aus Regelbewertung und Übereinstimmungsgrad der Vorbedingung. Diese Verknüpfung entspricht einem UND-Operator und ist zweckmäßig, da die

²²Dieser Teil des regelbasierten Ansatzes wurde auf Seite 25 als Schritt 3 beschrieben

Konklusion genau dann gültig ist, wenn die entsprechende Regel *und* die Vorbedingung gilt.

Falls eine Schlußfolgerung durch mehrere Regeln hergeleitet wird, so müssen die jeweiligen Gültigkeitswerte miteinander verrechnet werden. Da diese Werte Alternativcharakter haben, bietet sich eine ODER-Verknüpfung an. Es wird der bei den Certainty Factors verwendete Operator benutzt:

Seien K_1 und K_2 zwei zu verrechnende Konklusionsbewertungen. Dann ergibt sich die Gesamtbewertung K zu:

$$K = \begin{cases} K_1 + K_2 - (K_1 \cdot K_2) & \text{falls } K_1, K_2 > 0 \\ K_1 + K_2 + (K_1 \cdot K_2) & \text{falls } K_1, K_2 < 0 \\ \frac{K_1 + K_2}{1 - \min\{|K_1|, |K_2|\}} & \text{sonst} \end{cases}$$

Die beiden in Kapitel 7 vorgestellten Verfahren zur Bestimmung von Gültigkeitswerten lieferten Regelbewertungen im Intervall $[-1, 1]$. Übereinstimmungswerte der Prämisse können im Bereich $[0, 1]$ liegen. Insgesamt ergeben sich damit für Schlußfolgerungen Gültigkeitswerte zwischen -1 und 1.

Änderung des Gültigkeitsbereichs: Bei allen vorangegangenen Untersuchungen wurde von einem Gültigkeitsbereich von $[-1, 1]$ für die Regeln ausgegangen. Häufig wird zur Propagierung von Fuzzyregeln auch das Intervall $[0, 1]$ zugrunde gelegt. Hier besitzen Regeln mit einer Bewertung von 0 keine Gültigkeit im Gegensatz zum Ansatz der Certainty Factors, bei dem diese Regeln eine negative Bewertung nahe -1 aufweisen. Der in Kapitel 7.1 vorgestellte Ansatz wurde bereits auf Seite 59 derart modifiziert, daß alle gültigen Regeln eine Bewertung größer 0 besitzen. Der in Kapitel 7.2 vorgestellte neuronale Ansatz kann ebenfalls leicht modifiziert werden, um eine Regelbewertung im Bereich $[0, 1]$ zu erzeugen. Bei den betreffenden Testläufen stellte sich heraus, daß generell die Tefferquote um drei Prozentpunkte abnimmt. Es ist anzumerken, daß hier Schranken von $[-2, 1]$ bzw. $[-1, 1]$ verwendet werden müssen, die auf den Bereich $[0, 1]$ abgebildet werden, damit Lösungen im akzeptablen Bereich entstehen.

Heuristik bei der Regelpropagierung: Die in den Aufgabenstellungen vorgegebenen Rühraufgaben setzen sich häufig aus mehreren Grundrühraufgaben zusammen. Jede dieser Teilaufgaben, die die Vorbedingung einer Regel erfüllt, erhält den gleichen Übereinstimmungswert von 1.0, d.h. daß jeder Teilaufgabe eine gleich große Bedeutung zukommt. In der Praxis gilt diese Annahme jedoch nicht unbedingt: Bei einer Kombination von Grundrühraufgaben, die unter anderem aus Homogenisieren bzw. Wärmeübergang besteht, besitzen diese Teilaufgaben in der Regel eine untergeordnete Bedeutung. Diese Rühraufgaben stellen in der Kombination sehr allgemeine Anforderungen dar, während die übrigen drei Grundaufgaben spezielle Rührereigenschaften erfordern.

Umsetzung: Um diese Differenzierung bei der rechnergestützten Verarbeitung zu berücksichtigen, werden zunächst bei jeder Aufgabenstellung die Teilaufgaben untersucht. Handelt es sich um eine Aufgabe, die außer Homogenisieren bzw. Wärmeübergang auch andere Grundaufgaben beinhaltet, so wird bei der Regelverarbeitung die Gültigkeit der Prämissen wie folgt bestimmt:

- Vorbedingung enthält Homogenisieren \Rightarrow Übereinstimmungsgrad = 0.2
- Vorbedingung enthält Wärmeübergang \Rightarrow Übereinstimmungsgrad = 0.5
- sonst gilt: Übereinstimmungsgrad = 1.0

Setzt sich die Rühraufgabe allein aus Homogenisieren bzw. Wärmeübergang zusammen, so wird ein Übereinstimmungsgrad von 1.0 verwendet, da keine andere Grundaufgabe hervorzuheben ist.

Die Gewichtungen wurden von Experten angegeben.

Auf diese Weise wird die Bedeutung derjenigen Resultate gestärkt, die aus den Grundaufgaben Suspendieren, Emulgieren und Begasen abgeleitet werden.

8.1.2 Propagieren der Zahl begründender Fälle

Die Aussagekraft einer Regel hängt neben dem Gültigkeitswert von der Zahl der Fallbeispiele ab, die den Regelzusammenhang unterstützen (siehe 7.1.5)²³. Bei der Regelverarbeitung sollten diese Zahlen mitberücksichtigt werden.

Analog zum vorangegangenen Abschnitt müssen Operatoren festgelegt werden, die die Propagierung der Zahl begründender Fälle ermöglichen.

- Leitet man aus den gegebenen Variablen der Aufgabenstellung anhand einer Regel eine Schlußfolgerung ab, so sollte die Zahl begründender Fallbeispiele der Regel auf die Konklusion übertragen werden.
- Handelt es sich um eine Regelverkettung so wird die Gültigkeit einer Schlußfolgerung von der Glaubwürdigkeit der Regel *und* der Vorbedingung bestimmt.
 1. Setzt sich die Vorbedingung aus mehreren, mit einem Operator verknüpften Variablen zusammen, so bietet sich folgende Verrechnungsvorschrift der Anzahl Beispiele an:
 - UND-Verknüpfung \Rightarrow Mittelwert der Zahlen begründender Fälle
 - ODER-Verknüpfung \Rightarrow Summe der Zahlen begründender Fälle
 2. Die Zahl der Fallbeispiele, die die Regel und die Prämisse unterstützen, wird verrechnet und auf die Konklusion übertragen. Analog zum vorangegangenen Abschnitt entspricht dies einer UND-Verknüpfung, so daß hier die Bildung des Mittelwertes dieser Zahlen angebracht ist.

²³Wird eine Bewertung der Regeln anhand neuronaler Netze zugrunde gelegt, so kann dieser Verarbeitungsschritt entfallen. Man kann von der Annahme ausgehen, daß hier allen Regeln 0 begründende Fälle zugeordnet sind.

- In der Situation, daß eine Konklusion anhand von mehreren Regeln gefolgert wird, ist eine Verrechnung der Zahl begründender Fälle angebracht. Die Tatsache, daß eine Aussage durch mehrere Regeln abgeleitet wird, erhöht die Gültigkeit dieser Aussage. Daher bietet sich eine Addition der Anzahl begründender Fälle an.

Beispiel: Seien folgende Regeln gegeben:

“IF Aufgabe=Homogen. AND Viskosität=niedrig THEN Drehzahl=groß”

gestützt von 24 Beispielen;

“IF Aufgabe=Suspend. AND Viskosität=niedrig THEN Drehzahl=groß”

gestützt von 16 Beispielen;

“IF Drehzahl=groß THEN Rührer=Propeller” gestützt von 30 Beispielen.

Die Aufgabenstellung Homogenisieren und Suspendieren eines niederviskosen Produktes führt mit den ersten zwei Regeln zu der Aussage “Drehzahl=groß” mit $24 + 16 = 40$ Beispielen. Damit kann die dritte Regel angewendet werden, so daß die Aussage “Rührer=Propeller” durch $(40 + 30) / 2 = 35$ Beispiele unterstützt wird.

8.2 Klassifikation

Nachdem die Verarbeitung der vorliegenden Regeln beschrieben wurde, werden nachfolgend Überlegungen hinsichtlich der Klassifikation der propagierten Aussagen angestellt.

8.2.1 Auswahlkriterien

Die ausgehend von einer konkreten Aufgabestellung gefolgerten Schlußfolgerungen besitzen zwei Attribute, die Aufschluß über die Gültigkeit der Aussage geben: der Gültigkeitswert und die Zahl der unterstützenden Beispiele. Es stellt sich die Frage, auf welche Weise eine geeignete Auswahl der Schlußfolgerungen mit Hilfe dieser Attribute getroffen werden kann.

Idee: Gültigkeitswerte geben größeren Aufschluß in Bezug auf die Gültigkeit einer Aussage als die Zahl begründender Beispiele, da sie auf qualitativen Übereinstimmungswerten basieren und nicht allein auf Häufigkeiten. Trotzdem bietet die Anzahl begründender Fallbeispiele die Möglichkeit einer Vorauswahl der gefolgerten Aussagen, da dieses Attribut Schlußfolgerungen geringer Gültigkeit erkennt: Alle Aussagen, die von verhältnismäßig wenig Beispielen gestützt werden, sollten bei der weiteren Verarbeitung unberücksichtigt bleiben.

Anschließend kann eine Klassifikation anhand der Gültigkeitswerte erfolgen.

1. Zahl begründender Fälle

Inwieweit eine Vorauswahl der resultierenden Aussagen hinsichtlich der Zahl begründender Fälle vorgenommen werden kann, hängt von dem Maximalwert dieser Zahlen ab.

Beispiel: Wird für eine Aufgabenstellung ein Sigmarührer mit 14 begründenden Fällen und ein Kreuzbalkenrührer mit 8 begründenden Fällen gefolgert, so sollten beide Aussagen in die Klassifikation einbezogen werden, da die Attributwerte ähnlich groß sind. Ergeben sich jedoch Anzahlen von 50 gegenüber 8 begründenden Fällen, so deutet dies darauf hin, daß der erste Rührer eine wesentlich höhere Aussagekraft besitzt als der zweite. Damit kann der zweite Rührer aus der Menge möglicher Kandidaten gestrichen werden.

Ergebnis: Untersuchungen der vorliegenden Fallsammlung haben gezeigt, daß folgende Auswahlmethode in Abhängigkeit des Maximums MAX der begründenden Fälle sinnvoll ist:

$MAX < 5$	\Rightarrow Verwendung aller gefolgerten Elemente
$5 \leq MAX \leq 10$	\Rightarrow Elemente mit Anzahl Beispiele > 2
$10 < MAX \leq 30$	\Rightarrow Elemente mit Anzahl Beispiele $\geq MAX/3$
$30 < MAX$	\Rightarrow Elemente mit Anzahl Beispiele $\geq MAX/2$

2. Gültigkeitswerte

Nach der Vorauswahl der gefolgerten Aussagen wird eine Klassifikation abhängig von den zugehörigen Gültigkeitswerten durchgeführt. Diese Werte bewegen sich im Bereich $[-1,1]$, wobei Werte nahe 1 für eine hohe Gültigkeit stehen. Es bieten sich folgende Klassifikationsmöglichkeiten an:

- (a) Die Wahl der Schlußfolgerung mit der größten Konfidenz stellt eine begründete Klassifikation dar, obwohl damit keine Alternativlösungen aufgezeigt werden.
- (b) Die Angabe aller Schlußfolgerungen, die absteigend nach Gültigkeitswerten sortiert sind, entspricht zwar der Forderung nach Alternativlösungen, allerdings stellt dieses Konzept keine sinnvolle Klassifikation dar.

Bei der direkten Klassifikation entsteht hier keine Auswahl von zwei oder drei passenden Rührern, da in der Regel fast alle abgeleiteten Rührer aufgelistet werden, wobei der Gültigkeitswert allein die Position in der Liste bestimmt.

- (c) Ein geeigneter Ansatz, der die Vorteile der ersten beiden Methoden verbindet, besteht darin, zunächst die Gültigkeitswerte der Schlußfolgerungen absteigend zu sortieren und anschließend diejenigen Folgerungen als Ergebnis der Klassifikation anzugeben, die die "besten" Bewertungen aufweisen.

Zur Spezifikation des Begriffs "beste" Bewertungen wurden Tests anhand der vorliegenden Datensammlung durchgeführt. Es stellte sich heraus, daß für die in den vorangegangenen Kapiteln beschriebenen Lösungsansätze unterschiedliche Kriterien sinnvoll sind. Allgemein führt die Selektion aller Schlußfolgerungen, deren Konfidenzen höchstens um 0.1 bzw. um 0.2 vom Maximum abweichen, zu günstigen Ergebnissen.

Bezeichnung: Im folgenden wird diese Methode als "*best-of*" Auswahl, die maximale Abweichung als *Distanz* bezeichnet.

Bemerkung 1: Die in den vorausgegangenen Kapiteln angeführten Testergebnisse wurden ausnahmslos mit einer Distanz von 0.1 ermittelt.

Bemerkung 2: Bei der Klassifikation über Rührermerkmale unter Verwendung der Ähnlichkeitsfunktion entstehen Ähnlichkeitswerte bis zu einem Wert von 5.0 für die einzelnen Rührer. In diesem Fall sollte eine andere Distanz eingesetzt werden. Alle bereits angegebenen Testergebnisse wurden hier mit einer Distanz von 0.9 erzielt.

- (d) Bei der vorliegenden Problemstellung müssen sowohl Variablen mit diskretem Wertebereich, als auch linguistische Variablen klassifiziert werden.

Betrachtet man die propagierten Schlußfolgerungen zu einer linguistischen Variablen, so können diese Aussagen mit entsprechendem Gültigkeitswert als Term-Zugehörigkeitswert-Verbindungen aufgefaßt werden. Damit bietet sich als weitere Auswahlmethode die *Rückübersetzung* der als Term-Zugehörigkeitswert ausgelegten Schlußfolgerungen an. Diese Methode kann anhand der in Kapitel 5.1.3 beschriebenen Defuzzifizierungstechnik durchgeführt werden.

Begründung: Alle Schlußfolgerungen werden aus einer Aufgabenstellung abgeleitet und sind nebeneinander zu einem gewissen Grad gültig. Da es sich um linguistische Terme handelt, die miteinander durch die Zugehörigkeitsfunktion verbunden sind, ist eine Selektion durch Defuzzifizierung angebracht.

Beispiel: Mit Hilfe des regelbasierten Ansatzes soll das Durchmesser Verhältnis für eine bestimmte Aufgabenstellung ermittelt werden. Nach der Regelverarbeitung ergeben sich die in Tabelle 8.1 angegebenen Schlußfolgerungen. Die Defuzzifizierung anhand der Flächenschwerpunktmethode lie-

Tab.8.1: Gefolgerte Konklusionen
zum Parameter *Durchmesser Verhältnis*

Term	Gültigkeitswert
sehr groß	0.42
groß	0.67
mittel	0.36
klein	0.63
sehr klein	0.0

fert für diese Werte ein Durchmesser Verhältnis von 0.524. Ist statt eines konkreten Parameterwertes allein ein linguistischer Term gesucht, so wird der gefundene Parameterwert anhand der Zugehörigkeitsfunktion in eine Term-Zugehörigkeitswert-Verbindung transformiert. Für das Durchmesser Verhältnis ergibt sich damit als Resultat: (“mittel” 0.905), (“groß” 0.094).

Betrachtet man die Variablen, die einen diskreten Wertebereich besitzen, so fällt auf, daß manche Wertebereiche aus voneinander unabhängigen Termen

bestehen (Grundrühraufgaben und Förderrichtung), während sich andere Wertebereiche aus voneinander abhängigen Termen zusammensetzen (Stufenzahl und Form der Rührer). Für die zweite Gruppe bietet sich eine Klassifikation durch Defuzzifizierung an. Die dazu notwendige Zugehörigkeitsfunktion kann man sich als Rechtecklinienzüge der Höhe eins vorstellen.

8.2.2 Vergleich von Defuzzifizierung und “best-of” Auswahl

Als Grundlage für den Vergleich der Methoden 3 und 4 zur Objektauswahl dienen Testläufe auf der Basis der repräsentativen Beispielmenge.

Dazu wird der Klassifikationsansatz über Rührparameter betrachtet.

Die Regeln der Form “IF Aufgabe= a AND Viskosität= v THEN Merkmal $_i$ = m_i ” werden anhand des in Kapitel 7.1 *Gewichtung der Regeln durch Verrechnung der Fuzzywerte* aufgezeigten Ansatzes bewertet.

Die Abbildung der gefolgerten Rührermerkmale erfolgt mit Hilfe der vorgestellten Ähnlichkeitsfunktion.

Zur Objektauswahl werden die Methoden

1. Klassifikation der linguistischen Variablen *Drehzahl*, *Durchmesser Verhältnis*, *Viskosität* und der Variablen *Form* durch Defuzzifizierung; Selektion der übrigen Parameter durch “best-of” Auswahl mit einer Distanz von 0.1,
2. Klassifikation *aller* Parameter anhand der “best-of” Auswahl mit einer Distanz von 0.1

gegenübergestellt.

- Betrachtet man die Zuordnungen der Rührparameter mit der “best-of” Auswahlmethode, so fällt auf, daß häufig diskontinuierliche Alternativlösungen angegeben werden. Beispielsweise werden für einen Drehzahlbereich gleichzeitig die Werte “klein” und “sehr groß” vorgeschlagen.

Die “best-of” Auswahl gewährleistet somit keine angeglichenen Lösungsvorschläge, da stets diejenigen Terme mit den größten Gültigkeitswerten angeführt werden, auch wenn damit eine sprunghafte Änderung verbunden ist.

Im Gegensatz dazu führt die Klassifikation durch Defuzzifizierung stets zu zusammenhängenden Lösungen aufgrund der Flächenschwerpunktmethode. Damit werden gegensätzliche Bereichsangaben vermieden, so daß verbesserte Ergebnisse erzielt werden.

Konkret führen die diskontinuierlichen Ergebnisse der “best-of” Auswahl in 9% der Fälle zu einer falschen Rührerauswahl, bei denen durch Defuzzifizierung eine korrekte Zuordnung erfolgt.

- Vorteilhaft erweist sich die Defuzzifizierungsmethode außerdem hinsichtlich der Anzahl aufgeführter Ergebnisse. Es handelt sich immer um ein oder zwei Vorschläge für einen Parameterwert, was angemessen bei Wertebereichen mit maximal fünf Elementen ist.

Der “best-of” Ansatz liefert hier keine Beschränkung: Alle gefolgerten Terme einer Variablen, die eine ähnlich hohen Gültigkeit besitzen, werden vorgeschlagen. So liefern zahlreiche Testfälle für die Stufenzahl gleichzeitig eine einstufige als auch eine mehrstufige Auslegung. Die Defuzzifizierung führt hier ausnahmslos zu eindeutigen Zuordnungen.

Ebenso werden mit dem “best-of” Ansatz häufig drei von möglichen fünf Termen gewählt. Dies ist eine unangemessen große Anzahl, die wenig Differenzierungen erlaubt.

- Während mit der “best-of” Auswahl immer die Terme mit dem größten Gültigkeitswert klassifiziert werden, entsteht durch Defuzzifizierung eine Art Mittelung der Terme. Diese Mittelung führt teilweise zu ungünstigen Ergebnissen.

Beispiel: In der Fallsammlung der wird zum Homogenisieren und Suspendieren eines mittelviskosen Produktes ein Sigmarührer eingesetzt, der eine *mittlere Form* aufweist.

Durch Anwendung der Regeln werden für den Parameter Form die Terme *mittel* und *sehr strömungsgünstig* mit ähnlich hohem Gültigkeitswert gefolgert. Dies weist auf einen Sigma- oder Propellerrührer hin, der eine sehr strömungsgünstige Form besitzt. Anhand der Defuzzifizierungsmethode resultiert jedoch der Wert *strömungsgünstig*. Damit stimmt das Merkmal Form bei Anwendung der Ähnlichkeitsfunktion weder mit dem Sigma- noch mit dem Propellerrührer überein. Dies führt zum unerwünschten Ergebnis, das einen Kreuz- oder Trapezrührer favorisiert. Beide Rührer haben eine strömungsgünstige Form.

Die “best-of” Methode klassifiziert bei der obigen Aufgabenstellung dagegen den erwarteten Sigmarührer.

Dieser Effekt läßt sich ebenfalls bei linguistischen Variablen erkennen, allerdings wirkt sich die Defuzzifizierung hier geringfügiger auf das Gesamtergebnis aus, da die abgelegten Rührerdefinitionen für diese Variablen mehrere Werte zulassen im Gegensatz zum Parameter *Form*.

Bis auf den zuletzt genannten Aspekt erweist sich die Defuzzifizierungsmethode gegenüber der “best-of” Auswahl als vorteilhaft.

Idee: Zur Verbesserung der ungünstigen Ergebnisse, die durch Defuzzifizierung des Merkmals *Form* hervorgerufen werden, sollte diese Variable mit dem “best-of” Ansatz ausgewählt werden.

Diese Modifikation führt bei der Testreihe zu einer Verbesserung der Ergebnisse um fünf Prozentpunkte auf eine Trefferquote von 86%.

Legt man statt einer Distanz von 0.1 eine erweiterte Distanz von 0.2 zugrunde, so kann oftmals eine Optimierung der Testergebnisse um drei Prozentpunkte erzielt werden. Die Vergrößerung der Distanz führt zu einer erhöhten Angabe von Lösungsvorschlägen, wobei teilweise inakzeptable Resultate mit bis zu acht Rühreralternativen entstehen. Der Distanzwert von 0.1 ist daher zweckmäßig.

9 Erweiterte Problemstellung

Die in den vorangegangenen Kapiteln angeführten Überlegungen beziehen sich speziell auf Lösungsmethoden zur Rührorganauswahl. In den nachstehenden Betrachtungen werden darauf aufbauende bzw. verallgemeinernde Problemstellungen untersucht.

9.1 Eignung von neu entwickelten Rührern

Die vorliegende Datensammlung und das regelbasierte Verfahren zur Lösung des Rührproblems können als Grundlage verwendet werden, um andersartigen Fragestellungen im Zusammenhang mit Rührwerksauslegungen nachzugehen.

Zur Klassifikation von Rührorganen hinsichtlich einer Rühraufgabe wurden Regeln erzeugt, die eine vorwärtsverkettende Verarbeitung erfordern. Betrachtet man die Verbindung *Rühraufgabe* \rightarrow *Rührorgan* in der entgegengesetzten Richtung, so entsteht die Beziehung *Rührorgan* \rightarrow *Rühraufgabe*. Da die eingangs ermittelten Regeln und deren Gültigkeit aus ungerichteten Zusammenhängen entstanden sind, macht es Sinn, diese Umkehrung zu verfolgen.

Sowohl Regeln als auch deren Bewertungen wurden aus der vorhandenen Fallsammlung erzeugt und spiegeln somit die dort enthaltenen Beziehungen wider. Schließt man nun mit Hilfe der Regeln ausgehend von bestimmten Rührorganen auf Aufgabenstellungen, so gewinnt man Informationen darüber, auf welche Weise diese Rührorgane innerhalb der Datensammlung eingesetzt werden. Beispielsweise erhält man Aufschluß über die konkrete Verwendung von Schrägblattrührern in Bezug auf Rühraufgabe und Viskositätsbereich. Da diese Angaben ebenfalls als Expertenwissen bei der Charakterisierung der einzelnen Rührer abgelegt worden sind, kann ein Vergleich gezogen werden zwischen der Einschätzung des Experten und der durch die Fallsammlung repräsentierten Auslegung.

Untersucht man in gleicher Weise die Verbindung *Rührermerkmal* \rightarrow *Rühraufgabe*, die einer Umkehrung der Regeln zur tieferen Modellierung entspricht, so deutet sich ein neuer Anwendungsbereich an: Die Möglichkeit, aus einer Menge von Rührereigenschaften auf Rühraufgaben zu schließen, kann derart ausgenutzt werden, als daß Aussagen bezüglich des Einsatzbereichs von neu entwickelten Rührorganen gemacht werden können. Kon-

kret handelt es sich darum, bei dem Entwurf eines neuen Rührers im Vorfeld zu prüfen, inwieweit dieser für die angestrebten Rühraufgaben geeignet sein wird.

figure=back-chain.ps

Abb.9.1: Zusammenhang *Merkmale* \rightarrow *Rührer*

An dieser Stelle zeigt sich, in welchem Maße die generierte Regelmenge Allgemeingültigkeit besitzt: Die Frage nach dem Einsatzgebiet neu konzipierter Rührer stützt sich auf Zusammenhänge, die aus anderen, bereits bekannten Auslegungen gewonnen werden und in einem völlig neuen Kontext Anwendung finden. Mit Hilfe dieser Problemstellung wird somit eine Einschätzung bezüglich der Universalität der erzeugten Regeln möglich.

Realisierung: Um die oben angeführten Fragestellung mit Hilfe der generierten Regelmenge lösen zu können, müssen Überlegungen hinsichtlich der Verarbeitung angestellt werden.

Ausgehend von den bereitgestellten Merkmalen *Drehzahl*, *Durchmesser Verhältnis*, *Förder richtung*, *Form* und *Stufenzahl* eines neu entwickelten Rührorgans werden alle Regeln untersucht, die rückwärtsverkettend anwendbar sind. Das bedeutet, falls die Konklusion der Regeln "IF Aufgabe=a AND Viskosität=v THEN Merkmal_i=m_i" zutrifft, werden die Aussagen der Vorbedingung anerkannt. Der Grad der Akzeptanz ist abhängig von der Gültigkeit der Regel. Daher wird jedem Prämissenelement der entsprechende Gültigkeitswert zugeordnet. Bei der rechnergestützten Realisierung wurden diejenigen Regeln zugrunde gelegt, die anhand der in Kapitel 7.1 beschriebenen Methode durch Verrechnung von Fuzzywerten bewertet worden sind. Folglich wird den Komponenten *Aufgabe* und *Viskosität* der Gültigkeitswert und die Zahl begründender Fallbeispiele zugewiesen.

Da eine Reihe von Regeln derart Anwendung finden, kommt es vor, daß gleiche Rühraufgaben und Viskositätsbereiche mehrfach aus Regeln abgeleitet werden. Aus diesem Grunde muß eine Verrechnungsvorschrift für diese Werte gefunden werden. Bei der Umsetzung am Rechner sind verschiedene Verfahren getestet worden, wobei sich die Verrechnung durch Mittelwertbildung als geeignet erwiesen hat, sowohl für die Gültigkeitswerte als auch für die Anzahl begründender Beispiele.

Beispiel: Seien für einen neuen Rührer u.a. die Merkmale

Form = *mittel* und *Stufen* = *einstufig* gegeben.

Damit können die Regeln

"IF Aufgabe = Suspendieren AND Viskosität = niedrig THEN Form = mittel"
zu 0.83 mit 16 begründenden Fällen und

"IF Aufgabe = Suspendieren AND Viskosität = niedrig THEN Stufen = einstufig"
zu 0.65 mit 8 begründenden Fällen angewendet werden.

Falls der Rühraufgabe *Suspendieren* und dem Viskositätsbereich *niedrig* noch keine Werte zugeordnet worden sind, ergibt sich für beide Aussagen eine Bewertung von 0.74 mit 12 begründenden Fällen.

Nach Anwendung aller Regeln, ergibt sich häufig für jede der fünf Rühraufgaben und Viskositätsbereiche ein Gültigkeitswert und eine Zahl unterstützender Fallbeispiele. Um daraus Aussagen bezüglich des Einsatzbereichs für einen neuen Rührer ableiten zu können, müssen die “besten” Objektwerte selektiert werden. Dazu erweist sich folgendes Auswahlverfahren als vorteilhaft:

Für die *Rühraufgaben* wählt man das Maximum der Gültigkeitswerte und alle Rühraufgaben mit Gültigkeitswerten, die einen Abstand von höchstens 0.6 zum Maximum aufweisen (siehe “best-of” Auswahl zu Kapitel 8).

Demgegenüber wird für den *Viskositätsbereich* zunächst eine Vorauswahl in Abhängigkeit von der Zahl begündender Fallbeispiele getroffen, die sich analog zu dem Auswahlverfahren auf Seite 65 verhält. Danach werden die besten der verbleibenden Viskositätsbereiche anhand der Gültigkeitswerte mit der “best-of” Auswahl (Distanz = 1.0) bestimmt.

Die Verwendung unterschiedlicher Selektionskonzepte für Grundrühraufgabe und Viskosität ist aus der Motivation entstanden, die Rühraufgaben nicht von der Zahl begründender Fallbeispiele abhängig zu machen. Da die in der Datensammlung auftretenden Fallbeispiele nicht alle Aufgaben mit gleicher Häufigkeit repräsentieren, sind selten auftretende Rühraufgaben, wie *Begasen* und *Emulgieren* den “Standardaufgaben” *Homogenisieren* und *Wärmeübergang* anzahlmäßig unterlegen. Daher wird hier allein der Gültigkeitsgrad betrachtet.

Für den Viskositätsbereich gilt dies nicht. Eine Vorauswahl aufgrund der Zahl unterstützender Fälle ist angebracht, um selten vorkommende Bereiche ausklammern zu können.

Testergebnisse: Um einen Eindruck von der Eignung dieses Ansatzes zu erhalten, werden einige Testergebnisse angeführt. Als Grundlage wurden zwei nicht in der Datensammlung enthaltene und damit als “neu” einzuschätzende Rührorgane geprüft, bei denen der Einsatzbereich allerdings bekannt ist (*Zahnscheibe* und *Ekato Interprop*). Außerdem wurden einige Rührer aus der Fallsammlung zur Kontrolle verwendet. Tabelle 9.1 stellt die erzielten Ergebnisse für die unbekannteren Rührer dar. Diese Resultate geben das tatsächliche Einsatzgebiet der Rührer wieder. Die Zahnscheibe wird in der Regel für Dispergieraufgaben (*Begasen* und *Emulgieren*) im niederviskosen Bereich eingesetzt. Der Interprop-Rührer weist bis auf das Merkmal *Form* ähnliche Eigenschaften wie der Propellerührer auf. Daher kommt ein Einsatz in Produkten niedriger bis mittlerer Viskosität in Frage, wobei vor allem die Rühraufgaben *Homogenisieren* und *Wärmeübergang* gelöst werden. Die Testläufe zu den geläufigen Rührern führten ebenfalls zu günstigen Ergebnissen: Schnellläufer wurde ein niedrigviskoser Bereich zugeordnet, Langsamläufer ein Viskositätsbereich von niedrig bis hoch. Typische Grundaufgaben, wie Wärmeübergang

Tab.9.1: Ermittelter Einsatzbereich für zwei neue Rührorgane

Rührorgan	Merkmal	Wert	Viskosität	Rühaufgaben
Zahnscheibe	Drehzahl	mittel groß sehr groß	<i>niedrig</i> <i>sehr niedrig</i> <i>mittel</i>	<i>Begasen</i>
	Durchm.-Verh.	mittel klein		
	Förd.-Richt.	radial		
	Form	sehr kantig		
	Stufenzahl	einstufig		
Interprop	Drehzahl	mittel groß	<i>niedrig</i> <i>sehr niedrig</i> <i>mittel</i>	<i>Homogenisieren</i> <i>Wärmeübergang</i> <i>Begasen</i>
	Durchm.-Verh.	sehr klein klein mittel		
	Förd.-Richt.	axial		
	Form	mittel		
	Stufenzahl	einstufig		

für Ankerrührer und Begasungsvorgänge für Scheibenrührer, wurden erkannt.

Diese Resultate veranschaulichen, daß die generierte Regelmenge in Bezug auf neue Fragestellungen geeignet ist. Außerdem stellen die Testergebnisse eine Bestätigung für die Allgemeingültigkeit der generierten Regeln dar.

9.2 Verallgemeinerte Fragestellungen

Neben der Betrachtung des Problems der Rührerauswahl aus einem anderen Blickwinkel bietet es sich an, das in den Kapiteln 6 bis 8 entwickelte Lösungsverfahren auf andersartige Fragestellungen mit vergleichbaren Voraussetzungen zu übertragen.

Im Zusammenhang mit der Auslegung von Rührwerken steht die Frage nach der erforderlichen spezifischen Leistung. In Kapitel 2.4 wurde bereits erläutert, auf welche Weise sich die spezifische Leistung berechnen läßt. Diese Größe gibt Aufschluß über die Komponenten Rührwelle und Antrieb beim Entwurf eines Rührwerks.

Die Größe der spezifischen Leistung kann nur schwer im vorhinein für ein Rührwerk abgeschätzt werden und muß mit der aufgezeigten Rechenmethode bestimmt werden. Es wäre daher vorteilhaft, allgemeine Zusammenhänge zwischen den Daten eines Rührwerks und der zugehörigen spezifischen Leistung aufzudecken, um so einen Vorstellung von dieser Größe zu gewinnen. Als Einflußfaktoren gelten die Größen *Viskositätsbereich*, *Rührorgan* und *Rühaufgabe*. Da jeder Fall der vorliegenden Datensammlung sowohl die Einflußgrößen als auch die erforderliche *Leistung* und das *Füllvolumen* beinhaltet, können aus

diesen Informationen allgemeine Zusammenhänge herausgefiltert werden. Damit entsteht eine Grundlage für neue Rührwerksauslegungen, auf der eine direkte Ableitung der Größe der spezifischen Leistung möglich wird.

Realisierung: Die Vorgehensweise zur Konstruktion eines entsprechenden Lösungsverfahrens verläuft analog zur Problemlösung der Rührorganauswahl, die in den Kapiteln 6 bis 8 vorgestellt wurde:

Zunächst wird eine Regelmenge generiert, die aus Regeln der Form

“IF Aufgabe= a AND Viskosität= v AND Rührer= r THEN spezifische Leistung= l ”

besteht. Die Variable *spezifische Leistung* wird als linguistische Variable vereinbart. Der entsprechenden Zugehörigkeitsfunktion liegt ein logarithmischer Maßstab zur Basis 2 zugrunde, wobei sieben linguistische Terme definiert sind (“extrem niedrig” bis “extrem hoch”).

Für die hier angeführten Testergebnisse wurde die Regelmenge durch Einbindung neuronaler Netze bewertet und nach der in Kapitel 8 beschriebenen Methode propagiert.

Auf dieser Grundlage wurde in 72% aller Fälle der repräsentativen Beispielmenge eine richtige Angabe zur spezifischen Leistung gemacht. Da dieses Ergebnis ebenso wie die Resultate zur Rührorganauswahl auf einer inkonsistenten Datensammlung basieren, ist die Bedeutung der Trefferquoten höher einzuschätzen.

Das Beispiel der spezifischen Leistung verdeutlicht, daß eine Übertragung des entwickelten Lösungsverfahrens auf andere Problemstellungen möglich ist und günstige Ergebnisse erzielen kann.

10 Zusammenfassung

In dieser Arbeit werden unterschiedliche Vorgehensweisen zur Lösung von fallbasierten Klassifikationsproblemen im Bereich der Verfahrenstechnik untersucht. Beispielfhaft wird bei der Auslegung von Rührwerken das Problem der Rührorganauswahl für die verschiedenen Verfahren zugrunde gelegt.

Die Wahl eines Rührorgans beruht auf dem Erfahrungswissen eines Ingenieurs, wobei die Einflußgrößen *Rühraufgabe* und *Viskosität des Rührguts* ausschlaggebend sind. Da eine vollständige Angabe und verbale Formulierung des Expertenwissens nur bedingt möglich ist, wird zur Wissensakquisition eine Menge von Rührwerksauslegungen verwendet, in der implizit Informationen über die Vorgehensweise des Ingenieurs beim Entwurf der Anlagen enthalten sind.

Grundsätzlich eignen sich für eine Aufgabenstellung mehrere Rührorgane. Aus diesem Grund zeichnet sich die vorliegende Sammlung von Rührwerksauslegungen dadurch aus, daß bei gleichen Anforderungen unterschiedliche Rührer selektiert werden. Die vorgestell-

ten Lösungsverfahren müssen in geeigneter Weise diese Inkonsistenzen handhaben.

Die Grundlage für alle betrachteten Ansätze bilden zwei Modelle:

- A. Die direkte Abbildung der Einflußgrößen der Aufgabenstellung auf ein Rührorgan.
- B. Die Herleitung der Zusammenhänge
(*Rühraufgabe, Viskosität*) \rightarrow *Rührermerkmale* \rightarrow *Rührorgan*.

Zunächst wird in dieser Arbeit eine *statistische Auswertung* der Datensammlung durchgeführt, wobei Testläufe ein Resultat von 75% richtig ausgewählten Rührern erzielten. Nachteilig erweist sich bei dem statistischen Verfahren jedoch die fehlende Transparenz des Lösungsweges für den Benutzer und die häufige Wahl der Rührer mit hoher a-priori Wahrscheinlichkeit. Ferner stehen keine alternativen Lösungsvorschläge zur Verfügung, obwohl dies bei der betrachteten Problemstellung wünschenswert ist.

Zur Vermeidung dieser Mängel wird eine regelbasierte Vorgehensweise entwickelt, die auf dem Konzept der *unscharfen Logik* beruht. Der Lösungsweg dieses Ansatzes setzt sich aus drei Schritten zusammen:

1. Regelerzeugung
2. Regeltbewertung
3. Regelverarbeitung

Zur Regelgenerierung erweist sich eine *direkte Ableitung* der Regeln aus den in der Fallsammlung enthaltenen Zusammenhängen gegenüber dem *ID3 Algorithmus* als vorteilhaft.

Um eine kompakte, aussagekräftige Regelmenge zu erhalten, die einen Vergleich der Regeln in Bezug auf die Güte erlaubt, werden zwei Verfahren gegenübergestellt, die *Gültigkeitswerte* zu den Regeln mit Hilfe der Fallsammlung auffinden:

Einerseits wird eine Methode aufgezeigt, die Bewertungen durch *Verrechnung der Zugehörigkeitswerte* bestimmt. Zum anderen wird eine *Einbettung neuronaler Netze* in die Regelmenge durchgeführt, wobei die Gewichte der Netzwerkkanten auf die Gültigkeitswerte der Regeln abgebildet werden.

Der Vergleich der beiden Ansätze zeigt, daß neuronale Netze schlechtere Ergebnisse bei Testläufen bewirken und die Konfidenzen aufwendiger zu ermitteln sind. Zwar entsteht durch die Einbindung eines neuronalen Netzes eine kompakte Regelmenge, jedoch kann dies ebenfalls bei der Methode der Verrechnung der Zugehörigkeitswerte durch Einführung einer unteren Schranke erzielt werden, so daß insgesamt letzteres Verfahren zu favorisieren ist.

Bei der Regelverarbeitung über Rührermerkmale ist generell die Abbildung der Merkmale auf einen Rührer anhand einer *Ähnlichkeitsfunktion* gegenüber der Abbildung mit Hilfe von *zusätzlichen Regeln* zu bevorzugen. Es ergibt sich ein Testergebnis von 86% korrekter

Klassifikationen, wobei die fehlerhaft eingestuften Ergebnisse in den meisten Fällen als akzeptabel einzuordnen sind. Die Formulierung über Rührermerkmale erweist sich auch in Bezug auf weitere Eigenschaften positiv gegenüber der direkten Klassifikation.

Überträgt man das entwickelte, regelbasierte Konzept auf andere verfahrenstechnische Probleme, die eine Aufdeckung versteckter Zusammenhänge aus einer Datensammlung erfordern, so führt dies ebenfalls zu günstigen Resultaten. Desweiteren lassen sich Fragestellungen, die eng mit der Rührorganauswahl verwandt sind, wie die Bestimmung des Einsatzbereichs neu entwickelter Rührer, auf der Basis der generierten Regelmenge lösen.

Insgesamt liefert der auf eine Fallsammlung aufbauende, regelbasierte Ansatz eine geeignete Grundlage, den Ingenieur bei Problemstellungen zu unterstützen, deren Lösungen sonst durch Erfahrungswissen oder hohen Rechenaufwand gefunden werden. Dieses Verfahren genügt den Gesichtspunkten *Erklärungskomponente*, *Alternativlösungen* und liefert *günstige Resultate*.

Literatur

- [BrSe87] Bronstein, I.N., Semendjajew, K.A.: Taschenbuch der Mathematik, Verlag Harri Deutsch, Thun 1987
- [CoHu92] Cohen, M.E., Hudson, D.L.: Approches to the Handling of Fuzzy Input Data in Neuronal Networks, in IEEE International Conference on Fuzzy Systems
- [ED89] Elmer-Dewitt, P.: Time for Some Fuzzy Thinking, in TIME, 25. September 1989
- [Ga91] Gariglio, D.: Fuzzy in der Praxis, in Elektronik 1991, Heft 20
- [Gr69] Grasselli, A.: Automatic Interpretation and Classification of Images, S.97-105, Academic Press Inc. 1969
- [HeHo91] Hetzmann, H., Hommel, G.: Fuzzy Logik für die Automatisierungstechnik? in atp 1991, Heft 10
- [KB91] Kleine Büning: Vorlesung "Wissensbasierte Systeme", WS 1991/92
- [KB92] Kleine Büning: Vorlesung "Diagnose und Konfiguration", SS 1992
- [KiKrFr91] Kiendl, H., Krabs, M., Fritsch, M.: Rule-Based Modelling of Dynamical Systems, Analysis and Control of Industrial Processes, Vieweg 1991
- [Ma88] Mayer, O.: Programmieren in COMMON LISP, BI-Wiss.-Verlag 1988
- [Na1993] Narazaki, H.: Implementation of Fuzzy Systems using Multi-layered Neuronal Network, in IEEE International Conference on Neuronal Networks 1993
- [NN91] N.N.: Neuronal Computing, Neuronal Ware 1991
- [SchHaGa90] Schönberg, E., Hansen, N., Gawelczyk, A.: Neuronale Netzwerke, Markt&Technik Verlag 1990, Kapitel 1-4
- [St84] Steele, G.: Common LISP: The Language, Digital Equipment Corporation 1984
- [Ste93] Firma Stelzer, Unterlagen zum Seminar: Praxis der Rührtechnik, 27. September 1993
- [Sti94] Stieß: Mechanische Verfahrenstechnik 1, S.198-240, Springer 1994
- [SuTschVe93] Sulzberger, S., Tschichold-Gürman, N., Vestile, S.: FUN: Optimization of Fuzzy Rule Based Systems Using Neuronal Networks, in IEEE International Conference on Neuronal Networks 1993

- [Tr90] Trautzel, G.: Unscharfe Logik: Fuzzy Logik, in der elektroniker 1990, Heft 3
- [vAl91] von Altrock, C.: Über den Daumen gepeilt, in c't 1991, Heft 3
- [Wa89] Waller, L.: Fuzzy Logic: It's comprehensible, it's praktikal - and it's commercial, in Electronics, März 1989
- [WeZi91] Weber, R., Zimmermann, H.J.: Automatische Akquisition von unscharfem Expertenwissen, in KI 1991, Heft 2
- [Wo91] Wolf, Th.: Fuzzy Logik, in mc, März 1991
- [Won88] Wonneberger, R.: Kompaktführer \LaTeX , Addison-Wesley 1988
- [YoCa74] Young, T.Y., Calvert, T.W.: Classification, Estimation and Pattern Recognition, Elsevier 1974
- [Za84] Zadeh, L.A.: Making computers think like people, in IEEE Spectrum, August 1984

A Hinweise zur Implementation

Alle in dieser Arbeit beschriebenen Lösungsansätze sind in der Programmiersprache LISP realisiert worden.

Im Anschluß findet man Informationen zu den dabei verwendeten Datenstrukturen und implementierten Funktionen. Desweiteren werden im Anhang A.3 Hinweise zur Benutzung des erstellten Systems gegeben.

A.1 Statistischer Ansatz

Die Realisierung der Klassifikation von Rührorganen mit Hilfe des statistischen Ansatzes, der in Kapitel 4 aufgezeigt worden ist, basiert auf den Datenstrukturen und Funktionen der folgenden Dateien:

- **“data/all-cases”** enthält die gegebene Sammlung von Rührwerksauslegungen, wobei die einzelnen Fälle nummeriert und in einer Liste zusammengefaßt sind. Ein Fallbeispiel besteht wiederum aus einer Liste von Objekt-Wert-Listen.
- **“tool-def.lisp”** beinhaltet die vom Ingenieur gelieferten Eigenschaften der Rührer: *Propeller-, Schrägblatt-, Scheibe-, Sigma-, MIG-, Trapez-, Kreuz-, Anker- und Alpharührer*. Desweiteren sind die Zugehörigkeitsfunktionen der linguistischen Variablen und die Wertebereiche der Aufzählungstypen abgelegt.
- **“convert-input-data.lisp”** weist die in “data/all-cases” abgelegte Datensammlung der globalen Variablen ***stored-cases*** zu. Anschließend werden die Daten derart transformiert, daß die Werte der linguistischen Variablen in Term-Zugehörigkeitswert-Verbindungen umgeformt werden. Ferner werden die Rührermerkmale teilweise anhand der von Ingenieur gelieferten Rührerdefinitionen bestimmt (siehe “tool-def.lisp”). Die globale Variable ***case*** dient zur Erfassung der transformierten Daten. Diese Variable ist ein Array [1..188] von Hashtabellen. Die Objektwerte eines Falles i werden unter dem Schlüssel “Objektname” in der Hashtabelle ***case*[i]** eingetragen.
- **“stat-classification.lisp”** enthält Funktionen, die der Trainings- und Klassifikationsphase des statistischen Ansatzes entsprechen:
 - **(init-all-density-fun)** dient zur Ermittlung der Dichtefunktionen und a-priori-Wahrscheinlichkeiten.

Grundlegende Datenstrukturen:

(a) *aufg.stat.fun*, *visko.stat.fun* = Hashtabellen mit den Schlüsseln *Rührer*, *Drehzahl*, *Durchmesser Verhältnis*, *Förderrichtung*, *Form*, *Stufenzahl* von Matrizen. Die Matrizen repräsentieren die Dichtefunktionen $f_{k,auf}$ und $f_{k,vis}$ der Klassen C_k .

(b) *a.prior* Hashtabelle mit den in (a) genannten Schlüsseln von Arrays, die die a-priori-Wahrscheinlichkeiten der einzelnen Klassen beinhalten.

- (**classify-object** *class c.auf c.vis*) führt eine Klassifikation des Merkmalvektors (*c.auf c.vis*) durch, wobei eine Klasse aus der Menge *class* als Resultat geliefert wird.

Desweiteren sind in der Datei “stat-classification.lisp” zwei Funktionen enthalten, die die Testläufe auf der Basis der repräsentativen Datensammlung durchführen:

- * (**test-stat-direct-tool** *para*) führt eine direkte Klassifikation des Parameters *para* durch. Neben der Angabe von *para = ruehrer* können auch einzelne Rührermerkmale klassifiziert werden.
- * (**test-stat-para-tool**) realisiert die Klassifikation über Rührparameter (tiefe Modellierung). Dabei werden zunächst alle Rührermerkmale anhand des statistischen Ansatzes klassifiziert, so daß anschließend eine Abbildung der Merkmale auf ein Rührorgan anhand der Ähnlichkeitsfunktion möglich wird. Näheres zur Ähnlichkeitsfunktion findet man auf Seite 87.

A.2 Regelbasierter Ansatz

Alle implementierten Verfahren bezüglich des regelbasierten Ansatzes sind derart realisiert worden, daß grundsätzlich beliebige Klassifikationsprobleme gelöst werden können. Voraussetzungen für diese Problemstellungen sind

1. Zu den relevanten Objekten (Variablen) muß entweder eine Zugehörigkeitsfunktion vorhanden sein (stetiger Wertebereich), oder die (diskrete) Wertemenge muß bekannt sein.
2. Der gesuchte Klassifikationszusammenhang muß in Form von Regeln formulierbar sein. Bei mehrstufigen Regelmodellen muß eine Regelverkettung möglich sein. 187

figure=conn-rules.ps
3. Es muß eine Datensammlung zur Verfügung stehen, die diese Zusammenhänge in bereits gelösten Fallbeispielen beinhaltet.

A.2.1 Datenstrukturen

(a.) Dem in den Kapiteln 5 bis 6 vorgestellten regelbasierten Ansatz liegen nachstehende Datenstrukturen zugrunde:

- ***counter*** Ganze Zahl, die die Anzahl vorliegender Fallbeispiele der Datensammlung enthält.
- ***fuzzy*** Liste aller linguistischen Variablen (Objekte).

enum Liste aller Variablen, die einen diskreten Wertebereich besitzen (endliche Menge von Termen).

indep Liste aller Variablen, deren zugehörige Terme unabhängig voneinander sind (z.B.: *Förderrichtung*).

dep Liste aller Variablen, deren zugehörige Terme voneinander abhängig sind (z.B.: *Form* und alle linguistischen Variablen).

number-of-objects Liste, die für ein n -stufiges Regelmodell $n + 1$ ganzzahlige Elemente besitzt, die die Anzahl verschiedener Variablen der Stufe $i = \{0, \dots, n\}$ angeben (nachfolgend sei $n + 1$ stets die Stufenzahl des Regelmodells).

variables-at-layer Array $[0..n]$ von Listen;
Komponente i enthält die Liste aller Variablen der i -ten Stufe.

object-operand Liste mit Elementen $\in \{\text{“AND”}, \text{“OR”}\}$. Das Listenelement an Position i gibt den Verknüpfungsoperator der Aussagen der Stufe i an.

number-of-terms Liste von $n+1$ ganzen Zahlen, wobei die Zahl an Position i die Gesamtzahl der Terme bzw. Termkombinationen (bei “AND”- Verknüpfung) der Stufe i angibt.

coded-case Array $[1..*\text{counter}^*]$ von Listen;
Komponente k enthält eine Liste von $n + 1$ Listen $\{l_0, \dots, l_n\}$. Jede der Listen l_i stellt die codierten Variablenwerte von Stufe i dar, die im Fallbeispiel k auftreten.

Die Codierung ist abhängig von dem ***object-operand*** der entsprechenden Stufe und wird wie folgt bestimmt:

Sei i die eine Stufe aus $\{0, \dots, n\}$ mit m Variablen, wobei sich der Wertebereich der Variablen v_j aus o_j Termen $\{t_{j,0}, \dots, t_{j,o_j-1}\}$ zusammensetzt.

1. Sei das i -te Element der Liste ***object-operand*** gleich “AND”.

Dann berechnet sich der Code der Elemente $v_0 = t_{0,w_{v_0}}$ bis $v_m = t_{m,w_{v_m}}$ zu

$$\sum_{c=0}^m w_{v_c} \cdot o_c$$

2. Sei das i -te Element der Liste ***object-operand*** gleich “OR”.

Dann berechnet sich der Code des Elementes $v_j = t_{j,w_{v_j}}$ zu

$$w_{v_j} + \sum_{c=0}^{j-1} o_c$$

Eine Liste l_i setzt sich nun aus Code-Wert-Listen zusammen. Der *Code* (siehe oben) spiegelt die Eigenschaften des entsprechenden Fallbeispiels wider, während der *Wert* bei linguistischen Variablen dem Zugehörigkeitswert entspricht und sonst gleich 1.0 ist.

- (b.) Zur Bestimmung von Gültigkeitswerten nach der in Kapitel 7.1 dargestellten Methode der Verrechnung von Fuzzywerten werden folgende globale Variablen verwendet:

confidences *Array für die Konfidenzen der Regeln*

Array $[0..n-1]$ von Matrizen $[0..p_i-1, 0..p_{i+1}-1]$ von Konfidenzen.

p_i ist dabei die i -te Komponente der Liste **number-of-terms**, d.h. p_i gibt die Zahl verschiedener Terme der Stufe i an.

confidences $[i][x, y]$ repräsentiert die Regel "IF (decode x) THEN (decode y)²⁴", die zu den Regeln der i -ten Regelstufe gehört. Der Eintrag von **confidences** $[i][x, y]$ besteht aus einer Liste (Regelbewertung, Zahl begründender Fallbeispiele).

low-bound Zahl aus $[-1, 1]$, die als untere Schranke für die Regelbewertungen verwendet wird, um eine Reduzierung der Regelmenge zu erreichen (siehe Seite 59).

- (c.) Bei Einbettung von neuronalen Netzen zur Regelbewertung (Kapitel 7.2) werden nachstehende Datenstrukturen eingesetzt:

connections *Array für die Kantengewichte eines neuronalen Netzes*

Array $[0..n-1]$ von Matrizen $[0..p_i-1, 0..p_{i+1}-1]$ von Kantengewichten.

p_i ist dabei die i -te Komponente der Liste **number-of-terms**, d.h. p_i gibt die Zahl verschiedener Terme der Schicht i an²⁵.

connections $[i][v, w]$ repräsentiert die Kante zwischen den Neuronen $v^{[i]}$ und $w^{[i+1]}$ und beinhaltet das zugehörige Kantengewicht.

vertex-layer Array $[0..n]$ von Arrays $[0..p_i-1]$ von Ausgangszuständen.

p_i wie oben.

vertex-layer $[i][w]$ repräsentiert das Neuron w der i -ten Schicht und enthält während der Propagierung der Eingangsinformationen zur Ausgabeschicht die Aktivität dieses Neurons. Nach der Fehlerermittlung an der Ausgabeschicht beinhaltet **vertex-layer** $[i][w]$ eine zweielementige Liste: (Aktivität x_w^i ; Fehler e_w^i)

net-type $\in \{ \text{"vollstaendig"}, \text{"relevant"} \}$

Gibt Aufschluß über den Netzwerktyp. Falls "relevant", werden nur die Kanten im Netzwerk betrachtet, die mit der erzeugbaren Regelmenge übereinstimmen. Default = "relevant".

net-edge-init $\in \{ \text{"random"} \} \cup [-1, 1]$

Bestimmt die Initialgewichte der Netzkanten. Falls "random", werden die Kanten zufällig mit Gewichten aus $(-0.1, 0.1)$ initialisiert. Default = "random".

neuro-bound $\in \{ \text{"normalize"} \} \cup \text{Liste } (l, u); l, u = \text{Gleitpunktzahlen.}$

Art der Umwandlung der Netzkantengewichte in Regelbewertungen. Falls "nor-

²⁴Die Funktion (decode x) liefert für den Code x die entsprechende Variable-Term-Liste.

²⁵Die Stufe i ist im neuronalen Netz äquivalent zur i -ten Netzwerkschicht.

malize”, werden die Gewichte normalisiert, sonst werden die in der Liste angegebenen Schranken zur Abbildung verwendet.

- (d.) Bei der Regelverarbeitung (Kapitel 8) werden nachfolgende globale Variablen benutzt:

rules *Array für alle erzeugten Regeln*

Array [1..n] von Arrays [0.. $p_{i+1}-1$] von Regellisten.

p_{i+1} ist dabei die $(i+1)$ -te Komponente der Liste **number-of-terms**, d.h. p_{i+1} gibt die Zahl verschiedener Terme der Stufe $i+1$ an.

rules[$i+1$][y] enthält eine Liste aller Regeln der $(i+1)$ -ten Regelstufe, die (decode y) als Konklusion aufweisen.

pos-rules *Array für alle gültigen Regeln*

Array [1..n] von Arrays [0.. $p_{i+1}-1$] von Regellisten.

p_{i+1} wie oben.

pos-rules[$i+1$][y] enthält eine Liste aller gültigen Regeln der $(i+1)$ -ten Regelstufe, die (decode y) als Konklusion aufweisen.

results *Array für alle anwendbaren, gültigen Regeln zu einer gegebenen Aufgabenstellung*

Array [1..n] von Arrays [0.. $p_{i+1}-1$] von Regellisten.

p_{i+1} wie oben.

results[$i+1$][y] enthält eine Liste aller für eine bestimmte Aufgabenstellung anwendbaren Regeln der $i+1$ -ten Regelstufe, die (decode y) als Konklusion aufweisen.

new-case Hashtabelle, die die gültige Datenmenge repräsentiert. Die Menge aller Variablen bildet die Menge möglicher Schlüssel.

Diese Hashtabelle beinhaltet zunächst die vorgegebenen Variablenwerte für eine zu lösende Aufgabenstellung. Während der Regelpropagierung werden alle gefolgerten Terme eingetragen bis zur letzten Regelstufe.

Die Einträge von **new-case** beinhalten damit sowohl die Lösung einer Aufgabenstellung, als auch alle Zwischenergebnisse, die bei der Erklärungskomponente Verwendung finden.

distance Gleitpunktzahl aus dem Intervall [0, 1], die die in Kapitel 8 beschriebene Distanz zur “best-of”-Auswahl angibt. Default = 0.1.

- (e.) Für das spezielle Klassifikationsproblem der Rührorganauswahl werden bei der Verwendung der auf Seite 35 beschriebenen Ähnlichkeitsfunktion folgende globale Variablen benutzt:

weight-vector Liste, die die Gewichte der einzelnen Rührermerkmale enthält (Gleitpunktzahlen). Ein Gewicht an der j -ten Position der Liste ist dem Merkmal zugeordnet, das sich an der j -ten Position der Liste **variables-at-layer**[1] befindet (vorausgesetzt, die Merkmale befinden sich auf Stufe 1).

ana-distance Gleitpunktzahl, die die in Kapitel 8 beschriebene Distanz zu den ermittelten Ähnlichkeitswerten bei der Klassifikation liefert. Default = 0.9.

A.2.2 Funktionen

Im Anschluß wird ein Überblick über die erstellten Dateien und die wichtigsten Funktionen gegeben.

(a.) Allgemeine Dateien

“**global-variables.lisp**” Deklaration aller globalen Variablen.

Initialisierung mit NIL.

“**basic-fun.lisp**” Diese Datei beinhaltet die Basisfunktionen, die allgemein Verwendung finden. Diese Funktionen gliedern sich wie folgt:

- Grundlegende Listenoperationen
- Funktionen zur Initialisierung von Arrays und Hashtabellen
- Regelbasierte Funktionen:
 - (**create-rule** *prem concl r.value r.ex.sum*) zur Erzeugung einer Regel der Form
(*r.value r.ex.sum*) (IF *prem*) (THEN *concl*)
 - Funktionen zur Ermittlung einzelner Regelkomponenten (Prämisse, Konklusion, Gültigkeitswert, Anzahl begründender Fallbeispiele)
 - Verknüpfungsoperatoren für Gültigkeitswerte und Zahlen begründender Fälle (**connect-...**)

(b.) Datei zum Einladen und Konvertieren der Datensammlung

“**load-and-convert-data.lisp**”

Grundlegende Funktionen:

- (**load-collection** *file.name*)
 - *file.name* ist ein String, der den Namen der Datei angibt, die die Variablen definitionen und die Datensammlung enthält (Details siehe A.3).

Verarbeitung: Definition aller auf Seite 80 f aufgeführten globalen Variablen gemäß der in *file.name* abgelegten Daten. Konvertierung der einzelnen Fallbeispiele in die für **coded-case** vorgegebene Form anhand der Funktion (**load-examples** *elem*).
- (**gen-fun-code** *code.list len.list op*)
 - *code.list* ist eine Liste von Einzelcodes für jeden Variablenwert einer bestimmten Stufe.
 - *len.list* ist eine Liste, die für jede Variable dieser Stufe die Anzahl der Terme des zugehörigen Wertebereichs angibt.

- *op* ist der Verknüpfungsoperand der entsprechenden Stufe.

Verarbeitung: Ermittlung der Gesamtcodierung gemäß der Code Funktion, die bei der Beschreibung der Datenstruktur **coded-case** angegeben worden ist.

Ergebnis: Ermittelter Gesamtcode.

- (**gen-inv-fun-code** *code len.list op*)
 - *code* ist der Gesamtcode eines Variablenwertes einer bestimmten Stufe.
 - *len.list* ist eine Liste, die für jede Variable dieser Stufe die Anzahl der Terme des zugehörigen Wertebereichs angibt.
 - *op* ist der Verknüpfungsoperand der entsprechenden Stufe.

Verarbeitung: Ermittlung der Liste aller Einzelcodes für jeden Variablenwert (inverse Abbildung zur Funktion (**gen-fun-code** ...)).

Ergebnis: Ermittelte Liste von Einzelcodes.

- (**object-decode** *code ob.list op*)
 - *code* ist der Gesamtcode eines Variablenwertes einer bestimmten Stufe.
 - *ob.list* ist eine Liste aller Variablen dieser Stufe.
 - *op* ist der Verknüpfungsoperand der entsprechenden Stufe.

Ergebnis: Variable-Wert-Liste, die den *code* repräsentiert, falls *op* = “OR”.
Liste von Variable-Wert-Listen, die gemeinsam den *code* repräsentieren, falls *op* = “AND”.

(c.) Dateien zur Anwendung der Fuzzy Logic

“**memb-ship.lisp**” Bestimmung von Zugehörigkeitswerten linguistischer Variablen

Grundlegende Funktion: (**membership** *linguist.var value*)

- *linguist.var* ist der Name einer linguistischen Variablen, deren Auswertung die entsprechende Zugehörigkeitsfunktion liefert.
- *value* ist eine Gleitpunktzahl.

Ergebnis: Liste von Term-Zugehörigkeitswert-Listen, die sich aus *value* für die Variable *linguist.var* ergeben.

“**grav-point**” Defuzzifizierung durch Flächenschwerpunktmethode

Grundlegende Funktion: (**defuzzy** *'linguist.var 't.z.list*)

- *linguist.var* wie oben.
- *t.z.list* ist eine Liste von Term-Zugehörigkeitswert-Listen zu der linguistischen Variablen *linguist.var*.

Verarbeitung: Flächenschwerpunktmethode für linguistische Variablen, deren Zugehörigkeitsfunktionen aus Dreieck- oder Trapezlinienzügen bestehen. Aufteilung der Gesamtfläche in Teilflächen $i = \{1, \dots, r\}$ mit dem Flächeninhalt m_i .

Diese Teilflächen sind trapez- bzw. dreieckförmig, so daß der jeweilige Schwerpunkt (x_i, y_i) leicht berechenbar ist.

Ermittlung des Gesamtflächenschwerpunktes (x_G, y_G) durch den Zusammenhang

$$x_G = \frac{\sum_{i=1}^r m_i \cdot x_i}{\sum_{i=1}^r m_i} \quad \text{und} \quad y_G = \frac{\sum_{i=1}^r m_i \cdot y_i}{\sum_{i=1}^r m_i}$$

Ergebnis: Auflistung einer Liste l und der Position x_G des resultierenden Flächenschwerpunktes. Die Liste l besteht aus Term-Zugehörigkeitswert-Listen, die sich aus der Abbildung (**membership linguist.var x_G**) ergeben.

(d.) Datei zur Regelbewertung anhand neuronaler Netze

“**neuro-net-conf.lisp**” Erzeugung eines neuronalen Netzes mit Einstellung der Kantengewichte anhand des Backpropagation Algorithmus:

- Definition der globalen Variablen *connections* und *vertex-layer*.
Funktion: (**init-gen-net**)
- Möglichkeit für den Benutzer, bereits propagierte Kantengewichte, die in einer Datei abgelegt worden sind, wiederzuverwenden.
- Anwendung des Backpropagation Algorithmus (benutzergesteuert).
- Möglichkeit, die neu eingestellten Kantengewichte in einer Datei abzulegen.

Grundlegende Funktion: (**gen-layer-net-modification**)

(e.) Datei zur Erzeugung einer Regelmenge aus den Netzkantengewichten

“**rule-creation.lisp**”

Grundlegende Funktion: (**rule-creation**)

Verarbeitung: In Abhängigkeit der definierten globalen Variablen *neuro-bound* werden die Netzkantengewichte, die im Array *connections* abgelegt worden sind, zu Regelbewertungen transformiert.

Ergebnis: Das Array *rules* beinhaltet die komplette Regelmenge. Im Array *pos-rules* sind alle gültigen Regeln abgelegt, während *neg-rules* die ungültigen Regeln enthält.

(f.) Datei zur Regelbewertung durch Verrechnung der Fuzzywerte

“**memb-comp-conf.lisp**”

- Erzeugung einer Regelmenge, die im Array *rules* abgelegt wird.
- Ermittlung von Gültigkeitswerten nach der in Kapitel 7.1 beschriebenen Methode.
- Bestimmung der Zahl begründender Fälle zu den einzelnen Regeln.
- Reduzierung der Regelmenge gemäß der in *low-bound* definierten unteren Schranke (*pos-rules*).

Funktion: (**create-valid-rule-set**)

Grundlegende Funktionen: (**init-confid-matrix**),
(**generate-and-modificate-rule-set**)

(g.) Datei zur Regelverarbeitung

“**select.lisp**” Die Regeln werden aufgrund der in *new-case* abgelegten Eingangsinformationen propagiert. Selektion der abgeleiteten Terme durch Defuzzifizierung kombiniert mit der “best-of” Auswahl (Distanz gemäß der globalen Variablen *distance*). Alle so erzeugten Zwischenergebnisse werden in der Hashtabelle *new-case* abgelegt, bis keine Regeln mehr angewendet werden können und das Endergebnis gefolgert worden ist.

Grundlegende Funktion: (**selection**)

Verarbeitung: Der Benutzer wird nach den Variablenwerten einer zu lösenden Aufgabenstellung gefragt. Die Regelpropagierung erfolgt nach der in Kapitel 8 beschriebenen Vorgehensweise.

Ergebnis: Dem Benutzer werden alle Zwischenergebnisse und das Endresultat mit den entsprechenden Gültigkeitswerten angegeben.

(h.) Dateien speziell zur Lösung des Problems der Rührorganauswahl

“**tool-select.lisp**” Diese Datei dient zur Regelverarbeitung der Regeln zur Rührorganauswahl.

Grundlegende Funktion: (**select-object** *object.type c.step*)

- *object.type* gibt den Variablennamen an (*ruehrer* oder Rührermerkmal).
- *c.step* enthält die Nummer der Stufe dieser Variablen.

Verarbeitung: Die Verarbeitung verläuft analog zur Funktion (**selection**). Der Unterschied liegt in der Einbettung einer Heuristik in Abhängigkeit von der Rühraufgabe. Sind die Grundrühraufgaben *Suspendieren*, *Emulgieren* oder *Begasen* Bestandteil der Aufgabe, so wird die Bedeutung der allgemeinen Grundaufgaben *Homogenisieren* um den Faktor 0.2 und *Wärmeübergang* um den Faktor 0.5 reduziert (siehe (**compute-total-heuristic-val ...**)).

Desweiteren wird für jede erhaltene Schlußfolgerung getestet, ob diese alle verlangten Grundrühraufgaben erfüllt oder allein einen Teil der Aufgaben. In diesem Fall wird die Schlußfolgerung aus der Lösungsmenge entfernt (siehe (**auf-filter ...**)).

Ergebnis: Eintrag der klassifizierten Objekte in *new-case*.

“**ana-function.lisp**” Diese Datei enthält die Ähnlichkeitsfunktion, die die ermittelten Rührermerkmale auf einen Rührer abbildet. Der Benutzer kann den Gewichtungsvektor *weight-vector* und die Distanz *ana-distance* vorgeben. Zur Ermittlung des Ähnlichkeitswertes zum Parameter *Rühraufgabe* werden zwei unterschiedliche Funktionen bereitgestellt:

(**equ-aufgabe** *computed.val defined.val*) und
(**simple-equ-aufgabe** *computed.val defined.val*).

Einerseits werden die in den Rührerdefinitionen gegebenen Beiwerte der Rühraufgaben verrechnet, andererseits werden Binärwerte für die Erfüllbarkeit der Rühraufgaben zugrunde gelegt (Erläuterungen siehe Seite 0a). Das resultierende Rührorgan wird dann in der Hashtabelle **new-case** abgelegt.

(i.) Datei aller Benutzerfunktionen zum regelbasierten Ansatz

“app-functions.lisp”

Benutzerfunktionen:

- (**load-data**) Datensammlung laden:
 - Erfragung des Dateinamens, der die Datensammlung enthält.
 - Laden und konvertieren der Datensammlung durch die Funktion (**load-collection**).
 - Initialisierung der Datenstrukturen zur Regelbewertung anhand der Funktionen (**init-gen-net**) und (**init-confid-matrix**).
- (**generate-rule-confidences**) Erzeugung einer Regelmenge, deren Regelbewertungen durch Verrechnung von Fuzzywerten ermittelt werden und Bestimmung der Zahl begründender Fallbeispiele. Erfragung des Wertes von **low-bound** zur Bestimmung der Menge gültiger Regeln im Array **pos-rules**.
- (**neuro-net-propagation**) Modifikation der Netzkanten anhand des Backpropagation Algorithmus:
 - Falls bereits abgelegte Kantenbewertungen verwendet werden sollen, wird der Name der entsprechenden Datei erfragt.
 - Anschließend wird die Zahl der Zyklen erfragt, die durchgeführt werden sollen (= Anzahl Durchläufe des Backpropagation Algorithmus mit vollständiger Beispielmenge).
 - Während der Propagierung werden die Nummern der gerade betrachteten Fallbeispiele angezeigt.
 - Nach jedem Durchlauf werden die Kantenbewertungen ausgegeben, die alle Knoten der Inputebene mit dem ersten Knoten der nachfolgenden Ebene verbinden.
 - Der Algorithmus wird nach Ausführung der vorgegebenen Zyklenzahl abgebrochen, und der Name der Datei wird erfragt, in dem die gewonnenen Kantenbewertungen abgelegt werden sollen. Die Gewichte werden darin gespeichert.
 - Erfragung der Transformationsmethode der Kantengewichte zu Regelbewertungen. Folgende Angaben sind möglich:
 - * “normalize” zur Normalisierung der Gewichte.

- * Liste (l u); l = untere Schranke, u = obere Schranke, $l < u$ falls feste Schranken verwendet werden sollen.
- Erzeugung der Regelmengen **rules**, **pos-rules** und **neg-rules**.
- **(show-rule-set rules)** Ausgabe der Regelmenge *rules* mit Bewertungen.
Bei Aufruf der Funktion muß ein Parameter angegeben werden:
 - **(show-rule-set *rules*)**: Angabe aller erzeugten Regeln.
 - **(show-rule-set *pos-rules*)**: Angabe aller gültigen Regeln.
 - **(show-rule-set *neg-rules*)**: Angabe aller ungültigen Regeln.
- **(selection)** Klassifikation anhand einer Aufgabenstellung:
 - Erfragung der Werte aller Eingabeparameter
 - Klassifikation der gesuchten Parameter mit Hilfe der Regelmenge (forward reasoning)
 - Ausgabe aller Einträge der Hashtabelle **new-case**
- **(tool-selection)** Spezielle Lösung des Rührproblems. Regelverarbeitung anhand der Funktionen aus “tool-select.lisp”.
Ablauf prinzipiell wie **(select)**. Unterschied:
Falls es sich um eine tiefe Modellierung des Rührzusammenhangs handelt, wird die Ähnlichkeitsfunktion aus “ana-function.lisp” eingesetzt.
- **(show-results)** Ausgabe der Parameterwerte einer Stufe:
 - Erfragung der Nummer der Stufe, deren Parameterwerte angezeigt werden sollen (0 = Inputebene).
 - Angabe der ermittelten Parameterwerte dieser Stufe.
- **(show-applied-rules)** Angabe der anwendbaren Regeln:
Es werden alle Regeln zu einer vorgegebenen KONKLUSION angezeigt.
 - Erfragung des Variablennamens, zu dem die Regeln angezeigt werden sollen (Variable erscheint als Schlußfolgerung).
 - Erfragung des relevanten Variablenwertes. Bei Eingabe “ALLE” werden alle Regeln ausgegeben, in denen die Variable vorkommt.
 - Ausgabe der Regel, der Regelbewertung, des Gültigkeitswertes der Prämisse und evt. Angabe der Zahl der begründenden Beispiele.

figure=benutzer.ps

Abb.A.1: Benutzerfunktionen im Ablaufdiagramm

A.3 Benutzungshinweise

A.3.1 Darstellung der Datensammlung

SYSTEM $\langle \text{Systemname} \rangle$ IS

$\langle \text{Anzahl Beispiele} \rangle$

OBJECTS

$\langle \text{Objekttyp} \rangle \langle \text{Objektname } 1 \rangle \langle \text{Abhängigkeit} \rangle$

$\langle \text{Definition} \rangle$

...

$\langle \text{Objekttyp} \rangle \langle \text{Objektname } k \rangle \langle \text{Abhängigkeit} \rangle$

$\langle \text{Definition} \rangle$)

RULES

(IF $\langle \text{Prämisse } 1 \rangle$ THEN $\langle \text{Konklusion } 1 \rangle$)

...

(IF $\langle \text{Prämisse } r \rangle$ THEN $\langle \text{Konklusion } r \rangle$)

CASES

$\langle \text{Beispiel } 1 \rangle \dots \langle \text{Beispiel } n \rangle$

END

$\langle \text{Systemname} \rangle$	= Zeichenkette zur Bezeichnung des Systems
$\langle \text{Anzahl Beispiele} \rangle$	= Anzahl der unten aufgeführten Fallbeispiele n
$\langle \text{Objekttyp} \rangle$	$\in \{\text{ENUMERATE, FUZZY}\}$
$\langle \text{Objektname } i \rangle$	= Name des Parameters $i \in \{1, \dots, k\}$
$\langle \text{Abhängigkeit} \rangle$	$\in \{\text{DEP, INDEP}\}$
$\langle \text{Definition} \rangle$	= $(\text{elem}_1, \dots, \text{elem}_s)$, falls $\langle \text{Objekttyp} \rangle$ gleich ENUMERATE = $((\langle \text{Maßstab} \rangle \langle \text{Einheit} \rangle) (\langle \text{elem}_1 \rangle \langle \text{Punktliste} \rangle) \dots$ $(\langle \text{elem}_s \rangle \langle \text{Punktliste} \rangle)$ falls $\langle \text{Objekttyp} \rangle$ gleich FUZZY
$\langle \text{Maßstab} \rangle$	$\in \{\text{"LOG", "LB", "EINS.ZU.EINS"}\}$
$\langle \text{Punktliste} \rangle$	= $\langle \text{Punkt}_1 \rangle \dots \langle \text{Punkt}_p \rangle$
$\langle \text{Punkt} \rangle$	= (x-pos y-pos)
$\langle \text{Prämisse } i \rangle$	= $\langle \text{Objektname } p_{i,1} \rangle \langle \text{Operator } p_i \rangle \dots \langle \text{Operator } p_i \rangle \langle \text{Objektname } p_{i,n} \rangle$
$\langle \text{Konklusion } i \rangle$	= $\langle \text{Objektname } k_{i,1} \rangle \langle \text{Operator } k_i \rangle \dots \langle \text{Operator } k_i \rangle \langle \text{Objektname } k_{i,n} \rangle$
$\langle \text{Operator } j \rangle$	$\in \{\text{"OR", "AND"}\}$
$\langle \text{Beispiel } i \rangle$	= $\langle \text{Code } 1 \rangle \dots \langle \text{Code } k \rangle$
$\langle \text{Code } i \rangle$	= $\langle \text{Brcode} \rangle$, falls $\langle \text{Objekttyp } i \rangle$ gleich ENUMERATE

$\langle \text{Brcode} \rangle$ NUMBER, falls $\langle \text{Objektyp } i \rangle$ gleich FUZZY
 = 0 oder 1 für jeden Objektwert

A.3.2 Beispiel

Die Datensammlung zur Darstellung des Zusammenhangs
 (VISKOSITÄT \times AUFGABE) \rightarrow RÜHRORGAN
 mit einer Zahl von 100 Beispielen sieht folgendermaßen aus
 (Zeilenumbrüche und Leerzeichen sind beliebig setzbar):

SYSTEM rührerauswahl IS

100

OBJECTS

FUZZY visko DEP

((“LOG” “mPas”)

 (“SEHR.NIEDRIG” (0.0001 1) (10 1) (100 0))

 (“NIEDRIG” (10 0) (100 1) (1000 0))

 (“MITTEL” (100 0) (1000 1) (10000 0))

 (“HOCH” (1000 0) (10000 1) (100000 0))

 (“SEHR.HOCH” (10000 0) (1000001)(1000000 0)))

ENUMERATE aufgabe INDEP

(“HOMO” “SUSP” “WAERME” “BEGASEN” “EMUL”)

ENUMERATE ruehrer INDEP

(“PROPELLER” “SCHRAEGBLATT” “SCHEIBE” “SIGMA” “KREUZ” “TRAPEZ”
 “MIG” “ANKER-COMBI” “ALPHA” “SONDERFORM”)

RULES

(IF aufgabe “AND” visko THEN ruehrer)

CASES

100 1 0 0 0 0 1 0 0 0 0 0 0 0 0

50000 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0

⋮

END

A.3.3 Vorliegende Datensammlungen

data/ruehr-collection Datensammlung für direkte Rührerauswahl

data/i-o-conn Kantengewichte bei der direkten Rührerauswahl

data/para-collection Datensammlung zur Bestimmung der Rührermerkmale

data/i-p-conn Kantengewichte bei der Bestimmung der Rührermerkmale

data/para-ruehr-collection Datensammlung des erweiterten Ansatzes (3 Ebenen)

data/long-i-p-conn Gewichte der Kanten zwischen Ebenen 0 und 1

data/long-p-o-conn Gewichte der Kanten zwischen Ebenen 1 und 2

data/performance-collection Datensammlung zur Bestimmung der spezifischen Leistung (AUFGABE \times VISKOSITÄT \times RÜHRER \rightarrow SPEZIFISCHE-LEISTUNG)

data/i-perm-conn Kantengewichte bei der Bestimmung des Strömungsbereichs