

DIE REPRÄSENTATION VON VAGEM WISSEN IN NATÜRLICHSPRACHLICHEN SYSTEMEN DER KÜNSTLICHEN INTELLIGENZ

WOLFGANG WAHLSTER

|f|-HH-B-38/77

Juli 1977
Institut für Informatik
Universität Hamburg
Schlüterstraße 70
D-2000 Hamburg 13

ABSTRACT

Besides its ambiguity the vagueness inherent in natural language discourse is one of the central problems involved in the design of natural language AI systems. Whereas problems of ambiguity have been taken into account in most natural language systems, the issues of fuzziness of natural language and human knowledge have been largely ignored.

The aim of this report is to provide the first complete survey of the various, somewhat isolated approaches to the representation and processing of fuzzy knowledge within the field of Artificial Intelligence and to suggest new solutions and approaches to these problems in several representation languages.

The paper includes a discussion of the theoretical aspects of the problem (formal methods of fuzzy-set theory and fuzzy logic). Problems of fuzziness in the various components of a natural language system are also investigated.

Requirements for the representation language and the inference system of an understanding system are derived from a review of relevant philosophical, psychological and linguistic investigations of the problems of vagueness.

It is shown how fuzzy inference rules and the semantics of vague natural language terms can be represented and interpreted in the following representation languages:

- extended Predicate Calculus
- production systems
- AI programming languages
- semantic nets
- frame systems

Solutions to the following problems are proposed: Zadeh's fuzzy linguistic variables in extended semantic nets, reference set inferences for the processing of relative adjectives, fuzzy inference rules in generalized implication graphs, linguistic hedges as deduce procedures in the AI programming language FUZZY and matching modes on 'features' and 'meta-descriptions' in KRL-Frames, predicate-limiting adverbials as 'perspectives' in KRL. The report also deals with objections to the use of fuzzy logic in AI systems.

ZUSAMMENFASSUNG

Neben der Mehrdeutigkeit ist die Vagheit der natürlichen Sprache eines der zentralen Probleme beim Entwurf natürlichsprachlicher Systeme. Während die Mehrdeutigkeit in den meisten natürlichsprachlichen Systemen berücksichtigt wurde, ist das Problem der Vagheit der natürlichen Sprache und des menschlichen Wissens meist unberücksichtigt geblieben.

Ziel der Arbeit ist es, den ersten vollständigen Überblick über die verschiedenen, bisher isolierten Ansätze zur Repräsentation und Verarbeitung von vagem Wissen im Rahmen der Künstlichen Intelligenz zu geben und vor allem neue, weitergehende Lösungsvorschläge in unterschiedlichen Repräsentationssprachen zu entwickeln.

Neben einer ausführlichen Einführung in das behandelte Gebiet werden die theoretischen (formale Methoden der Theorie der 'fuzzy sets' und der 'fuzzy logic') und die anwendungsbezogenen (Vagheit in den einzelnen Komponenten eines natürlichsprachlichen Systems) Aspekte des Themas diskutiert.

Aus philosophischen, psychologischen und linguistischen Untersuchungen zum Problemkreis Vagheit werden Forderungen an die Repräsentationssprache und das Inferenzsystem eines natürlichsprachlichen Systems hergeleitet. Es wird gezeigt, wie vage Inferenzregeln und die Semantik vager sprachlicher Ausdrücke in den Repräsentationssprachen

- erweiterter Prädikatenkalkül
- Produktionensysteme
- sehr hohe Programmiersprachen
- semantische Netze
- 'Frames'

dargestellt und interpretiert werden können.

Es werden Lösungen u.a. zu folgenden Problemen vorgestellt: Zadehs unscharfe, linguistische Variablen in erweiterten semantischen Netzen, Referenzmengeninferenzen zur Verarbeitung relativer Adjektive, vage Inferenzregeln in verallgemeinerten Implikationsgraphen, Linguistische Hecken als FUZZY-Theoreme und Match-Modi über sortierten Deskriptoren in KRL-Frames, Limitierende Adverbiale als 'perspectives' in KRL.

Die Arbeit setzt sich auch kritisch mit den Einwänden gegen eine Verwendung der 'fuzzy logic' in Systemen der Künstlichen Intelligenz auseinander.

HINWEISE AN DEN LESER

Der eilige Leser kann zunächst große Teile von Kapitel 2 überspringen und sich mit der Durchsicht der Zusammenfassungen 2.1.4, 2.2.4, 2.3.4 begnügen.

Abschnitt 3.1 kann von einem Leser, der über fundierte Kenntnisse der Theorie der 'fuzzy sets' verfügt, übersprungen werden.

Ein Leser, der sich bereits mit dem Kerngebiet 'Repräsentation von Wissen' und dem Forschungsschwerpunkt 'Natürlichsprachliche Systeme' im Rahmen der Künstlichen Intelligenz intensiv beschäftigt hat, kann sich auf eine flüchtige Durchsicht der Abschnitte 1.1, 1.2, 4.1.1, 4.2.1, 4.3.1, 4.4.1 und 4.6.1 beschränken.

Umgekehrt erhält der Leser, der sich nur über die allgemeine Fragestellung der Repräsentation von Wissen in natürlichsprachlichen Systemen der Künstlichen Intelligenz informieren möchte, einen knappen Überblick, indem er lediglich die Abschnitte 1.1, 4.1.1, 4.2.1, 4.3.1, 4.4.1 und 4.6.1 durcharbeitet.

Ein Leser, der sich nach Kapitel 1 noch eine weitere, praktische Motivation für das in der Arbeit behandelte Thema wünscht, sollte die Lektüre von Abschnitt 5.2 vorverlegen.

Wichtige neue Einsichten und Ergebnisse der Arbeit findet der Leser in den Abschnitten 1.3, 2.1.4, 2.2.4, 2.3.4, 3.2, 4.1.3, 4.3.2, 4.4.2.2, 4.4.3, 4.4.5.1, 4.6.2 und 5.1.

Alle Definitionen und Beispiele sind durch Hervorhebung am Textrand deutlich gekennzeichnet. Mit Großbuchstaben geschriebene Autorennamen verweisen auf Literaturangaben im Literaturverzeichnis.

Ein Teil dieser Arbeit wurde während der Arbeitstagung der GI-Fachgruppe Künstliche Intelligenz in Bad Honnef, die vom 7.-11. März 1977 stattfand, vorgetragen.

Diese Arbeit entstand als Diplomarbeit am Institut für Informatik. Herrn Dr. P. Schefe und Herrn Prof. Dr. W. Brauer danke ich für die Betreuung und Begutachtung der Arbeit. Außerdem danke ich den Herren Dipl.-Inform. H. Boley, Prof. Dr. K. Brunnstein, Prof. Dr. W. v.HAHN, W. Hoepfner, Prof. Dr. H.-H. Nagel, Dr. J. Schmidt und Dr. T. Wittig sowie den Teilnehmern der Linguistik-Informatik-Psychologie-Treffen und den Mitgliedern der Fachgruppe Künstliche Intelligenz für Hinweise, Kritik, Anregungen und Unterstützung

INHALTSVERZEICHNIS

1	EINFÜHRUNG IN DEN PROBLEMBEREICH UND ABGRENZUNG DES THEMAS	7
1.1	KÜNSTLICHE INTELLIGENZ (KI)	7
1.1.1	KERNGEBIETE UND FORSCHUNGSSCHWERPUNKTE DER KI	7
1.1.2	DIE REPRÄSENTATION VON WISSEN ALS ZENTRALES PROBLEM DER KI	9
1.1.3	EINIGE GRUNDBEGRIFFE DER REPRÄSENTATIONSTHEORIE	10
1.2	NATÜRLICHSPRACHLICHE SYSTEME (NSS)	13
1.2.1	NATÜRLICHSPRACHLICHE SYSTEME UND FRAGEANTWORT-SYSTEME: DEFINITIONEN UND BEISPIELE	13
1.2.2	PRAKTISCHE EINSATZMÖGLICHKEITEN VON NSS	14
1.2.3	AKTUELLE FORSCHUNGSPROBLEME AUF DEM GEBIET DER NSS	17
1.3	VAGES WISSEN	18
1.3.1	MEHRDEUTIGKEIT UND VAGHEIT ALS DIE BEIDEN ZENTRALEN PROBLEME IN NSS	18
1.3.2	KOMMUNIKATIVE FUNKTIONEN VAGER FORMULIERUNGEN	20
1.3.3	VAGES WISSEN IN DEN EINZELNEN KOMPONENTEN EINES NSS	23
2	UNTERSUCHUNGEN BENACHBARTER DISZIPLINEN ZUM PROBLEMKREIS 'VAGHEIT'	26
2.1	VAGHEIT AUS DER SICHT DER PHILOSOPHIE	26
2.1.1	DAS PARADOXON VON DER KLEINEN ZAHL	26
2.1.2	DEFINITION UND BEURTEILUNG VON VAGHEIT ALS SEMANTISCHES PROBLEM	27
2.1.3	WICHTIGE ERGEBNISSE DER PHILOSOPHISCHEN ARBEITEN ÜBER VAGHEIT	28
2.1.4	ZUSAMMENFASSUNG	29

2.2	VAGHEIT AUS DER SICHT DER PSYCHOLOGIE	30
2.2.1	EMPIRISCHE BELEGE FÜR DIE VAGHEIT SPRACHLICHER KATEGORIEN	30
2.2.2	AUSGEZEICHNETE ELEMENTE IN UNSCHARFEN KATEGORIEN: DIE PROTOTYPENTHEORIE	31
2.2.3	ZUR STRUKTUR VAGER BEGRIFFE: DIE FAMILIEN- ÄHNLICHKEITS-HYPOTHESE	32
2.2.4	ZUSAMMENFASSUNG	33
2.3	VAGHEIT AUS DER SICHT DER LINGUISTIK	34
2.3.1	ZUR MODIFIKATION VAGER PRÄDIKATIONEN DURCH LINGUISTISCHE HECKEN	34
2.3.2	LINGUISTISCHE UNTERSUCHUNGEN VAGER AUS- DRÜCKE AM BEISPIEL DER RELATIVEN ADJEKTIVE	35
2.3.3	INHALTSSEMANTISCHE UND REFERENZSEMANTISCHE ASPEKTE DER VAGHEIT	38
2.3.4	ZUSAMMENFASSUNG	38
3	DIE THEORIE DER 'FUZZY SETS' UND IHRE ANWENDUNG AUF DIE REPRÄSENTATION VON VAGEM WISSEN	39
3.1	ZUR THEORIE DER 'FUZZY SETS'	39
3.1.1	EINFÜHRENDE ÜBERSICHT	39
3.1.2	GRUNDLEGENDE DEFINITIONEN	41
3.2	DIE FORMALE DARSTELLUNG VON VAGEM WISSEN DURCH 'FUZZY SETS'	43
3.2.1	DIE REPRÄSENTATION DER EXTENSIONALEN BEDEU- TUNG VAGER SPRACHLICHER AUSDRÜCKE	43
3.2.2	DIE DARSTELLUNG VON LINGUISTISCHEN HECKEN	47
3.2.2.1	Linguistische Hecken vom Typ 1	48
3.2.2.2	Linguistische Hecken vom Typ 2	50
3.2.3	DAS KONZEPT DER LINGUISTISCHEN VARIABLEN	51
3.3	PSYCHOLINGUISTISCHE ASPEKTE DER 'FUZZY SETS'	53
4	ZUR FORMALEN DARSTELLUNG VON VAGEM WISSEN IN VERSCHIEDENEN REPRÄSENTATIONSSPRACHEN DER KI56	
4.1	DIE DARSTELLUNG VON VAGEM WISSEN IN DEN AN DER PRÄDIKATENLOGIK ORIENTIERTEN REPRÄSENTATIONS- SPRACHEN	56
4.1.1	EINFÜHRENDE ÜBERSICHT	56
4.1.2	MEHRWERTIGE LOGIK UND 'FUZZY LOGIC'	57
4.1.3	DIE REPRÄSENTATION VAGER INFERENZREGELN UND IHRE VERARBEITUNG DURCH 'FUZZY RESOLUTION'	58

4.1.4	DIE PRÄDIKATENLOGISCHE DARSTELLUNG DER INTENSIONALEN BEDEUTUNG RELATIVER ADJEKTIVE UND VERGLEICHE	61
4.2	DIE REPRÄSENTATION VON VAGEM WISSEN IN PRODUKTIONENSYSTEMEN	62
4.2.1	EINFÜHRENDE ÜBERSICHT	62
4.2.2	DIE DARSTELLUNG VAGER INFERENZREGELN IM CF-MODELL	64
4.2.3	INFERENZNETZE ZUR REPRÄSENTATION VAGER SCHLUSSREGELN	67
4.3	DIE REPRÄSENTATION VON VAGEM WISSEN IN KI-PROGRAMMIERSPRACHEN	70
4.3.1	EINFÜHRENDE ÜBERSICHT	70
4.3.2	DIE MEHRWERTIGE PLANNER-ARTIGE PROGRAMMIERSPRACHE FUZZY	71
4.3.2.1	KI-Programmiersprachen mit mehrwertiger Logik	71
4.3.2.2	Die Repräsentation und Verarbeitung von vagen sprachlichen Ausdrücken in FUZZY	72
4.3.2.3	Die Darstellung vager Inferenzregeln in FUZZY	75
4.3.3	DEDUCE: VAGE PRÄFERENZBEDINGUNGEN IN EINER ANFRAGESPRACHE FÜR RELATIONALE DATENBASEN	76
4.3.4	REPRÄSENTATIONSKONSTRUKTIONEN FÜR VAGES WISSEN IN EINER QA4-ERWEITERUNG	78
4.4	DIE REPRÄSENTATION VON VAGEM WISSEN IN SEMANTISCHEN NETZEN (SN)	80
4.4.1	EINFÜHRENDE ÜBERSICHT	80
4.4.2	ELEMENTARE MODIFIKATIONEN DER KANTEN UND KNOTEN EINES SN	81
4.4.2.1	Gültigkeitsmarkierungen in Quillians SN	82
4.4.2.2	F-Graphen in propositions- und objektzentrierter Darstellung	82
4.4.2.3	I- und U-Markierungen im SCHOLAR-System	83
4.4.2.4	Unvollständige Mengen	84
4.4.3	DIE DARSTELLUNG DER EXTENSIONALEN BEDEUTUNG VAGER KONZEPTE IN SN	85
4.4.3.1	Linguistische Variablen in erweiterten SN	85
4.4.3.2	Die Referenzmengeninferenz in einer einfachen Generalisierungshierarchie	87
4.4.3.3	Die Repräsentation von Wissen über die Referenzmengen relativer Adjektive	88

4.4.4	DIE DARSTELLUNG DER INTENSIONALEN BEDEUTUNG VAGER KONZEPTE IN VERSCHIEDENEN FORMEN VON SN	90
4.4.4.1	Relative Skalen in Schanks Konzeptuelle-Dependenz-Graphen	90
4.4.4.2	Die Repräsentation von Referenzmengen in Schuberts Formalismus	91
4.4.4.3	Die Modifikation vager Konzepte in gerichteten rekursiven Bewertungsknoten-Hypergraphen	92
4.4.5	DIE DARSTELLUNG VON VAGEN INFERENZREGELN IN SN	94
4.4.5.1	Verallgemeinerte Implikationsgraphen	94
4.4.5.2	Die 'Davon-wüßte-ich'-Inferenz als Beispiel einer komplexen Schlußregel	95
4.5	EIN VERGLEICH VON VIER REPRÄSENTATIONSSPRACHEN AM BEISPIEL EINER VAGEN INFERENZ	96
4.6	DIE REPRÄSENTATION VON VAGEM WISSEN IN DER REPRÄSENTATIONSSPRACHE 'FRAMES'	97
4.6.1	EINFÜHRENDE ÜBERSICHT	97
4.6.2	DIE DARSTELLUNG VON VAGEM WISSEN IN KRL	97
4.6.2.1	Das Framekonzept von KRL und einige Erweiterungen	97
4.6.2.2	Meta-Merkmale und Hecken vom Typ 2	100
4.6.2.3	Die Darstellung der extensionalen Bedeutung vager Konzepte durch objekt-bezogene Prozedurdefinitionen	101
4.6.3	MÖGLICHKEITEN UND PROBLEME DES 'FUZZY MATCHING'	104
5	ABSCHLIESSENDE ÜBERSICHTEN	106
5.1	ZUR KRITIK AN DER VERWENDUNG DER 'FUZZY LOGIC' IN KI-SYSTEMEN	106
5.2	DIE REPRÄSENTATION VON VAGEM WISSEN IN BESTEHENDEN NSS: PERFORMANZ UND ANWENDUNG DER SYSTEME	109
	LITERATURVERZEICHNIS	112

1. EINFÜHRUNG IN DEN PROBLEMBEREICH UND ABGRENZUNG DES THEMAS

1.1. KÜNSTLICHE INTELLIGENZ

Zunächst wird das Thema der vorliegenden Arbeit in den größeren Rahmen der Forschung auf dem Gebiet der Künstlichen Intelligenz (Abk.: KI, engl.: Artificial Intelligence) eingebettet. Nach einem allgemeinen Überblick zu den Fragestellungen und Kerngebieten der KI werden einige Grundbegriffe des Forschungsschwerpunktes 'Repräsentation von Wissen' eingeführt.

1.1.1. KERNGEBIETE UND FORSCHUNGSSCHWERPUNKTE DER KI

Die Künstliche Intelligenz kann als ein Teilgebiet der Informatik betrachtet werden. Während sich die Informatik ganz allgemein in mit der systematischen Verarbeitung von Information mit Hilfe von Digitalrechnern beschäftigt, werden in der Künstlichen Intelligenz speziell solche Informationsverarbeitungsprozesse untersucht, die in besonders starkem Maße Strategien menschlicher Intelligenz voraussetzen.

In diese Richtung tendieren auch sehr schlichte Definitionen der Künstlichen Intelligenz wie

The term Artificial Intelligence is widely used as the name of the branch of computer science that studies how to make computers smarter. (RAPHAEL76, S.2)

Eine stärker differenzierende Betrachtung zeigt, daß es mindestens zwei unterschiedliche Motivationen zur Forschung im Bereich KI gibt:

- eine technische Motivation: Entwicklung von Programmen, die Aufgaben übernehmen, von denen gegenwärtig gesagt wird, sie erforderten Intelligenz.¹
- eine theoretische Motivation: KI als Teil einer Grundlagenwissenschaft, die als 'cognitive science' (vgl. BOBROW/COLLINS75) bezeichnet wird und mit Methoden der Linguistik, Psychologie und KI zur wissenschaftlichen Beschreibung und Erklärung kommunikativer und kognitiver Prozesse beitragen soll. Die Methodik der KI unterscheidet sich u.a. von der traditionellen linguistischen oder psychologischen Vorgehensweise dadurch, daß sie ausschließlich Prozeßmodelle entwickelt, welche - als Programm codiert - durch Ausführung auf einer Rechenanlage getestet und verbessert werden.

Während die technische Motivation zur Entwicklung von Leistungsmodellen führt, in denen das Ein/Ausgabeverhalten im Vordergrund steht, wird bei einer theoretischen Motivation eher die Simulation kognitiver Prozesse durch Funktionsmodelle angestrebt, die stets Analogien zwischen der Arbeitsweise des Modells und der des nachgebildeten Originals enthalten.

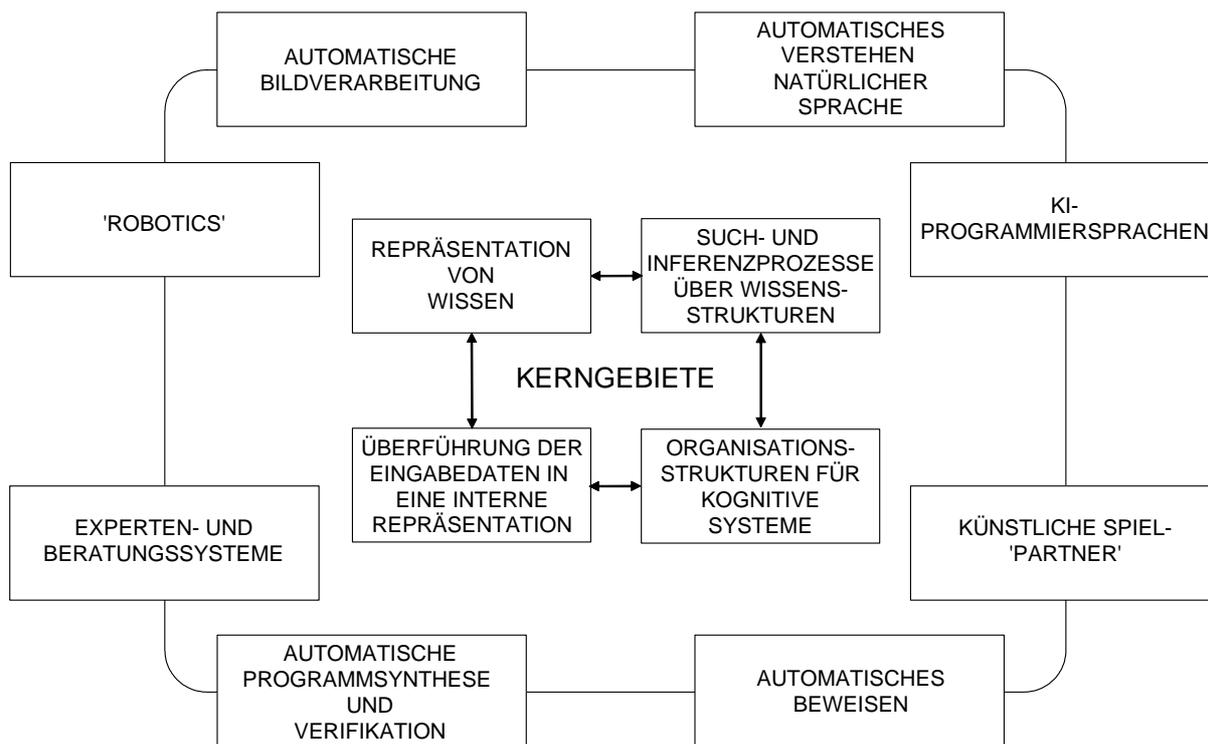
Auf eine wissenschaftstheoretische Diskussion der unterschiedlichen Ansätze und Definitionen der Künstlichen Intelligenz wird hier verzichtet (vgl. NEWELL70, SCHEFE73, BOLEY76a).

Da es eine Reihe von Einführungen und Lehrbüchern zur KI gibt (z.B. NILSSON71, JACKSON74, HUNT75, ITZINGER76, RAPHAEL76, WINSTON77) und der Forschungsstand in Übersichtsartikeln (z.B. NILSSON74, WINOGRAD74, CHANDRASEKARAN75, HART76) und Aufsatzsammlungen (vgl. Literaturverzeichnis) gut dokumentiert ist, werden Geschichte, Grundlagen und bisherige Ergebnisse der KI-Forschung in dieser Arbeit nicht auf breiter Basis zusammengefaßt.

Fig. 1 enthält einen Überblick zu Kerngebieten und Forschungsschwerpunkten der KI. Vier Fragestellungen stehen im Zentrum der Forschung zur KI:

¹Dieser Bezug auf Urteile von Informanten kann als Operationalisierung des sonst vagen Begriffs der Intelligenz aufgefaßt werden.

- Durch welche Prozesse lassen sich z.B. in visueller oder akustischer Form eingegebene Nachrichten in eine interne, formale Repräsentation übersetzen, die eine effektive maschinelle Verarbeitung möglich macht ?
- Welche Datenstrukturen sind zur Repräsentation von dem zur Verarbeitung der Nachrichten notwendigen Wissen geeignet ?
- Welche Such- und Inferenzprozesse sind zur Verarbeitung der Nachrichten notwendig, und welche Algorithmen und Strategien liegen diesen Prozessen zugrunde ?
- Wie sind komplexe kognitive Systeme zu organisieren ?



FIGUR 1

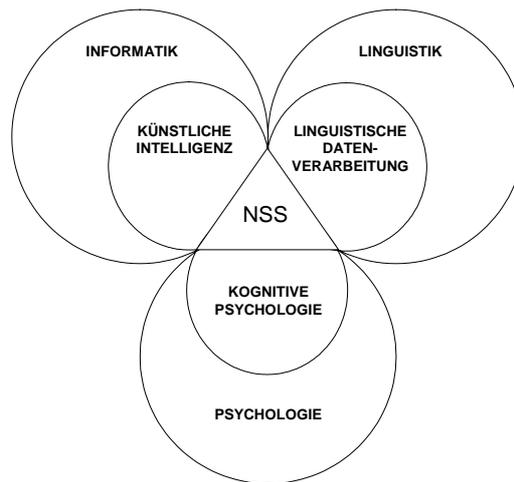
Die Konstruktion natürlichsprachlicher Systeme (Abk.: NSS) ist eine der Aufgaben im Bereich der KI, deren Bearbeitung sich trotz intensiver Forschung als besonders schwierig erwiesen hat:

One of the significant discoveries of AI has been how computationally difficult are the elementary tasks of vision, language, and common sense reasoning which we perform continually in the course of our everyday activities. (KUIPERS75, S.152)

Die Forschung über NSS steht in so enger Beziehung zu den anderen in Fig. 1 dargestellten Gebieten, daß LeFavre behauptet:

Clearly, to 'solve' the problem of natural language is to solve the problem of artificial intelligence.(LEFAIVRE74, S.133)

Die Aufgabenstellung des Forschungsschwerpunktes 'Automatisches Verstehen natürlicher Sprache', in dem die vorliegende Arbeit anzusiedeln ist, kann nur interdisziplinär erfolgreich bearbeitet werden (vgl. Fig. 2).



FIGUR 2

1.1.2. DIE REPRÄSENTATION VON WISSEN ALS ZENTRALES PROBLEM DER KI

Obwohl schon 1958 von McCarthy als Problem erkannt, wurde die Frage nach der Repräsentation von Wissen erst in den letzten Jahren im Zusammenhang mit der Implementation größerer Systeme (z.B. SHRDLU, MARGIE, MYCIN, TORUS) und der Entwicklung einer neuen Klasse von Programmiersprachen (vgl. 4.3) intensiv bearbeitet. Auch die im Bereich der theoretischen und systembezogenen Informatik geführte Diskussion über Datenstrukturen hat die Entwicklung einer Repräsentationstheorie positiv beeinflusst (vgl. 4.6).

Gegenwärtig wird das Problem der Repräsentation von Wissen häufig als das zentrale Problem der KI angesehen.² Viele Autoren gehen sogar so weit, die KI mit Hilfe dieser Problemstellung von anderen Gebieten der Informatik abzugrenzen. So definiert z.B. Winograd:

Artificial Intelligence in general might be characterized as the part of computer science concerned with this problem of representation. (WINOGRAD74, S.7)

DEFINITION: Dabei wird Wissen in der KI als eine Ansammlung von Kenntnissen, Erfahrungen und Problemlösungsmethoden verstanden, die den Hintergrund für komplexe Informationsverarbeitungsprozesse bildet. Das Wissen besteht aus Daten über Objekte, Relationen und Prozesse.

Wir gehen hier davon aus, daß jegliches Wissen verbal präsentierbar ist (vgl. WUNDERLICH74, S.55). Da ohne soziale Kooperation der Gewinn neuen Wissens für eine einzelne Person außerordentlich begrenzt wäre, ist der Mensch ständig gezwungen, Wissen zu kommunizieren. Soll Wissen mitteilbar sein, so muß es dargestellt werden.

Obwohl die Wahl einer problemgerechten Darstellung von Daten eine allgemeine Fragestellung innerhalb der Informatik ist (vgl. z.B. WIRTH75, S.15ff), die besonders im Zusammenhang mit dem sog. 'Strukturierten Programmieren' relevant wird, stellt sich das Problem der Repräsentation von Daten bei der Konstruktion kognitiver Systeme verschärft, da solche Systeme auf einen umfangreichen Wissensfundus, der aus vielen unterschiedlich strukturierten Einheiten bestehen kann, zurückgreifen müssen. Raphael unterstreicht die Bedeutung von Wissen für Systeme der KI:

Knowledge seems to be the real basis, or at least an important prerequisite, for understanding and intelligence. Common sense may be merely a matter of having some relevant knowledge, and being able to make use of it. (RAPHAEL76, S.48)

² Man spricht in der KI in diesem Zusammenhang gelegentlich auch von 'knowledge engineering' (vgl. MICHIE74).

So ist z.B. das Verstehen natürlicher Sprache (Abk.: NS) ohne ein gewisses Maß an Allgemeinwissen und Kenntnissen über den besprochenen Gegenstandsbereich nicht möglich. Die Konstruktion von NSS setzt also die operationale und formale, und damit computergerechte Repräsentation von Wissen voraus.

1.1.3. EINIGE GRUNDBEGRIFFE DER REPRÄSENTATIONSTHEORIE

Im folgenden sollen einige Grundbegriffe aus dem Bereich der Repräsentation von Wissen möglichst einfach und direkt definiert werden. Dabei werden verglichen mit den Darstellungen von HAYES74, BOBROW75, BOLEY76a, die in Anlehnung an die Semiotik versuchen, eine exakte Terminologie einzuführen, viele Vereinfachungen und einige Erweiterungen vorgenommen.

DEFINITION: Die Gesamtheit an Wissen, die einem System der KI zur Verfügung steht, nennen wir Wissensbasis. Die Wissensbasis kann oft in einzelne Bereiche, die wir als Wissensquellen (vgl. NASH-WEBBER/BRUCE76, STEELS76) bezeichnen, unterteilt werden. Die Wissensquellen bestehen aus einzelnen Wissensseinheiten.

BEISPIEL: Wissensquellen in NSS: Grammatische Regeln, Lexikon, Faktenwissen, Inferenzregeln, Dialogregeln

DEFINITION: Wissen, das sich auf anderes Wissen innerhalb der Wissensbasis bezieht, nennen wir Metawissen. Es sind im allgemeinen mehrere Meta-Ebenen zu unterscheiden.

BEISPIEL: Angaben zur Unterscheidung von Wissensbereichen, die als erschöpfend betrachtet werden können und solchen, in denen nur partielle Information vorliegt (vgl. 4.4.2.4.).

DEFINITION: Ein Hilfsmittel zur systematischen Darstellung von Wissen bezeichnen wir als Repräsentationssprache (Abk.: RS).³ Wird ein zu repräsentierendes Original in einen bestimmten Ausdruck der unter Umständen auch mehrdimensionalen RS abgebildet, so bezeichnen wir diesen Ausdruck als Repräsentationskonstruktion (Abk.: RK).⁴

BEISPIEL: RS: Semantisches Netz, logische Kalküle, Planner-artige Programmiersprachen
RK: Teilgraph, Axiom, Programm

Jede RK muß den für eine RS gültigen Wohlgeformtheitsbedingungen genügen, die im einfachsten Fall durch eine kontextfreie Grammatik angegeben werden können. Im Vordergrund muß allerdings die Definition der Semantik einer RS stehen. Erst die Angabe einer semantischen Interpretation, wie sie in Modellen der extensionalen oder intensionalen Varianten einer modelltheoretischen Semantik vorgesehen ist (vgl. z.B. WUNDERLICH74, LEWIS74, HAYES74), führt zu einer formal befriedigenden Definition einer RK.

Gegenwärtig wird bei der Einführung neuer RS ähnlich wie in einer frühen Phase der Entwicklung von Programmiersprachen außer bei einigen an logischen Kalkülen orientierten Systemen auf eine formale Definition der Semantik verzichtet. Stattdessen wird die Semantik informell anhand von Beispielen oder Interpretationsalgorithmen eingeführt.⁵ Kennzeichnend für diese Situation ist auch, daß Sprachdefinitionen für diejenigen Programmiersprachen, die im Bereich der KI entwickelt wurden, den Anforderungen, die man heute im allgemeinen an die formale Definition einer Programmiersprache stellt (vgl. z.B. MARCOTTYetal.76), oft nicht genügen.

DEFINITION: Besonders in dialogfähigen NSS existiert häufig eine Wissensquelle, in der das Wissen über die Situation, auf die sich der Dialog bezieht, oder den angesprochenen Sachbereich

³ vgl. BOLEY76a für den Versuch einer formalen Explikation des Begriffs. HAYES74 führt dafür den Begriff 'Schema' ein.

⁴ HAYES74 führt dafür den Begriff 'Konfiguration' ein.

⁵ vgl. auch den Versuch von WOODS75 eine Semantik semantischer Netze zu entwickeln.

dargestellt wird. Eine solche Wissensquelle bezeichnen wir als Weltmodell (syn.: Weltausschnitt). Abhängig von der Anzahl der in einem Weltmodell dargestellten Objekte unterscheidet man zwischen Mikro-, Mini- und Makrowelten. Besteht ein Weltmodell nur aus RK für Objekte und Prädikationen über diesen Objekten, so sprechen wir von einem statischen Weltmodell. Ein dynamisches Weltmodell enthält zusätzlich zu den Angaben in einem statischen Weltmodell Operatoren, die angeben, wie bestehende Relationen dynamisch verändert werden können (vgl. HENDRIX73, S.145).

Wenn wir uns in einem einfachen Modell vorstellen, daß ein DV-System aus mehreren Schichten von dem materiellen Teil - der Rechenanlage - über viele Zwischenebenen bis hin zum Anwenderprogramm besteht, so ergeben sich in dieser Hierarchie abstrakter Maschinen auf vielen Ebenen und Detailstufen Probleme der Repräsentation von Daten. Das Prinzip der konzeptuellen Unsichtbarkeit der niedrigeren Ebenen, das typisch für DV-Systeme ist, erlaubt es uns, hier von der Darstellung des in einer RS auf hohem Niveau codierten Wissens in den unteren Schichten eines DV-Systems zu abstrahieren.

Da KI-Systeme häufig aus mehreren übereinandergeschichteten Interpretern bestehen, muß bei der Betrachtung von RS stets die genaue Trennung der verschiedenen Repräsentationsebenen beachtet werden.

BEISPIEL:

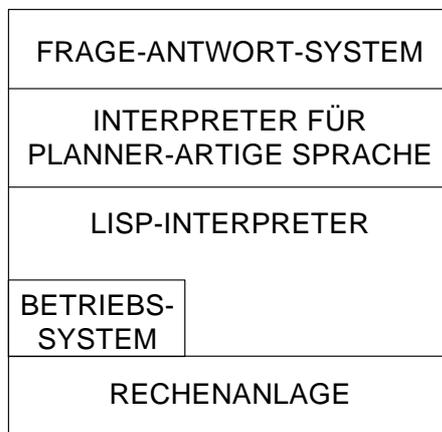


Fig. 3 zeigt in einem einfachen Schichtenmodell ein Frage-Antwort-System, das als Interpreter für die eingegebenen Fragen aufgefaßt werden kann (vgl. WAHLSTER/v.HAHN76). Das Frage-Antwort-System ist in einer Planner-artigen Sprache implementiert, für die ein Interpreter, der in LISP geschrieben ist, existiert.

FIGUR 3

Es gibt Repräsentationssprachen, die sich eher für eine analogische Darstellung und andere, die sich für eine linguistisch-logische Repräsentation anbieten.

DEFINITION: Bestehen die Repräsentationskonstruktionen weitgehend aus Programmen, so sprechen wir von einer prozeduralen Repräsentationssprache. Bestehen sie dagegen lediglich aus Daten, die in einem allgemeinen Interpretationsprozeß verarbeitet werden, so handelt es sich um eine deklarative Repräsentationssprache.

Außerdem lassen sich auf einer niedrigeren Ebene RS danach unterscheiden, ob Programme und Daten homogen (z.B. in LISP) oder auf sehr unterschiedliche Weise (z.B. in PASCAL) dargestellt werden.

Eine weitere Unterscheidung für RS ergibt sich aus dem Kriterium, ob bei der Darstellung zwischen allgemeinen Deduktionsregeln und speziellen Fakten unterschieden wird oder nicht.

Auf eine Darstellung der zahlreichen Kontroversen darüber, welche der genannten Merkmale eine 'ideale' RS haben sollte (vgl. z.B. die Prozedural-Deklarativ-Kontroverse, WINOGRAD75b), muß hier verzichtet werden, da eine differenzierte Betrachtung die unterschiedlichen Auffassungen vom menschlichen Problemlösungsverhalten, die als Grundlagen dieser Kontroversen zu betrachten sind, berücksichtigen müßte und damit den Rahmen dieser Arbeit sprengen würde.

DEFINITION: Eine Organisationsform, in der das Wissen um konzeptuelle Objekte herum gruppiert wird und alle das Objekt betreffenden Informationen in einer größeren Einheit zusammenhängend

dargestellt werden, bezeichnet man als objekt-zentrierte Strukturierung von Wissen (z.B. Eigenschaftslisten von LISP, 'units' in KRL, vgl. 4.6.2.1). Wird das Wissen dagegen nicht um die einzelnen in den darzustellenden Propositionen enthaltenen Objekte gruppiert sondern einfach als eine Menge von Propositionen organisiert, so sprechen wir von einer propositions-zentrierten Strukturierung von Wissen (z.B. assoziative Datenbasen in Planner-artigen Programmiersprachen).

Um den kumulativen Aufbau der Wissensbasis eines KI-Systems durch eine schrittweise Erweiterung zu fördern, sollte das Wissen möglichst modular repräsentiert werden.

Bei der Vielzahl in der KI vorgeschlagener Repräsentationssprachen ist die Frage, unter welchen Randbedingungen eine bestimmte RS einer anderen vorzuziehen ist, besonders relevant. So leisten z.B. DAVIS/KING75 einen nützlichen Beitrag, indem sie zeigen, welche Vor- und Nachteile der Einsatz der RS 'Produktionensysteme' (vgl. 4.2) hat.

Allzuoft wurde bisher auf der Suche nach einer 'universellen' Repräsentationssprache, ähnlich wie in einer bestimmten Phase der Entwicklung von Programmiersprachen oder Datenmodellen für Informationssysteme, auf eine kritische Diskussion der Vor- und Nachteile der vorgeschlagenen Sprachen verzichtet. Heute setzt sich allmählich die Erkenntnis durch, daß es eine 'universelle' Repräsentationssprache nicht gibt, und daß es sogar innerhalb eines einzelnen kognitiven Systems ein Nebeneinander verschiedener RS geben sollte, um die Information in den unterschiedlichen Wissensquellen jeweils adäquat darstellen zu können.

DEFINITION: Werden innerhalb der Wissensbasis eines KI-Systems verschiedene Repräsentationssprachen verwendet, so sprechen wir von einer heterogenen Wissensbasis.

Oft ist es vorteilhaft, bestimmte Wissenseinheiten in der Wissensbasis redundant zu codieren, d.h. die für Datenbanksysteme oft geforderte Redundanzfreiheit der Datenbasis (vgl. z.B. SCHWEPPE75, S.63) aufzugeben.

BEISPIEL: In dem NSS SOPHIE wird die Analysegrammatik sowohl in prozeduraler als auch in deklarativer Form gespeichert. Zunächst wird stets eine Syntaxanalyse der Eingabe von-oben-nach-unten vorgenommen, der die Grammatik in prozeduraler Form zugrunde liegt. Wird die Eingabe syntaktisch nicht akzeptiert, so wird in einem zweiten Durchlauf von-unten-nach-oben eine weniger effiziente Syntaxanalyse mit der Grammatik in deklarativer Form durchgeführt, um eine gute Fehlerdiagnose zu ermöglichen (vgl. BROWN/BURTON75, S.332).

Bobrow und Winograd stellen dazu fest:

We believe that multiple descriptions containing redundant information are used in the human representation system to trade off memory space for computation depth, and that computer systems can take advantage of the same techniques.

(BOBROW/WINOGRAD76, S.5)

DEFINITION: Wird in einer Wissensbasis die gleiche Information evtl. in verschiedenen Repräsentationssprachen mehrfach dargestellt, so sprechen wir von einer multiplen Repräsentation (vgl. auch BOBROW75 S.29).

Abschließend stellen wir fest, daß die Wahl einer Repräsentationssprache stets in bezug auf die über dem darzustellenden Wissen durchzuführenden Operationen zu sehen ist.

1.2. NATÜRLICHSPRACHLICHE SYSTEME

1.2.1. NATÜRLICHSPRACHLICHE SYSTEME UND FRAGE-ANTWORT-SYSTEME: DEFINITIONEN UND BEISPIELE

Für den in dieser Arbeit behandelten Bereich der KI hat sich noch keine einheitliche Terminologie durchsetzen können. Häufig zu findende allgemeine Bezeichnungen für die im Forschungsschwerpunkt 'Automatisches Verstehen natürlicher Sprache' entwickelten Systeme sind 'Natürlichsprachliches System' (engl.: natural language system), 'Natürlichsprachliche Schnittstelle' (engl.: natural language interface), 'Natürlichsprachlicher Front-End-Prozessor' (engl.: natural language front-end), 'Sprachverstehendes System' (engl.: understanding system) und 'Frage-Antwort-System' (engl.: question-answering-system). Obwohl diese Bezeichnungen in der umfangreichen Literatur unterschiedlich benutzt werden, lassen sich folgende Tendenzen erkennen: Die Bezeichnung 'Natürlichsprachliches System' wird sehr allgemein gebraucht und fungiert oft als Oberbegriff. Die Bezeichnungen 'Natürlichsprachliche Schnittstelle' und 'Natürlichsprachlicher Front-End-Prozessor' werden für Systeme gewählt, in denen linguistische Algorithmen einen eher peripheren Teilaspekt des Systems darstellen, während innerhalb von 'Sprachverstehenden Systemen' die linguistischen und damit oft verbundenen kommunikativen und kognitiven Aspekte von zentraler Bedeutung sind.

Um den Sprachgebrauch innerhalb der vorliegenden Arbeit zu regeln, werden im folgenden der allgemeine Begriff 'Natürlichsprachliches System' und der Begriff 'Frage-Antwort-System', dem bereits ein gewisser terminologischer Status zukommt, präzisiert.

DEFINITION: Ein System heißt 'Natürlichsprachliches System' (Abk.: NSS) genau dann, wenn

- eine Teilmenge der in das System eingegebenen oder vom System ausgegebenen Nachrichten natürlichsprachlich codiert sind und
- zur Verarbeitung der Nachrichten u.a. Algorithmen zur syntaktischen oder semantischen Analyse oder Generierung natürlicher Sprache eingesetzt werden.

Die erste Bedingung sichert, daß auch solche Systeme als NSS bezeichnet werden können, die neben einer sprachlichen Ein- oder Ausgabe auch Nachrichten über Spezialperipherie austauschen, wie sie besonders in den Bereichen Mustererkennung, Analyse und Generierung gesprochener Sprache und bei der Konstruktion von Robotern in der KI entwickelt wurde (vgl. z.B. REDDY73). Die zweite Bedingung verhindert, daß Systeme, die lediglich einfache Zeichenkettenoperationen über beliebig und damit auch natürlichsprachlich codierten Texten durchführen (z.B. Texteditoren) als NSS aufgefaßt werden können.

Die Vielfalt der durch den Begriff 'NSS' denotierten Programmsysteme soll durch die in Fig. 4 zusammengestellten Beispiele belegt werden.

ZIELSETZUNG	EINGABEBEISPIEL	AUSGABEBEISPIEL	LITERATURHINWEIS
NS-PROGRAMMIERUNG	NS-CODIERTER ALGORITHMUS	PROGRAMM IN LISP	MILLER/BECKER74
NS-BILDBESCHREIBUNG	DIGITALISIERTES BILD	NS-BILDBESCHREIBUNG	BADLER75
FRAGE-ANTWORT-SYSTEM	NS-ANFRAGE ÜBER EIN SACHGEBIET	ANTWORT	MYLOPOULOSetal.75
'ABSTRACT'-GENERIERUNG	NS-TEXT	KURZE INHALTSANGABE	SCHANKetal.75
'SPEECH UNDERSTANDING'	NS-SATZ ALS PHONETISCHE EINGABE	REAKTION ALS SPIEL-'PARTNER'	REDDY76
PARAPHRASEN-GENERIERUNG	NS-SATZ	PARAPHRASE DES SATZES	SCHANK75
AUTOM. ÜBERSETZUNG	NS-SATZ AUF ENGLISCH	BEDEUTUNGSGLEICHER SATZ AUF FRANZÖSISCH	WILKS73
'ROBOTERSTEUERUNG'	ANWEISUNG IN NS	AKTION EINES 'ROBOTERS'	WINOGRAD72
INFORMATIONSFORMATIERUNG	NS-TEXT	FORMATIERTER TEXTINHALT IN DER DATENBASIS	SAGERetal.77
NS IN GRAPHIK-SYSTEMEN	NS-ANWEISUNG UND EINGABE ÜBER LICHTGRIFFEL	GRAPHISCH-CODIERTE INFORMATION	BROWNetal.77

FIGUR 4

DEFINITION: Ein dialogfähiges natürlichsprachliches System heißt Frage-Antwort-System (Abk.: FAS), wenn es auf natürlichsprachlich codierte Anfragen hin aufgrund von Such- und Inferenzprozessen über einer Wissensbasis natürlichsprachliche Antworten generiert.

Die Grenzen zwischen den so definierten FAS und neueren Informationssystemen sind fließend. Während konventionelle Informationssysteme (vgl. z.B. DATE75) keine linguistischen und deduktiven Fähigkeiten aufwiesen, müssen einigen neueren Informationssystemen diese Merkmale zugesprochen werden (vgl. z.B. CODD74). Der Trend geht dahin, FAS als eine spezielle Klasse von Informationssystemen zu betrachten, die sich durch eine besonders komfortable Anfragesprache und zahlreiche Inferenzmechanismen von konventionellen Informationssystemen unterscheiden.

Da es auf dem Gebiet der NSS einige neuere einführende Übersichten gibt (z.B. DAMERAU76, CHARNIAK/WILKS76, STEELS76), kann hier auf eine zusammenfassende Darstellung der Grundbegriffe und Methoden in diesem Bereich der KI verzichtet werden.

1.2.2. PRAKTISCHE EINSATZMÖGLICHKEITEN VON NSS

Im folgenden soll kurz die Frage untersucht werden, unter welchen Randbedingungen ein Dialog Mensch-DV-System, in natürlicher Sprache anderen Dialogformen vorzuziehen ist (vgl. auch CODD74, MALHOTRA/WLADAWSKY75, DEHNING/MAASS77, WOODS77). Die Fragestellung impliziert, schon, daß vom Standpunkt der Informatik aus betrachtet ein NS-Dialog nicht in allen denkbaren Situationen günstiger, als andere Dialogformen zu bewerten ist. Es wird hier also nicht pauschal angenommen, daß NSS immer die beste Lösung des Problems darstellen, einer größeren Klasse von Benutzern den Zugang zum Instrument DV-System zu eröffnen.

Als Randbedingungen sind vor allem das Benutzerprofil, die Häufigkeit der Interaktion sowie der Formalisierungsgrad der zugrundeliegenden Problemstellung zu betrachten.

Für die folgenden nur grob unterschiedenen Gruppen von Benutzern, ist ein NSS besonders interessant:

- 'naive' Benutzer: Benutzer, die weder Erfahrung im Umgang mit digitalen Rechenanlagen noch Informatikkenntnisse haben (vgl. MANN75).
- gelegentliche Benutzer: Benutzer, die in unregelmäßigen Zeitabständen mit dem DV-System interagieren und nicht durch Ihren Beruf oder ihre soziale Stellung zur Interaktion motiviert sind (vgl. CODD74, S.38).
- unerfahrene Benutzer: Benutzer, die obwohl sie als Spezialisten häufig mit DV-Systemen interagieren keine Erfahrungen mit bestimmten Interaktionsformen (z.B. Kommandos für das Arbeiten mit Magnetbändern) haben (vgl. WOODS77).

Außerdem liegen mit den folgenden beiden Situationen Randbedingungen vor, die für einen NS-Dialog sprechen (vgl., auch MALHOTRA/WLADAWSKY74):

- der Benutzer hat nur vage Vorstellungen von der Problemstruktur (z.B. ein 'naiver' Benutzer möchte mit dem System seinen nächsten Urlaub planen). Es handelt sich dabei vielfach um Situationen, in denen zu Beginn der Problemlösung kein definiter Test vorliegt, mit dem der Benutzer eine vorgeschlagene Lösung prüfen könnte (vgl. auch SIMON73).
- für den Benutzer ist das im Dialog zu lösende Problem wohldefiniert, aber die Bedienung des Systems ist ihm unklar (z.B. Manager möchte sich über Lagerbestände eines Unternehmens informieren).

In Fig. 5 ist der Wert eines NS-Dialogs bei unterschiedlichen Randbedingungen durch die jeweilige Anzahl von Punkten dargestellt.

BENUTZERGRUPPE	PROBLEMSTRUKTUR IST FÜR BENUTZER	
	WOHLDEFINIERT	\neg WOHLDEFINIERT
'NAIVER' BENUTZER	● ● ●	● ● ● ●
GELEGENTLICHER BENUTZER	● ●	● ● ●
UNERFAHRENER BENUTZER	-	●

FIGUR 5

Folgende Vorteile von Systemen mit natürlichsprachlicher Interaktionssprache werden häufig genannt:

- kein Zeitaufwand zum Erlernen einer formalen Sprache (besonders wichtig für 'naive' und gelegentliche Benutzer)
- der Benutzer ist frei von dem Ballast, sich eine Menge von Konventionen über eine längere Zeitspanne zu merken oder sich bei Bedarf wieder neu Sprachkonventionen anzueignen (wichtig für: gelegentlichen Benutzer).
- die Flexibilität, Universalität und konzeptuelle Struktur der NS erlauben einen schnellen, auf einem Dialog basierenden Einstieg in neue Problembereiche (wichtig für: 'naive' und unerfahrene Benutzer).

NSS sind speicher- und rechenzeitintensiver und damit teurer als konventionelle Systeme. Besteht der angestrebte Dialog mit dem DV-System aus relativ wenigen, sich ständig wiederholenden Interaktionsmustern, so ist der Dialog in einer für diese Interaktionsform speziell konstruierte formale Sprache günstiger, da die Ausbildungskosten geringer sind als die Kosten für ein NSS und die formale Sprache eine effizientere Gestaltung des Dialogs ermöglicht.

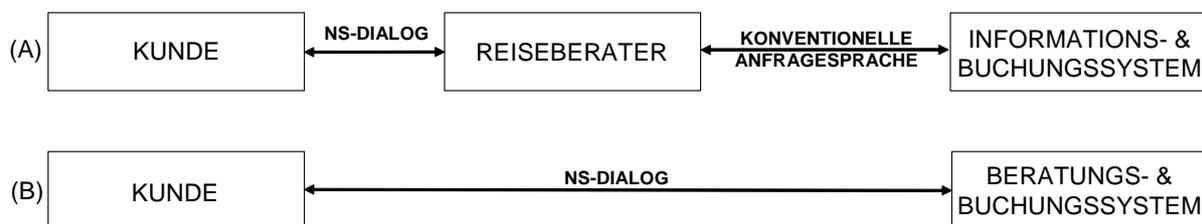
Der Versuch, konzeptuell sehr beschränkte Interaktionssprachen durch die Einführung von natürlichsprachlichen Bezeichnern oder einer an die natürliche Sprache angelehnte Syntax für die große, oben angesprochene Benutzergruppe attraktiver zu machen⁶, ist nicht sinnvoll, da er außer evtl. mnemotechnischer Vorteile die folgenden gravierenden Nachteile aufweist:

- Es besteht die Gefahr der Übergeneralisierung. Die Semantik der Bezeichner wird leicht verallgemeinert zur Semantik der entsprechenden NS-Ausdrücke. Auf diese Weise entsteht eine neue Fehlerquelle.
- Ein System, das, nachdem es dem Benutzer zunächst suggeriert hat, die stets mit der Benutzung von NS verknüpfte Ausdrucksvielfalt und ein gewisses Maß an Intelligenz zu besitzen, diesem ersten Eindruck nicht gerecht werden kann, stößt beim Benutzer schnell auf starke Ablehnung.

Am folgenden Beispiel soll die Begriffsbildung in diesem Abschnitt noch einmal veranschaulicht werden und anhand von konkret gegebenen Randbedingungen diese Frage nach der Nützlichkeit eines NS-Dialogs geprüft werden.

BEISPIEL: Situation (A): Ein Kunde eines Reisebüros hat sich spontan dazu entschlossen, irgendwo einen kurzen Erholungsurlaub zu machen. Der Reiseberater benutzt bei seiner Tätigkeit ein Informations- und Buchungssystem. Der Kunde tritt also mit einer Problemstellung, die nicht wohldefiniert ist, an den Reiseberater heran. Nach einem klärenden NS-Dialog mit dem Kunden ist für den Reiseberater das Problem wohldefiniert, und er interagiert mit dem System, um Informationen über die Buchungsmöglichkeiten für die in Frage kommenden Reiseziele einzuholen. In der gegebenen Situation wäre ein NSS für das Reiseunternehmen uninteressant.

Situation (B): Soll die Funktion des Reiseberaters aber durch die direkte Interaktion des Kunden mit einem Reiseberatungssystem ersetzt werden, so sollte dem Kunden als 'naivem' Benutzer mit seiner nur rudimentär definierten Problemstellung ein interaktives NSS zur Verfügung gestellt werden⁷ (vgl. Fig. 6).



FIGUR 6

Dieses Beispiel zeigt, daß ein solches natürlichsprachliches System, um den Kunden 'verstehen und beraten' zu können, zusätzlich zur Fähigkeit der syntaktisch-semantischen Analyse und Generierung mindestens auch über

- ein bestimmtes Allgemeinwissen
- Elemente einer Dialog- und Partnertaktik
- einige Inferenz- und Planungstechniken verfügen muß.

⁶Ein negatives Beispiel auf dem Gebiet der Programmiersprachen ist die Sprache COBOL

⁷Eine solche Lösung wird z.B. im GUS-System (vgl. BOBROWetal.76) angestrebt.

1.2.3. AKTUELLE FORSCHUNGSPROBLEME AUF DEM GEBIET DER NSS

Der Aufbau eines NSS, das auch nur annähernd die sprachlichen Ausdrucksmöglichkeiten menschlicher Kommunikationsformen bieten soll, setzt algorithmische Verfahren voraus, mit denen zumindest einige der im folgenden genannten sprachlichen Phänomene in einem maschinellen System berücksichtigt werden können.

Während in der traditionellen Linguistischen Datenverarbeitung und Computerlinguistik (vgl. DIETRICH/KLEIN74) häufig die Bearbeitung von Problemen der Morphologie und der Syntax im Vordergrund stand, bilden in den im Rahmen der KI entwickelten NSS Probleme der Semantik und Pragmatik der NS die Forschungsschwerpunkte, was auch in der Wahl von Bezeichnungen für dieses Forschungsgebiet wie 'Computer Semantics' (vgl. RAPHAEL/ROBINSON72) oder 'Computational Semantics' (vgl. CHARNIAK/WILKS76) zum Ausdruck gebracht wird.

Meine Auswertung einer von Waltz zusammengestellten Übersicht zu laufenden Forschungsprojekten auf dem Gebiet der natürlichsprachlichen Schnittstellen (vgl. WALTZ77) ergab, daß sich die gegenwärtige Forschung auf die Entwicklung theoretisch und praktisch adäquater Lösungen für folgende sprachlichen Probleme konzentriert:

- Behandlung von Mehrdeutigkeit und Vagheit (am häufigsten genannt)
- Probleme der Pronominalisierung und Deixis, Erzeugung von Textkohärenz, referenzsemantische Analyse von Nominalgruppen
- Berücksichtigung der metakommunikativen Ebene, Analyse von direkten und indirekten Sprechakten, Möglichkeiten zum kommunikativen Rollenwechsel, Verwendung von Partner- und Dialogtaktik
- Probleme des Tempussystems, der Adverbialsemantik, einer allgemeinen Behandlung von Konjunktion und Negation, der natürlichen Quantoren und der Behandlung einschränkender Relativsätze
- Möglichkeiten zur automatischen Korrektur von Eingabefehlern und zur Generierung von NS-Fehlerhinweisen, Einführung von Abkürzungen, neuen Definitionen und Begriffen auf einer metasprachlichen Ebene
- Berücksichtigung von sprachlich-gebundenen Inferenzen, Präsuppositionen und metaphorischem Sprachgebrauch
- Analyse von elliptischen Konstruktionen, Paraphrasen und Synonymen sowie von modalen und kausalen Fragen

Für jedes der aufgeführten Probleme gibt es in bereits implementierten Systemen Lösungsansätze. Es existiert gegenwärtig aber noch kein System, in dem alle angesprochenen Probleme behandelt werden, und - was noch wichtiger ist - es gibt kein System, welches eines der genannten sprachlichen Probleme in voller Allgemeinheit löst.

Die in der Literatur vorgeschlagenen 'Lösungen' sind immer an bestimmte Programmumgebungen, Mikrowelten oder Wissensstrukturen gebunden, was von den Autoren oft nicht in genügendem Maße deutlich gemacht wird.

Wegen dieser Einschränkungen ist eine Portierung oder Adaptierung von Programmmodulen, die eine semantische oder pragmatische Sprachverarbeitung leisten, im Gegensatz zu morphologischen und syntaktischen Routinen gegenwärtig kaum möglich.

Da die genannten linguistischen Probleme oft viele Teilaspekte umfassen, besteht eine weitere Einschränkung der meisten technischen Lösungen innerhalb von NSS der KI darin, daß nur bestimmte Teilaspekte des sprachlichen Phänomens behandelt werden. Dies muß für viele Forschungsprojekte der KI darauf zurückgeführt werden, daß häufig ohne Berücksichtigung der umfangreichen linguistischen Literatur von einem intuitiven Vorverständnis ausgehend Programme konzipiert und implementiert werden, die dann nur für einen kleinen Teilbereich des bearbeiteten sprachlichen Phänomens einsetzbar sind.

1.3 VAGES WISSEN

1.3.1. MEHRDEUTIGKEIT UND VAGHEIT ALS DIE BEIDEN ZENTRALEN PROBLEME IN NSS

Neben der Mehrdeutigkeit wird die Vagheit der natürlichen Sprache als eines der zentraler Probleme bei der Konstruktion von NSS betrachtet. Während die Mehrdeutigkeit in den meisten NSS berücksichtigt wurde, ist das Problem der Vagheit bisher meist unberücksichtigt geblieben.

Beispielsweise hat Winograd, obwohl er in seinen Arbeiten auf Probleme der automatischen Analyse unscharfer Formulierungen hinweist (vgl. WINOGRAD72, S. 129ff, WINOGRAD73, S.168), in seinem SHRDLU-System Vagheiten nicht systematisch berücksichtigt. Dies wird deutlich, wenn Winograd zur Darstellung der Bedeutung eines Wortes wie Millionär eine prozedurale RK der Form

(THGOAL (#POSSESS \$?X1 \$1.000.000))

vorschlägt (vgl. WINOGRAD72, S. 148). Jemand, der \$999.999 besitzt, wird dabei nicht mehr vom Programm als Millionär erkannt, was dem Sprachgebrauch widerspricht, der eher von einer vagen Vorstellung wie 'ein Millionär besitze rund \$1.000.000' ausgeht.

Zunächst sollen die Begriffe Mehrdeutigkeit und Vagheit definiert werden.

DEFINITION: Wir sprechen von Mehrdeutigkeit (syn.: Ambiguität, Polysemie) eines natürlichsprachlichen Ausdrucks genau dann, wenn der Ausdruck in mehreren Bedeutungen (syn.: Lesarten), deren Zusammenhang noch erkennbar ist (sonst: Homonymie), verwendet wird.

BEISPIEL: *Läufer* als Sportler, Schachfigur oder Teppich

DEFINITION: Wir sprechen von der Vagheit (syn.: Unscharfe, Verschwommenheit, engl.: fuzziness, vagueness) eines natürlichsprachlichen Ausdrucks, wenn es für bestimmte Anwendungsbereiche und -situationen keine eindeutigen Konventionen dafür gibt, ob der Ausdruck zur Bezeichnung eines bestimmten Gegenstands, Sachverhaltes oder Prozesses benutzt werden kann oder ob seine Verwendung ausgeschlossen ist, oder wenn bestimmte potentiell relevante Aspekte der Aussage, die durch den Ausdruck formuliert wird, offen bleiben.

BEISPIEL: *jung, Mittelklassewagen, gering, groß, wohlhabend* sind vage Wörter

SCHEFE77 gliedert den Bereich der Mehrdeutigkeit folgendermaßen:

- (1) syntaktische Mehrdeutigkeit
- (2) semantische Mehrdeutigkeit
 - (a) Mehrdeutigkeit von Wörtern
 - (b) von syntakt. Konstruktionen, insbesondere Kasuskonstruktionen
- (3) referentielle Mehrdeutigkeit
 - (a) anaphorische Referenz (Textreferenz)
 - (b) kontextuelle Referenz (Situationsreferenz)

Als Beispiele für Systeme, die Mehrdeutigkeiten aus den drei genannten Bereichen behandeln, seien hier drei am Institut für Informatik der Universität Hamburg auf einem DECsystem10 (PDP-10) implementierte NSS genannt (vgl. Fig, 7).

ART DER AMBIGUITÄT	SYSTEMBESCHREIBUNG
SYNTAKTISCH	H A S Y (SCHEFE76; MITTELSTEINetal.76)
INHALTSSEMANTISCH	H A N S A (WITTIG7576)
REFERENZSEMANTISCH	H A M – R P M (v.HAHNetal.76, WAHLSTER/v.HAHN76)

FIGUR 7

Zur Auflösung von Mehrdeutigkeiten werden in NSS der Kontext ausgewertet, Inferenzen und Weltwissen angewandt und Klärungsdialoqe geführt.

Für den Bereich der Vagheit geben wir zunächst nur eine grobe Klassifikation an (vgl. auch STAMMERJOHANN75, S.547ff). In Abschnitt 2.3. gehen wir genauer auf die linguistischen Grundlagen vager NS-Ausdrücke ein.

Vage Prädikationen können modifiziert werden durch:

- Gradangaben, z.B. wird ein Mensch mit einer Körpergröße von 1.90m eher als *groß* bezeichnet als jemand, der nur 1.75m groß ist.
- eine bestimmte Sichtweise, z.B. wird ein Mann mit einer Körpergröße von 1.90m im allgemeinen als *groß* bezeichnet, nicht aber, wenn er durch die Äußerung in seiner Rolle als Basketballspieler angesprochen wird.
- den sozio-kulturellen Hintergrund des Sprechers, z.B. wird ein 1.68m großer Mensch von einem Italiener als groß bezeichnet, von einem Schweden aber nicht.

Außerdem kann man unterscheiden zwischen:

- Pragmatischen Vagheiten, d.h. Vagheiten von Äußerungen, die in einer Sprechhandlungssemantik zu behandeln sind. Beispielsweise kann in gewissen Situationen das Verkehrsschild mit der Aufschrift *80 km bei Nässe* beim Empfänger dieser Nachricht Unsicherheit darüber auslösen, was konkret zu tun ist (Wann genau liegt Nässe vor?).
- Semantische Vagheiten, d.h. Vagheiten von Sätzen und Termen. Es gibt zumindest folgende Ausprägungen von semantischer Vagheit:
 - die Gebrauchsregel eines Terms ist in einer entlegenen Situation nicht genau bestimmt (z.B. die Unsicherheit, ob man bei bestimmten modernen Sitzelementen von *Stühlen* sprechen kann).
 - die Gebrauchsregel) eines Terms stützt sich auf einen Vergleich mit einem Prototyp (vgl. 2.2.2.).
 - ein wohlgeformter NS-Ausdruck enthält eine partielle Spezifikation (z.B. ist *Fritz verläßt Hamburg* ein partiell spezifizierter Ausdruck, weil die Fragen wann, warum, wie usw. Fritz Hamburg verläßt offen bleiben).

Mehrdeutigkeit und Vagheit können in einem Ausdruck auch gepaart auftreten.

BEISPIEL: *Otto war ein großer Mann*, in diesem Satz ist *groß* ambig; es kann sich z.B. auf Ottos Körpergröße oder Ottos Bekanntheitsgrad beziehen. Nehmen wir an, daß der Satz zu *Otto war ein bezüglich der Körpergröße großer Mann* disambiguiert wurde, so ist der Ausdruck *großer Mann* nach der oben gegebenen Definition unscharf.

Da wir uns im folgenden nur mit dem Problem der Vagheit in NSS der KI beschäftigen wollen, vereinbaren wir, daß alle Beispiele bereits disambiguiert sein sollen.

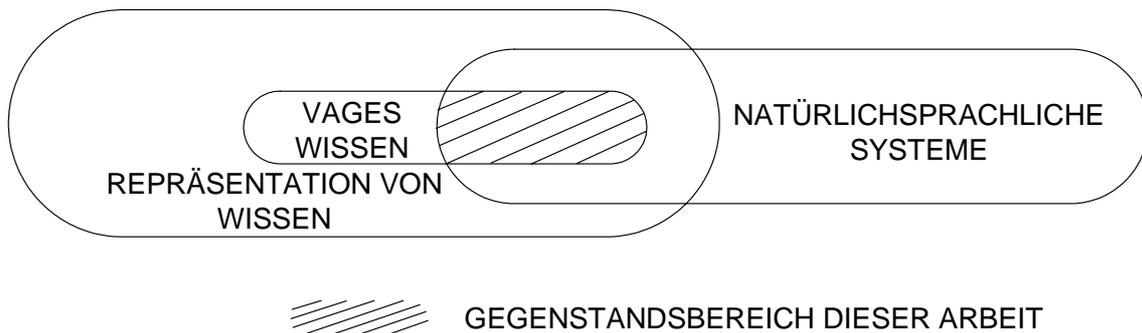
Nachdem wir definiert haben, was wir unter einem vagen NS-Ausdruck verstehen wollen, können wir nun den für diese Arbeit zentralen Begriff 'vages Wissen' definieren. Wir gehen davon aus, daß jegliches Wissen in Form von Wissensbehauptungen verbalisiert werden kann (vgl. 1.1.2.).

DEFINITION: Als vages Wissen bezeichnet man Wissen genau dann, wenn die dem Wissen entsprechenden Wissensbehauptungen vage natürlichsprachliche Ausdrücke enthalten oder durch Formulierungen wie *Ich vermute, daß p*, *Vielleicht p*, *Ich glaube, daß p* qualifiziert werden.

BEISPIEL: Folgende Wissensbehauptungen beziehen sich auf vages Wissen:

Ich bin mir ziemlich sicher, daß Bobrow jetzt bei Xerox arbeitet und viel verdient.
Minsky ist schon relativ alt.

Nachdem die Begriffe 'Repräsentation von Wissen', 'Natürlichsprachliches System' und 'vages Wissen' definiert wurden, soll die bereits aus dem Titel der Arbeit hervorgehende Abgrenzung des Gegenstandsbereichs in Fig. 8 durch ein vereinfachendes Venn-Diagramm verdeutlicht werden.



FIGUR 8

1.3.2. KOMMUNIKATIVE FUNKTIONEN VAGER FORMULIERUNGEN

Da die Mehrdeutigkeit von NS-Ausdrücken nur in wenigen Fällen eine kommunikative Funktion hat (z.B. in Witzen) und oft zu unbeabsichtigten Mißverständnissen führen kann, wird sie bei der Konstruktion von NSS nicht etwa als Vorteil sondern als eine besondere Schwierigkeit gesehen, die es zu bewältigen gilt, wenn man einen NS-Dialog Mensch-DV-System ermöglichen will.

Dagegen kann der Vagheit in zumindest folgenden Fällen eine besondere kommunikative Funktion zugeschrieben werden:

- der genaue Wert, nach dem gefragt wurde, ist dem System nicht bekannt. Sind unscharfe Formulierungen nicht vorgesehen, so kann die Frage vom System nicht beantwortet werden. Nur unter Verwendung vager Ausdrücke kann eine Antwort generiert werden.

BEISPIEL: Die Frage *Zu welcher Uhrzeit sah Meyer den Täter* kann nur vage mit *Es war um die Mittagszeit* beantwortet werden.

- aus bestimmten Gründen, die z.B. im Datenschutz bestehen, soll dem Empfänger einer Nachricht kein präziser Wert genannt werden. So ist es denkbar, daß ein FAS abhängig vom Status des Benutzers gewisse Informationen vage oder präzise formuliert ausgibt.

BEISPIEL: Benutzer A und B, die unterschiedliche Zugriffsrechte auf eine Datenbasis haben, fragen *Wieviel verdient Meyer?*. Antwort für Benutzer A: *Meyer verdient 5550 DM*. Antwort für Benutzer B: *Meyer verdient viel*.

- von partner- und dialogtaktischen Überlegungen ausgehend nimmt das System an, daß ein präziser Wert für den Empfänger irrelevant ist. Eine Verdichtung auf das Wesentliche unter Verwendung vager Prädikationen steigert die Benutzerfreundlichkeit eines FAS in bestimmten Situationen erheblich.

BEISPIEL: Ein Manager stellt die Anfrage *Was für einen Gewinn machten wir in diesem Jahr in der Sportabteilung?*. In vielen Fällen ist der Benutzer wesentlich zufriedener mit der Antwort *Geringfügige Gewinne* als mit einer extensionalen Antwort, die mehrere Seiten des Sichtgerätes mit der kompletten Bilanz für die Sportabteilung füllt. Erst wenn der Benutzer genauere Angaben anfordert, sollte ein 'intelligentes' FAS die komplette Bilanzübersicht ausgeben.

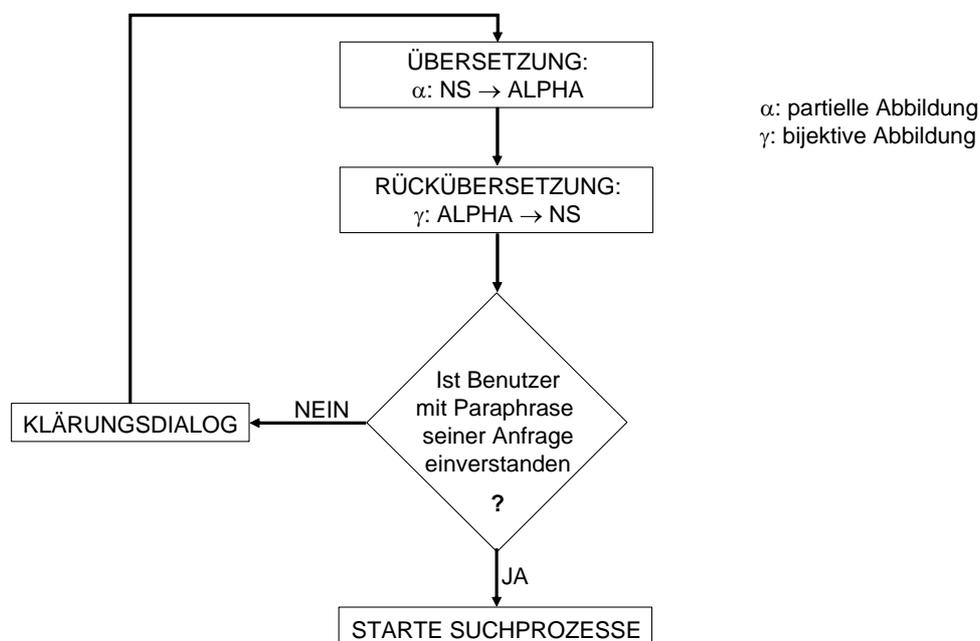
- auch wenn der Benutzer unscharfe Ausdrücke in der Anfrageformulierung verwendet, sollte das System die Frage nicht zurückweisen, sondern entweder eine Präzisierung der Anfrage verlangen oder versuchen, nach einer zwangsläufig idiosynkratischen Interpretation des vagen Ausdrucks eine Antwort zu generieren.

BEISPIEL: Frage: *Nenne alle Lieferanten, deren Bewertung 5 übertrifft.*
Übertreffen könnte durch die präzise Angabe *größer als* oder durch eine vage Formulierung wie *viel größer als* ersetzt werden (vgl. CODD74).

Auf die Frage, wie unscharfe Formulierungen eines Benutzers von einem FAS zu bearbeiten sind, lassen sich aus der Literatur zwei sehr unterschiedliche Antworten extrahieren:

- Die formale Anfragesprache, in die eine NS-codierte Frage des Benutzers übersetzt wird, enthält keine Repräsentationskonstruktionen für vage Terme. Daher ist gegebenenfalls in einem Klärungsdialog eine Präzisierung der Anfrage anzustreben. Besonders für solche Systeme mit einer NS-Schnittstelle, die eine Übersetzung der NS in eine formale Datensprache (z.B. ALPHA, SEQUEL, SQUARE, QUERY BY EXAMPLE) durchführen, ist dies auch eine technische Notwendigkeit. Die Technik der Rückübersetzung und des Klärungsdialogs wird zu diesem Zweck in Systemen wie RENDEZVOUS (vgl. CODD74) und MYCIN (vgl. SHORTLIFFEetal.75, S.316) erfolgreich eingesetzt.

Fig. 9 zeigt das Prinzip der Rückübersetzung und des Klärungsdialogs am Beispiel von RENDEZVOUS.



FIGUR 9

Diese Rückversicherungsstrategie entspricht in der alltäglichen Kommunikation Formulierungen wie *Verstehe ich Sie recht, Sie wollen...?* und ist besonders dort angebracht, wo wegen aufwendiger

Suchprozesse die Generierung einer evtl. den Benutzer nicht befriedigenden, vorläufigen Antwort nicht akzeptiert werden kann. Eine besondere Schwierigkeit ergibt sich, wenn das System das Ergebnis einer vom Benutzer angegebenen Präzisierung abspeichert und in späteren Dialogen wieder einsetzt.

BEISPIEL: So kann z.B. die Frage *Ist Fritz alt?* zu *Ist Fritz Alter größer als 55?* präzisiert werden. Da diese Präzisierung höchst subjektiv ist, sollte das System in einem späteren Dialog mit anderen Benutzern zwar in einer Rückübersetzung diese gespeicherte Präzisierung anbieten, aber nicht automatisch, ohne dem Benutzer eine Möglichkeit zu einer anderen Präzisierung zu geben, diesen Wert einsetzen.

Malhotra fordert, daß FAS auf unscharf formulierte Fragen wie *Stiegen die Kosten im letzten Jahr erheblich?*, die eine Ja/Nein-Entscheidung verlangen, nur indirekt antworten, indem sie dem Benutzer die notwendigen Informationen zur Verfügung stellen, aus denen er sich das gewünschte Urteil selbst ableiten kann. Das System sollte nach Malhotra kein Urteil bereitstellen (vgl. MALHOTRA75. S.229).

Auch wenn die Repräsentationssprache, in die übersetzt wird, keine Konstruktionen für vage Terme bereitstellt, müssen vage Prädikationen in der Oberflächenstruktur der Benutzeranfrage vom System erkannt werden, um in einem Klärungsdialog eine Präzisierung dieser Prädikationen zu erreichen.

- Die interne Repräsentationssprache läßt eine Darstellung vager Terme zu und das System kann Anfragen beantworten, in denen unscharfe sprachliche Ausdrücke verwendet werden.

In der KI werden häufig Systeme vorgeschlagen und implementiert, die über ein Überzeugungssystem (engl. belief System) verfügen und daher auch subjektiv gefärbte, idiosynkratische Antworten generieren (vgl. z.B. ABELSON73, BRUCE75, COLBY75). HEMPHILL75 betont, daß jedes System, das Prozesse des Sprachverstehens modelliert, notwendigerweise idiosynkratisch ist, da die an die modellierte Person gebundenen Erfahrungen und Kenntnisse bestimmen, in welcher Weise Äußerungen verstanden werden.

BEISPIEL: Auf die Frage *Ist Otto vermögend?* antwortet ein System *Nein*, da ein Vergleich zwischen der in einer Wissensquelle abgespeicherten Bedeutungsdefinition des Wortes *vermögend* und der in einer anderen Wissensquelle abgespeicherten Aussage *Otto besitzt 100000 DM* negativ ausfällt. Abgesehen vor der bei solchen Systemen angestrebten Simulation von Kommunikation zwischen menschlichen Partnern, die der weiteren Erforschung der Beziehungen zwischen Überzeugungssystem und kommunikativem Verhalten dienen soll, sind aber auch andere Anwendungen denkbar, wenn man sich klar macht, daß ein FAS lediglich ein Medium zur indirekten Kommunikation zwischen den Benutzern und den Konstrukteuren des Systems ist (vgl. auch GENRICH75, SCHEFE77) Die Datenbasis und das Programm, aus dem das System besteht, kann durchaus auf Norm- und Wertvorstellungen der Konstrukteure oder der von diesen modellierten Person oder Organisation Bezug nehmen. Wir werden in Kapitel 4.6. im Zusammenhang mit der Diskussion von Frame-Systemen sehen, daß ohne eine Verankerung von subjektiven Erfahrungswerten in der Wissensbasis eines Systems Aufgabenstellungen wie die Szenenanalyse oder automatisches Textverstehen nicht zu bewältigen sind.

Eine prinzipielle Beschränkung der Anwendbarkeit von Systemen, in deren Wissensbasis Normen verankert sind, ergibt sich dann, wenn die Normen sich laufend verändern. So müßte z.B. ein System, das, um die Frage *Ist das Auto X teuer?* beantworten zu können, Angaben über den üblichen Preis von Autos in einer Wissensquelle abgespeichert hat, ständig mit großem Aufwand an die wirtschaftliche Lage und das Preisniveau angepaßt werden.

Auf die wichtige gesellschaftliche Problematik, die in solchen Anwendungen zu sehen ist, kann in dieser Arbeit nicht eingegangen werden.⁸

⁸ Es wird mit Nachdruck auf die Arbeitspapiere des Werkstattgesprächs über 'Gesellschaftliche Auswirkungen großer Informationssysteme aus der Sicht verschiedener Disziplinen' (vgl. BRUNNSTEIN77) verwiesen.

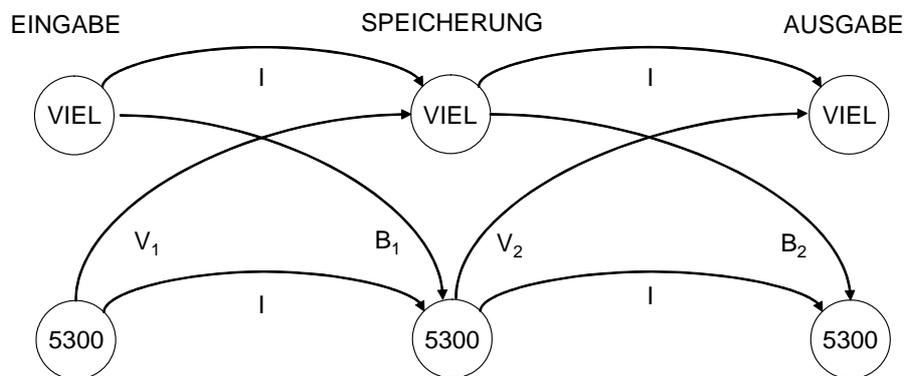
Bei der Vielfalt von kommunikativen Funktionen vager Formulierungen wird verständlich, warum gegenwärtig viele Autoren eine Berücksichtigung von Vagheiten in NSS fordern (vgl. z.B. MALHOTRA/WLADAWSKI75, WOODS77). Als typisches Beispiel für solche Forderungen zitieren wir Müllers Feststellung, die sich auf juristische FAS bezieht

Zur Gestaltung des natürlich-sprachlichen Dialogs sind Regeln nötig, die über sprachtypisch unscharfen Formulierungen operieren können. (MÜLLER77, S.12)

1.3.3. VAGES WISSEN IN DEN EINZELNEN KOMPONENTEN EINES NSS

Anhand von Fig. 10 soll gezeigt werden, in welchen Bereichen eines NSS vage sprachliche Terme vorkommen und auf welche Weise sie entstehen können.

Wir betrachten dazu die präzise Formulierung *Otto verdient 5300 DM* (Abk.: 5300) und die vage Formulierung *Otto verdient viel* (Abk.: VIEL).



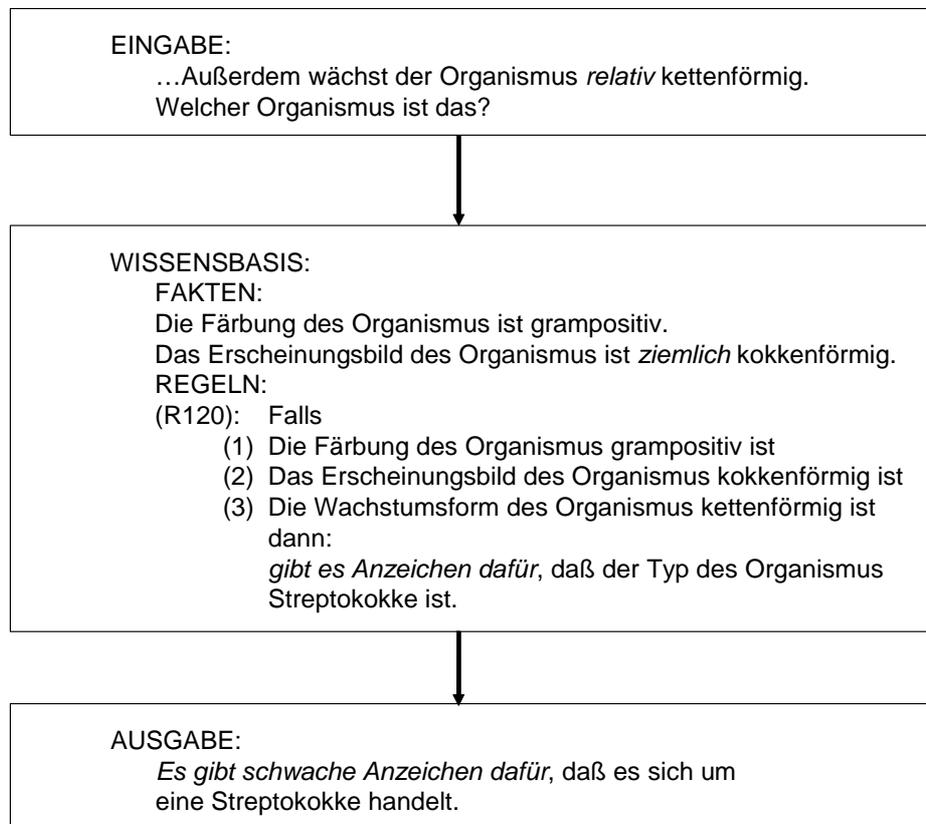
FIGUR 10

I ist die identische Abbildung. Bei Anwendung von I wird die eingegebene Nachricht im Speicher (Wissensbasis) aufbewahrt und auf Anfrage exakt reproduziert. Dieses Verhalten ist für klassische Informationssysteme das einzig mögliche. Wird eine Simulation des kommunikativen Verhaltens von Menschen angestrebt, so ist zusätzlich zu berücksichtigen, daß ein Mensch oft präzise Angaben innerhalb einer empfangenen Nachricht, falls diese genauen Angaben ihm irrelevant zu sein scheinen, mit Hilfe von auf Normen basierenden Funktionen (V_1) in vage Formulierungen übersetzt, die er sich dann besser merken kann. Auch das umgekehrte Phänomen, die Speicherung eines durch modellgesteuerte und beispielorientierte Inferenzen gewonnenen präzisen Wertes bei vagen Eingaben ist zu beobachten (B_1). Dieses Vorgehen, bei dem mit den Werten eines Prototyps gearbeitet wird, falls keine gegenteilige Information vorliegt, findet besonders in der Frame-Theorie (vgl. 4.6.) seine Anwendung.

V_2 bildet wie V_1 eine präzise Angabe auf eine unscharfe Formulierung ab und wird in den Fällen angewandt, in denen exakte Aussagen nicht angebracht zu sein scheinen (vgl. 1.3.2.). B_2 leistet eine vom jeweiligen Überzeugungssystem abhängige Übersetzung vager abgespeicherter Werte in präzise Angaben. Diese 'unzulässige' Präzisierung, die eine Vortäuschung von faktisch nicht vorhandenen Detailkenntnissen darstellt, wird in der menschlichen Kommunikation von einem Sprecher angewandt, um eine gewisse Aussage durch scheinbar präzise Angaben zu unterstreichen.

Schon diese stark vereinfachte, schematische Darstellung zeigt, wie komplex die kommunikative und kognitive Funktion von vagen sprachlichen Ausdrücken sein kann. Diejenigen NSS, die eine Simulation kognitiver und kommunikativer Prozesse zum Aufbau und Test wissenschaftlicher Modelle in diesem Bereich anstreben, müssen alle oben genannten Möglichkeiten berücksichtigen. In FAS sind die Funktionen V_1 , B_1 und B_2 normalerweise unerwünscht.

An dem in Fig. 11 dargestellten Beispiel, das sich an die Performanz des medizinischen Beratungssystems MYCIN (vgl. SHORTLIFFE74) anlehnt, soll gezeigt werden, in welchen Komponenten eines NSS Vagheiten berücksichtigt werden müssen.



FIGUR 11

In seiner Eingabe modifiziert der Arzt seine Prädikation 'Kettenförmig (Organismus)' durch die linguistische Hecke *relativ* (vgl. 2.3.1). Nach einer syntaktischen und semantischen Analyse wird die unscharfe Aussage des Arztes in eine semantische RK abgebildet, in der die in der sprachlichen Oberflächenstruktur ausgedrückte Modifikation der Prädikation berücksichtigt wird.

Die Frage des Benutzers wird in der Antwortbasis verarbeitet. Dabei werden Daten und Inferenzregeln verwendet, die in der Wissensbasis des Systems enthalten sind.

In der Wissensbasis können vage Propositionen enthalten sein (vgl. Fig. 11). Außerdem haben die Inferenzregeln unterschiedliche Aussagekraft, d.h. es handelt sich meist um abgeschwächte vage Implikationen (vgl. 4.2.). Ein weiteres Problem im Bereich Vagheit entsteht dadurch, daß Prämissen von Inferenzregeln häufig nur partiell erfüllt sind. In Fig. 11 sind die Prämissen 2 und 3 der vagen Schlußregel R120 nicht uneingeschränkt gültig. Der Wahrheitswert der durch die Inferenz abgeleiteten Aussage muß entsprechend eingeschränkt werden.

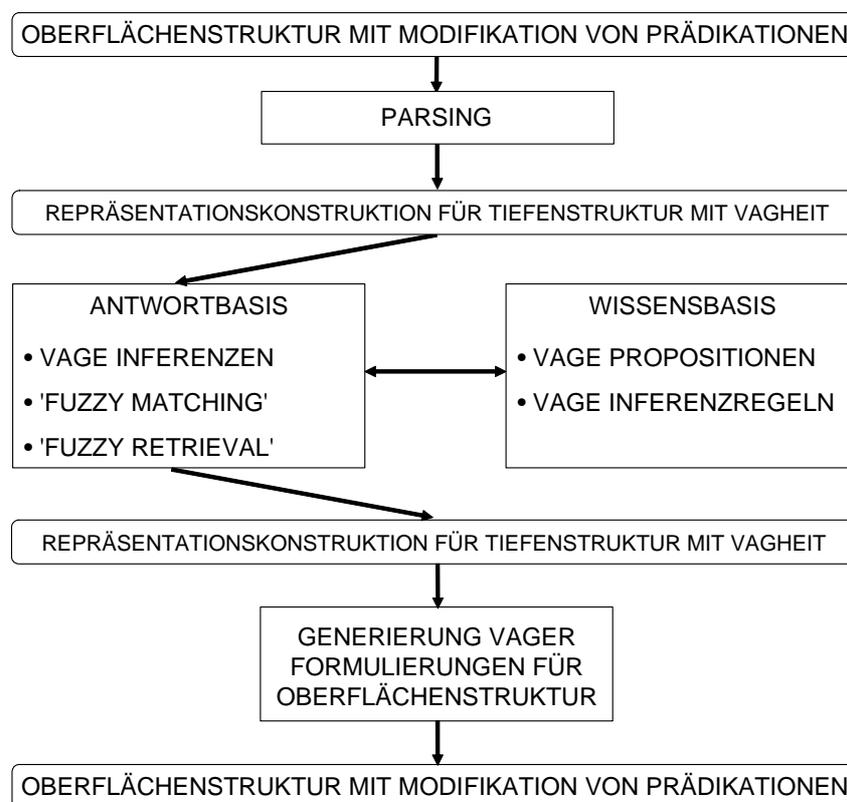
Eine weitere Komplikation bedeutet der Fall, daß sich über eine zweite Inferenzregel mit anderen Prämissen die gleiche Konklusion ableiten läßt, wobei allerdings ein anderer Wahrheitsgrad entsteht. Die beiden so gewonnenen Aussagen werden algorithmisch in eine einzige Aussage überführt (vgl. 4.2.).

Schließlich wird die auf diese Weise erhaltene RK für die Tiefenstruktur der Antwort durch Sprachgenerierungsalgorithmen in eine Oberflächenstruktur transformiert, in der die entsprechende Modifikation der Prädikation durch sprachliche Mittel ausgedrückt wird. Diese Oberflächenstruktur wird dann als Antwort ausgegeben.

Die Bedeutung der Repräsentation und Verarbeitung von vagem Wissen in den einzelnen Komponenten eines FAS wird in Fig. 12 zusammenfassend dargestellt.

Folgende Designziele für NSS ergeben sich im Zusammenhang mit der Repräsentation und Verarbeitung von vagem Wissen:

- der Entwurf von Algorithmen zur Verarbeitung natürlicher Sprache, die
- eine Analyse vager und partiell spezifizierter Anfragen durchführen
- eine Abbildung von vagen sprachlichen Ausdrücken in der Oberflächenstruktur auf formale RK ermöglichen
- eine von metakommunikativen Äußerungen oder vom Status des Benutzers abhängige Generierung von Antworten mit unterschiedlichen Genauigkeitsgraden bewirken
- der Entwurf von Repräsentationssprachen, die
- RK für vages Wissen zur Verfügung stellen
- eine formale Darstellung der Bedeutung vager sprachlicher Ausdrücke ermöglichen
- eine multiple Repräsentation von vagen und präzisen Werten eines Attributs zulassen
- eine effektive Verarbeitung und Anwendung von vagem Wissen fördern



FIGUR 12

2. UNTERSUCHUNGEN BENACHBARTER DISZIPLINEN ZUM PROBLEMKREIS 'VAGHEIT'

2.1. VAGHEIT AUS DER SICHT DER PHILOSOPHIE

2.1.1. DAS PARADOXON VON DER KLEINEN ZAHL

Schon seit der Antike beschäftigen sich Philosophen mit Problemen der Vagheit natürlichsprachlicher Ausdrücke. In der Antike waren es vor allem verschiedene natürlichsprachlich formulierte Paradoxa, die eine Diskussion über die Unschärfe sprachlicher Formulierungen auslösten.

Eines der bekanntesten Paradoxa, das in mehreren isomorphen Formulierungen vorliegt⁹, soll hier als Paradoxon von der kleinen Zahl diskutiert werden.

BEISPIEL: Paradoxon von der kleinen Zahl:

Prämisse (a): 1 ist eine kleine Zahl.

Prämisse (b): Wenn n eine kleine Zahl ist, dann ist auch $n + 1$ eine kleine Zahl.

Aus (a) \wedge (b) folgt: 10^{10} ist eine kleine Zahl.

Das Ergebnis der Folgerung, das intuitiv als ungültig eingestuft wird, ergibt sich durch wiederholte Anwendung des modus ponens auf Prämissen, die zunächst durchaus einleuchtend zu sein scheinen. Folgende Erklärungen dieses Paradoxons werden von Philosophen angeboten:

- Das Konzept 'kleine Zahl' ist sehr vage. Die Anwendung formaler Schlußverfahren auf unscharfe Prämissen führt zu unzulässigen Folgerungen.
- Die Prämisse (b) ist zwar keinesfalls falsch aber auch nicht völlig wahr. Die Wahrheit solcher Aussagen muß graduell angegeben werden. Da die Prämisse (b) nur eingeschränkt wahr ist, aber zur Herleitung einer Aussage wie ' 10^{10} ist eine kleine Zahl' $10^{10} - 1$ mal angewendet werden muß und für jede Anwendung des modus ponens 'gezahlt' werden muß, ist der Wahrheitsgehalt der gefolgerten Aussage sehr gering.¹⁰

Der erste Erklärungsversuch geht davon aus, daß es NS-Prädikate wie *arm - reich*, *groß - klein*, *viel - wenig*, *kurz - lang* usw. gibt, bei denen wir nicht wissen, an welchem Punkt der Hinzufügung oder der Wegnahme wir eine eindeutige Antwort geben können.

BEISPIEL: Es gibt keine Höhe h , so daß ein Mensch der Höhe h klein ist, während er mit einer Höhe von $h + \epsilon$, egal wie klein ϵ auch gewählt wird, nicht klein ist.

Es ist die mit der Vagheit von Prädikaten wie *arm - reich* verbundene Tatsache, daß für die Verwendung solcher Prädikate keine feste Grenzen festgelegt sind, die verhindert, daß aus einem 'reichen Mann' bei Verlust eines einzigen Pfennigs ein 'armer Mann' wird.

Der zweite Erklärungsversuch führt uns zu der ebenfalls schon in der Antike geführten Diskussion, die bis heute in der Philosophie fortgesetzt wurde, ob das Prinzip vom ausgeschlossenen Dritten (tertium non datur), d.h. $x \vee \neg x$ bzw. das Nichtwiderspruchsprinzip, d.h. $\neg(x \wedge \neg x)$ Tautologien sein sollten. Besonders seitdem in der mathematischen Logik Kalküle entwickelt wurden, die auf das

⁹ schon bei Cicero wurde es in zwei Formulierungen als Paradoxon vom Kahlköpfigen und Paradoxon vom Weizenhaufen vorgestellt; einen guten Überblick zu den verschiedenen Versionen gibt BLACK70.

¹⁰ Wir werden in einem späteren Kapitel auf die formale mehrwertige Logik eingehen (vgl. 4.1.) und schließlich die Formulierung dieses Paradoxons in einer Programmiersprache mit mehrwertiger Logik betrachten (vgl. 4.3.2.3.).

Axiom vom ausgeschlossenen Dritten verzichten (z.B. die intuitionistische Logik, die mehrwertige Logik), werden solche Paradoxa in der philosophischen Literatur wieder in stärkerem Maße besprochen.

Bereits G. Frege erkannte den Zusammenhang zwischen den beiden angeführten Erklärungsversuchen, indem er behauptete, daß das Axiom vom ausgeschlossenen Dritten nur eine andere Form der Forderung nach einer präzisen Definition aller Begriffe ist (vgl. FREGE62).

2.1.2. DEFINITION UND BEURTEILUNG VON VAGHEIT ALS SEMANTISCHES PROBLEM

Während einer besonders produktiven Phase der philosophischen Diskussion über den Problemkreis Vagheit zu Beginn des zwanzigsten Jahrhunderts (vgl. PEIRCE02, RUSSELL23, BLACK37, HEMPEL39) wurden die ersten Versuche zur formalen Beschreibung vager Terme unternommen (BLACK37, HEMPEL39). Auch die moderne Sprachphilosophie hat sich intensiv mit der Vagheit als semantisches Problem der NS auseinandergesetzt. Die verschiedenen Aspekte der Mehrdeutigkeit (vgl. 1.3.1), denen in der Linguistik besondere Aufmerksamkeit geschenkt wird, treten für Sprachphilosophen wie ALSTON64 gegenüber der Vagheit als zentrales semantisches Problem in den Hintergrund. Für BLACK73 besteht die Vagheit eines Wortes wie *Stuhl* darin, daß es Objekte gibt, für die die Frage, ob sie Element der Menge der Stühle sind, nicht eindeutig zu beantworten ist. Vergleichbare Definitionen der Vagheit von NS-Ausdrücken finden sich auch bei ALSTON64 und SCHAFF68.

DEFINITION:

A term is said to be vague if there are cases in which there is no definite answer as to whether the term applies.

(ALSTON64, S.84)

Ein Name ist dann ungenau, wenn er viele Gegenstände bezeichnet, deren Klasse nicht genau bestimmt ist.

(SCHAFF68, S.73)

Vereinfacht dargestellt existieren in der Philosophie zwei Haltungen gegenüber dem Phänomen der Unschärfe in der NS, wobei heute meist die zweite Haltung eingenommen wird.

- Die Unschärfe von Begriffen ist ein Defekt der NS (vgl. ADJUKIEWICZ75, S.52ff). Die Verschwommenheit vieler Begriffe muß beseitigt werden bzw. es muß eine 'ideale' Sprache konstruiert werden, die frei von jeglicher Vagheit ist und die als vollkommene Sprache die NS ersetzen kann. Philosophen wie Frege gingen sogar so weit zu behaupten, ein Begriff, der nicht scharf definiert ist, werde zu Unrecht 'Begriff' genannt (vgl. FREGE62).
- Die Unschärfe sprachlicher Ausdrücke ist ein wichtiges sprachliches Mittel, das eine Reihe kommunikativer Funktionen hat und z.B. eine Kommunikation bei unvollständiger Information und eine knappe Darstellung relevanter Informationen ermöglicht. Eine 'ideale' Sprache ist nicht konstruierbar, da es keine absolute Präzision gibt. Jede Messung birgt bekanntlich gewisse Fehler in sich.

Beide Haltungen hat der Philosoph Ludwig Wittgenstein in zwei verschiedenen Phasen der Entwicklung einer philosophischen Theorie vertreten. Nachdem er in seiner Arbeit *tractatus logico-philosophicus* (vgl. WITTGENSTEIN64) versucht hatte, eine ideale Sprache ohne Vagheit zu konstruieren, vertrat er später in den *Philosophischen Untersuchungen* die Auffassung, daß ein unexakter Begriff durchaus kein unbrauchbarer Begriff ist (vgl. WITTGENSTEIN71, S.59). Wittgenstein zeigt am Beispiel der Analyse des Begriffs *Spiel* (ebd. S.48ff), daß den Vorgängen, die als Spiel bezeichnet werden, nicht eine bestimmte Menge gemeinsamer Eigenschaften zugesprochen werden können, sondern daß zwischen den verschiedenen Vorgängen ein komplexes Netz von Ähnlichkeiten besteht, die *einander übergreifen und kreuzen* (ebd. S.48) und von ihm als *Familienähnlichkeiten* bezeichnet werden.

In vielen Kommunikationssituationen sind vage Ausdrücke sehr nützlich. So läßt ein Diplomat mit

der Drohung *Gegen jede Einmischung in die inneren Angelegenheiten unseres Staates gehen wir strengstens vor* durch die vage Formulierung *strengstens vorgehen* bewußt verschiedene Handlungsalternativen wie ökonomischen, politischen oder militärischen Druck offen. Außerdem bietet die NS die Möglichkeit, ad hoc Begriffe zu präzisieren. So kann z.B. arbiträr eine Vermögensgrenze festgelegt werden, um die Bedeutung von *wohlhabend* exakt zu definieren. Die gegenwärtige Nützlichkeit des Begriffes ginge bei einer allgemeinen Konventionalisierung dieser Präzisierung aber verloren.

2.1.3. WICHTIGE ERGEBNISSE DER PHILOSOPHISCHEN ARBEITEN ÜBER VAGHEIT

In Hinblick auf unsere Fragestellung der Repräsentation von vagem Wissen sind folgende Ergebnisse philosophischer Arbeiten über Vagheit besonders wichtig:

- Die NS ist ein so allgemeines und mächtiges Kommunikationswerkzeug, daß sie auch sprachliche Mittel bereitstellt, Ausdrücke in bestimmten Grenzen zu präzisieren.
- Für das praktische Handeln sind Präzisierungen von Termini auf dem Wege der Einführung von Konventionen und Normen z.B. im wissenschaftlich-technischen Bereich zwar wichtig, verschieben aber oft die unscharfen Randzonen lediglich auf andere Bereiche.

BEISPIEL: der Versuch, ein 'kaltes Getränk' als Getränk, das eine Temperatur von 2° Celcius hat, zu definieren, oder festzulegen, daß ein Bach ab 5m Breite in einen Fluß übergeht.

Die Notwendigkeit einer Präzisierung ist immer vom Ziel abhängig. Die Schärfung von Begriffen allein ist kein wissenschaftliches Ziel und kann in unangemessener Weise betrieben sogar zum Hindernis werden. Man sollte nie präziser sein, als das zu bearbeitende Problem es erfordert (vgl. POPPER76).

BEISPIEL: Es wäre äußerst hinderlich, im Alltag jede Messung mit einem Elektronenmikroskop vorzunehmen.

- Die Tatsache, daß in vielen Situationen Grade von Wahrheit oder zumindest ein Wert 'unbestimmt' (z.B. in der 3-wertigen Logik) eingeführt werden müssen, um eine adäquate Beschreibung zu ermöglichen, zeigt nicht, wie früher von Philosophen behauptet wurde, daß die zweiwertige Logik 'falsch' ist, sondern lediglich, daß in gewissen Situationen die Anwendung der zweiwertigen Logik als formale Beschreibungsmethode nicht angemessen ist. Bertrand Russell schreibt dazu:

All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life, but only to an imagined celestial existence.
(RUSSELL23, S.88)

Bei der Beschreibung komplexer Systeme ist eine gewisse Unschärfe unvermeidlich. Diese Aussage ergibt sich aus dem Zadehschen Inkompatibilitätsprinzip.

Zadehsches Inkompatibilitätsprinzip:

Sei $P \hat{=}$ der Grad der erreichbaren Präzision der Beschreibung eines Systems
 $K \hat{=}$ die Komplexität des zu beschreibenden Systems
dann gilt: $P \sim K^{-1}$

Stated informally, the essence of the principle is that as the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics. (ZADEH73, S.28)

- Der Versuch eine universelle, 'ideale' Sprache, die frei von jeglicher Vagheit ist, zu konstruieren, endet zwangsläufig in einer Sackgasse. Bei der Entwicklung von Repräsentationssprachen in der KI ist die Gefahr, in diese Sackgasse zu geraten, besonders groß, da in NSS Repräsentationskonstruktionen für vages Wissen formal interpretierbar sein müssen. In Kapitel 3 wird gezeigt, daß die notwendigerweise formale Darstellung von vagem Wissen nicht unbedingt mit einer arbiträren Präzisierung vager Begriffe verbunden sein muß.

2.1.4. ZUSAMMENFASSUNG

Die Ergebnisse philosophischer Untersuchungen zum Problem der Vagheit in der NS, die in den vorangegangenen Abschnitten vereinfachend ohne Bezug auf die entsprechenden philosophischen Schulen diskutiert wurden, können folgendermaßen zusammengefaßt werden:

- Die Vagheit ist kein Defekt der NS sondern trägt entscheidend zur Nützlichkeit des Kommunikationswerkzeugs NS bei.
- Die Komplexität unserer Umgebung macht es oft notwendig, von irrelevanten Details zu abstrahieren und Informationen durch vage Prädikationen zu komprimieren.
- Unscharfe Begriffe können, falls für bestimmte Zwecke notwendig, jederzeit durch sprachliche Mittel präzisiert werden.
- Der Versuch, eine universelle, vollkommene Sprache, die frei von jeglicher Vagheit ist, zu konstruieren, ist aus prinzipiellen Gründen zum Scheitern verurteilt und sollte daher durch die Entwicklung formaler Beschreibungsmethoden für Aspekte der Vagheit in der NS ersetzt werden.

2.2. VAGHEIT AUS DER SICHT DER PSYCHOLOGIE

2.2.1. EMPIRISCHE BELEGE FÜR DIE VAGHEIT SPRACHLICHER KATEGORIEN

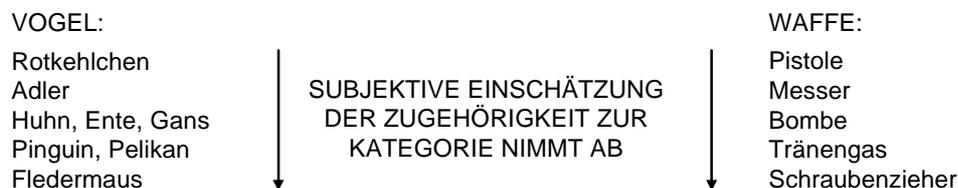
Neuere Arbeiten im Bereich der kognitiven Psychologie und der Psycholinguistik liefern empirische Belege dafür, daß eine Rekonstruktion natürlichsprachlicher Begriffshierarchien durch klassische mengentheoretische Konstruktionen (z.B. Obermenge, Untermenge, Element-von-Relation) nicht adäquat ist und durch Modelle ersetzt werden sollte, für die das Phänomen der Vagheit natürlichsprachlicher Bedeutungen konstitutiv ist.

Ausgehend von der Beobachtung, daß z.B. ein Stuhl als ein treffenderes Exemplar der Kategorie 'Einrichtungsgegenstand' empfunden wird als ein Radio und daß es z.B. Stühle gibt, die mehr und andere die weniger der jeweiligen Vorstellung von Stühlen entsprechen, wurden in der Psychologie verschiedene Experimente durchgeführt, um zu zeigen, daß die Mitglieder einer Sprachgemeinschaft die begriffliche Zugehörigkeit eher als graduellen Übergang denn als abrupten Sprung auffassen.

Die intensivsten psychologischen Untersuchungen zu diesem Problemkreis hat E. Rosch durchgeführt (vgl. z. B. ROSCH73/75ab, ROSCH/MERVIS75). Neben den Untersuchungen von Rosch gibt es noch eine Reihe experimenteller Arbeiten, die sich im Gegensatz zu Rosch unmittelbar auf Zadehs Theorie der 'fuzzy sets' beziehen und diese zur Darstellung und Interpretation ihrer Testergebnisse benutzen (vgl. 3.3.).

Zunächst soll auf Experimente von Rosch hingewiesen werden, welche die Hypothese bestätigen, daß die Zugehörigkeit zu einer Kategorie nicht als binäre sondern als stark gradierte Entscheidung aufgefaßt wird.

In einem ersten Experiment sollten die Testpersonen bestimmte ihnen vorgelegte Begriffe in Bezug auf eine vorgegebene Kategorie ordnen. Die Probanden lieferten weder eine einfache Zweiteilung noch eine völlig regellose Zuordnung sondern graduelle Abstufungen (vgl. Fig. 13).



FIGUR 13

Ein weiteres Experiment zeigte, daß die Antwortzeiten auf die Frage *Ist x ein y?* davon abhängt, ob x ein zentrales oder ein eher peripheres Element der Kategorie y ist. Bei den Elementen, für die in den vorausgegangenen Untersuchungen ein hoher Grad der Kategorienzugehörigkeit ermittelt wurde, waren die Antwortzeiten signifikant kürzer als bei peripheren Elementen.

Zunächst läßt sich also festhalten, daß Experimente der kognitiven Psychologie Indizien (keine Beweise!) dafür liefern, daß gewisse sprachliche Kategorien eine interne Struktur aufweisen und die Element-von-Relation für solche Kategorien stark gradiert ist (vgl. 3.1). Welche Abstufungen der Zugehörigkeit zu Kategorien im einzelnen vorgenommen werden, hängt u.a. vom sozio-kulturellen Hintergrund, dem Wissen und dem Überzeugungssystem der jeweiligen Testpersonen ab.

Es ist offensichtlich, daß diese Aussagen für Begriffshierarchien, die zu bestimmten Zwecken durch eindeutige Konventionen genormt wurden, einzuschränken sind.¹¹

¹¹ Allerdings werden auch in einer Wissenschaft wie der Informatik vage Kategorien benutzt. Man denke z.B. an die Unterscheidung von niederen, höheren und sehr hohen Programmiersprachen oder Klein- und Großrechnern.

2.2.3. ZUR STRUKTUR VAGER BEGRIFFE: DIE FAMILIENÄHNLICHKEITSHYPOTHESE

Nachdem gezeigt wurde, daß unscharfe natürlichsprachliche Kategorien eine interne Struktur haben und daß Prototypen eine besondere Stellung innerhalb einer Kategorie besitzen, soll jetzt gefragt werden, welchen Bedingungen Elemente einer Kategorie entsprechen müssen, um als Prototyp zu gelten.

Neben Faktoren wie die Häufigkeit des Auftretens eines Elements im alltäglichen Lebensbereich und perzeptuell besonders hervorstechenden Merkmalen eines bestimmten Elementes einer Kategorie dominiert folgendes in der Familienähnlichkeitshypothese formulierte strukturelle Prinzip die Bildung von Prototypen:

Familienähnlichkeitshypothese:

Prototypen einer Kategorie sind die Elemente, die

- am meisten Eigenschaften mit anderen Elementen aus ihrer Kategorie
- am wenigsten Eigenschaften mit Elementen aus anderen Kategorien gemeinsam haben.

Für diese Hypothese sprechen u.a. folgende von Rosch durchgeführten Experimente. Zunächst wurden die Probanden aufgefordert, aus Bildern von Objekten, die zu einer bestimmten Kategorie (z.B. Fahrzeuge) gehören, diejenigen auszuwählen, welche typische Beispiele für die Kategorie sind, und zusätzlich für jedes dargestellte Objekt möglichst viele Eigenschaften zu nennen. Es zeigte sich, daß die typischen Beispiele diejenigen Bilder waren, die am meisten Eigenschaften mit den anderen Bildern gemeinsam hatten. In einem weiteren Experiment wurden Bilder von kontrastierenden Kategorien hinzugenommen. Die Familienähnlichkeitshypothese wurde durch die Ergebnisse des Experiments bestätigt: Je typischer die dargestellten Objekte waren, desto weniger Eigenschaften hatten sie mit kontrastierenden Bildern gemeinsam.

Mit Hilfe der Familienähnlichkeitshypothese und der Prototypentheorie läßt sich auch erklären, warum in vielen Anwendungskontexten keine Schwierigkeiten bei der Verwendung unscharfer Begriffe wie *Stuhl* oder *Sessel* entstehen, in beispielorientierten Inferenzen wird gedanklich mit typischen Exemplaren der Kategorien operiert, so daß die beiden unscharfen Kategorien maximal diskriminiert sind. Erst wenn genaue Definitionen der Eigenschaften von Elementen einer sprachlichen Kategorie verlangt oder weniger typische Objekte vorgefunden werden, erkennt man bewußt die Vagheit vieler Begriffe.

Nachdem wir bisher nur die interne Struktur einer einzelnen Kategorie betrachtet haben, soll zum Schluß ein allgemeines Strukturprinzip für ein System von Kategorien vorgestellt werden.

Die grundlegenden kategorialen Einschnitte bei der sprachlichen Erfassung unserer Umwelt sind nicht arbiträr, sondern sie trennen informationsreiche Merkmalbündel voneinander. Es können mindestens drei Arten von sprachlichen Kategorien unterschieden werden:

DEFINITION: Grundkategorien sind die am stärksten umfassenden Kategorien. ihre Elemente haben untereinander viele Eigenschaften gemeinsam und sind gut von Elementen kontrastierender Kategorien abgegrenzt.

Übergeordnete Kategorien sind Kategorien, deren Elemente wesentlich weniger gemeinsame Eigenschaften als die Elemente einer Grundkategorie haben.

Untergeordnete Kategorien sind Kategorien, deren Elemente wesentlich mehr Attribute mit kontrastierenden Kategorien gemeinsam haben als Elemente einer Grundkategorie.

ÜBERGEORDNETE KATEGORIE:	Fahrzeug	wenig gemeinsame Eigenschaften z.B. Auto ↗ Fahrrad
GRUNDKATEGORIE:	Auto	viele gemeinsame Eigenschaften gute Abgrenzung
UNTERGEORDNETE KATEGORIE:	Sportwagen	schlechte Abgrenzung gegenüber kontrastierenden Kategorien z.B. Sportwagen ↗ Rennwagen

FIGUR 14

Im Sinne einer maximalen Diskrimination können also zwei Stufen der Kategorienbildung unterschieden werden:

- In Grundkategorien wie *Auto*, *Hund*, *Haus* usw. werden informationsreiche 'Cluster' beobachteter Eigenschaften der Umgebung abgebildet.
- Durch Prototypen innerhalb von Grundkategorien werden noch stärker diskriminierende Merkmalsbündel gebildet.

2.2.4. ZUSAMMENFASSUNG

Die wichtigsten Ergebnisse der psychologischen Untersuchungen zum Phänomen der Unschärfe in natürlichsprachlichen Kategorien können in folgenden Aussagen zusammengefaßt werden:

- Die Zugehörigkeitsrelation zu unscharfen sprachlichen Kategorien ist stark gradiert (vgl. 3.1.1.).
- Es gibt ausgezeichnete Elemente von unscharfen Kategorien, sog. Prototypen, die als typische Beispiele für die Elemente einer sprachlichen Kategorie, als Anhaltspunkte für Klassifikationsaufgaben oder als Repräsentanten der Kategorie in beispielorientierten Inferenzen fungieren.
- Prototypen werden nach einem bestimmten strukturellen Prinzip (Familienähnlichkeitshypothese) durch Informationsverarbeitungsprozesse über den Merkmalen von Elementen einer Kategorie gebildet.
- Es können zumindest drei Arten von Kategorien unterschieden werden: Grundkategorien, übergeordnete und untergeordnete Kategorien. Auf der Ebene der Kategorien haben Grundkategorien die gleiche Funktion, die Prototypen auf der Ebene der Elemente einer Kategorie haben.

Für die Repräsentation von vagem Wissen in NSS ergeben sich aus diesen Feststellungen folgende Konsequenzen:

- in der RS sollten RK zur Darstellung von 'Mengen' mit stark gradierter Zugehörigkeitsrelation vorhanden sein. Dieser Aspekt der Vagheit wird uns in Kapitel 3 zentral beschäftigen.
- Matching-, Inferenz- und Retrievalalgorithmen über vagem Wissen sollten die Prototypentheorie berücksichtigen. Wir werden feststellen, daß die 'Frame'-Theorie und besonders die Verwendung von Standardwerten in 'Frames' dies ermöglicht (vgl. 4.6.). in der sehr hohen Programmiersprache KRL werden komplexe Datenstrukturen zur Darstellung von Prototypen zur Verfügung gestellt (vgl. 4.6.2.).
- Die besondere Struktur von Grundkategorien sollte in die RS eingehen und für Retrieval-, Matching- und Inferenzalgorithmen eingesetzt werden können.

2.3. VAGHEIT AUS DER SICHT DER LINGUISTIK

2.3.1. ZUR MODIFIKATION VAGER PRÄDIKATIONEN DURCH LINGUISTISCHE HECKEN

Als Ausgangspunkte für die linguistischen Untersuchungen zum Problemkreis Vagheit (vgl. auch SCHMIDT74) sind neben den Impulsen aus der Sprachphilosophie (vgl. 2.1.2.) vor allem die Versuche zur formalen semantischen Beschreibung von zwei eng miteinander verknüpften sprachlichen Phänomenen zu nennen:

- die Möglichkeit zur Modifikation von Prädikationen durch sog. linguistische Hecken
- die Existenz sog. relativer Adjektive und die damit verbundenen Möglichkeiten zur Graduierung und zum Vergleich

DEFINITION: Als linguistische Hecke (kurz: Hecke, engl.: linguistic hedge) bezeichnen wir sprachliche Einheiten, die Prädikationen nach Grad oder Hinsicht ihres Zutreffens modifizieren und als Operatoren interpretiert werden können, welche die Vagheit des sprachlichen Konzeptes, auf das sie angewendet werden, verstärken oder abschwächen.

BEISPIELE: *sehr, besonders, relativ, eine Art, grob gesprochen, mehr oder weniger, ziemlich, praktisch, in mancher Hinsicht, letztlich, gewissermaßen, vor allem, genau betrachtet, höchst, äußerst, prinzipiell, par excellence, außergewöhnlich, regelrecht, im Grunde genommen, quasi, „_“, typisch, ungefähr, im wesentlichen, etwa, nahezu, enorm, zirka, halbwegs annähernd, rund, hochgradig, einigermaßen*

Die Bezeichnung 'Hecke' unterstreicht eine der möglichen Funktionen der genannten Klasse von sprachlichen Elementen: Der Blick auf ein durch eine (nicht-linguistische) Hecke verdecktes Objekt ergibt ein ungenaues Bild.

Nicht nur in der Umgangssprache sondern z.B. auch in fachsprachlichen Texten werden Hecken häufig verwendet.

Auf die von ZADEH71/72/73/74/75/76 und LAKOFF75 entwickelten Formalismen zur Beschreibung von Hecken im Rahmen der Theorie der 'fuzzy sets' gehen wir erst in Abschnitt 3.2.2. ein. Zunächst wird die Funktion von Hecken informell an Beispielen erläutert.¹²

Die von Rosch untersuchten Abstufungen der Zugehörigkeit von Objekten zu einer sprachlichen Kategorie spielen bei der semantischen Interpretation der in Fig. 15 verwendeten Beispielsätze, in denen Hecken enthalten sind, eine entscheidende Rolle.

BEISPIELSATZ	Beurteilung des Wahrheitswertes
Ein Rotkehlchen ist eine Art Vogel.	FALSCH
Ein Huhn ist eine Art Vogel.	WAHR
Ein Rotkehlchen ist ein Vogel par excellence.	WAHR
Ein Huhn ist ein Vogel par excellence.	FALSCH
In mancher Hinsicht ist eine Fledermaus ein Vogel.	WAHR
In mancher Hinsicht ist ein Rotkehlchen ein Vogel.	Präsuppositionsverletzung

FIGUR 15

Bereits aus diesen wenigen Beispielen lassen sich einige Aussagen herleiten über den Zusammenhang zwischen dem Zugehörigkeitsgrad eines Elements zu einer vagen sprachlichen Kategorie (hier: Vogel) und der Modifikation der entsprechenden Element-von-Relation durch linguistische Hecken:

- *eine Art* erfordert nur einen mittleren Zugehörigkeitsgrad

¹² Dabei wird auf eine Diskussion der Anwendung von Hecken auf performative Verben verzichtet, da diese Verwendung von Hecken noch kaum erforscht ist (vgl. FRASER75).

- *par excellence* erfordert einen sehr hohen Zugehörigkeitsgrad
- *in mancher Hinsicht* erfordert einen geringen Zugehörigkeitsgrad und präsupponiert, daß es sich nicht um ein typisches Beispiel der Kategorie handelt.

Die sog. limitierenden Adverbiale (vgl. BARTSCH72, S.63), wie *in privater Hinsicht*, *äußerlich*, *gesundheitlich* usw. bewirken häufig eine Verringerung der Unschärfe, indem sie die Gültigkeit einer vagen Prädikation auf einen bestimmten Bereich einschränken.

BEISPIEL: *Peter und Dieter sind sich äußerlich sehr ähnlich* kann als eine gewisse Präzisierung des vagen Prädikats *sehr ähnlich* in *Peter und Dieter sind sich sehr ähnlich* interpretiert werden (vgl. 4.6.2.3.).

Oft wird in der Literatur über Hecken der Vorschlag gemacht, die Frage, ob ein sprachlicher Ausdruck vage ist oder nicht, dadurch zu testen, daß man versucht Hecken auf den Ausdruck anzuwenden (vgl. ZADEH72, S.23, ZADEH75, S.251). So klassifiziert ZADEH75 *krank* als vages Konzept und *tot* als nicht vages Konzept, da man von einem *sehr kranken Mann*, nicht aber von einem *sehr toten Mann* sprechen könne.

Dagegen ist einzuwenden, daß eine Funktion von Hecken darin besteht, präzise definierte Konzepte wie *quadratisch* in vage Konzepte zu überführen, wie das Beispiel *dieser Tisch ist ziemlich quadratisch* zeigt (vgl. auch die Kritik von DAMERAU75). Besonders deutlich wird diese Wirkung von Hecken, wenn man die von KLOCKOW76 analysierten Anführungszeichen betrachtet, die in bestimmten Anwendungen als Alternative zu verbalen Hecken wie *eine Art* anzusehen sind. Falls ein Sprecher in einer Kommunikationssituation gerade nicht den richtigen Ausdruck findet, wendet er oft auf Ausdrücke die Hecke "_" an, um anzudeuten, daß keine konventionskonforme, scharfe Verwendung des Ausdrucks vorliegt. Am Beispielsatz *Das "Gehirn" eines Taschenrechners ist winzig* wird deutlich, daß die Einführung von Unschärfe durch den Gebrauch von Hecken oft mit einem metaphorischen Sprachgebrauch verbunden ist. Neben der häufig mit Vagheit gepaart auftretenden Mehrdeutigkeit wollen wir auch den metaphorischen Sprachgebrauch bei unseren Untersuchungen zur Repräsentation von vagem Wissen stets ausschließen (vgl. 1.3.1.).¹³

Hecken können nicht nur auf Substantive und Adjektive sondern auch auf Verben (z.B. *Fritz ähnelt seinem Vater sehr*) und Adverbiale (z.B. *Fritz versteht alles sehr schnell*) angewandt werden.

Es ist möglich, Hecken aufeinander anzuwenden (z.B. *sehr sehr typisch*) und vage Konzepte mit anderen vagen Konzepten zu kombinieren (z.B. *etwas weniger schnell*). Dabei sind allerdings die Kombinationsmöglichkeiten durch zahlreiche Restriktionen beschränkt; z.B. ist *sehr regelrecht* keine wohlgeformte, kombinierte Hecke. Fine spricht im Zusammenhang mit Sätzen wie *James has a few fewer hairs than his friend* von Vagheit höherer Ordnung (vgl. FINE75, S.287).

Auch die Verwendung von Indefinitpronomen (z.B. *ein paar*, *einige*, *wenige*), die häufig als 'natürliche' Quantoren bezeichnet werden, und von Gradpartikeln (vgl. ALTMAN76) führt zu unscharfen Formulierungen wie *einige lachten*, *wenige pfliffen*, *mehrere spendeten Beifall*.

2.3.2. LINGUISTISCHE UNTERSUCHUNGEN VAGER AUSDRÜCKE AM BEISPIEL DER RELATIVEN ADJEKTIVE

Am Beispiel der Wortart 'Adjektiv' werden im folgenden Ergebnisse neuerer linguistischer Untersuchungen im Bereich Vagheit¹⁴ dargestellt und ihre Bedeutung für Probleme der Repräsentation von vagem Wissen diskutiert.

Zunächst unterscheidet man in einer syntaktisch orientierten Klassifikation vergleichsfähige (syn.: graduierbare) Adjektive von vergleichsunfähigen, für die keine Komparativ-, Superlativ- oder

¹³ Erste Ansätze zu einer maschinellen Verarbeitung von Metaphern in NSS findet man z.B. in WILKS75 und RUSSELL76.

¹⁴ vgl. auch die Übersichten in BARTSCH/VENNEMANN73, Kap. 2, DAMERAU75, EISENBERG76, S. 117ff

Elativformen gebildet werden können und auf die weder Gradadjektive (z.B. *außergewöhnlich*) noch Gradadverbien (z.B. *sehr*) angewandt werden können (vgl. MOTSCH64). Die Möglichkeit zur Unterscheidung von Gradabstufungen kann als Indiz für eine gewisse Vagheitszone des Adjektivs interpretiert werden, die bei nichtgraduierbaren Adjektiven (z.B. *blind, gläsern, verheiratet*) seltener zu finden ist.

In einer semantischen Klassifikation werden sog. relative Adjektive¹⁵ von absoluten (z.B. *eine gerade Zahl*) und non-standard Adjektiven (z.B. *der angebliche Agent*) unterschieden.

DEFINITION: Als relative Adjektive bezeichnen wir Adjektive, mit deren Anwendung stets ein impliziter Vergleich verbunden ist. Das Objekt, auf das sich ein relatives Adjektiv bezieht, wird in eine Relation zu anderen Objekten derselben Klasse gestellt. Ohne kontextuelle Angabe der Referenzmenge, auf die sich der Vergleich bezieht, ist eine Prädikation, in der ein relatives Adjektiv benutzt wird, unvollständig.

BEISPIELE: *Fritz ist groß, Fritz ist reich, Fritz ist alt*

Die Referenzmenge ist jeweils aus dem sprachlichen bzw. außersprachlichen Kontext oder durch Inferenzen zu ermitteln. Beispielsweise kann der Satz *Fritz ist groß* in bestimmten Kontexten als *Fritz ist groß im Vergleich zu seinen Mitschülern* paraphrasiert werden. Die stark kontextabhängige Interpretation vager, relativer Adjektive ist eines der Hauptprobleme für die Repräsentation von vagem Wissen in NSS (vgl. 4.4.3.3.).

Besonders drastische Fälle dieser Abhängigkeit von Bezugsgrößen bei der Verwendung relativer Adjektive ergeben sich in Sätzen wie *Eine große Ameise ist kleiner als ein kleiner Elefant*, der interpretiert wird als *eine Ameise, die verglichen mit einer typischen Ameise groß ist, ist kleiner als ein Elefant, der verglichen mit einem typischen Elefanten klein ist*.

Ist aus dem Kontext die Referenzmenge nicht herleitbar, so kann sie z.B. auch durch Präpositionalkonstruktionen mit *für* explizit eingeführt werden (z.B. *Fritz ist klein für einen Basketballspieler*, vgl. SANDEWALL71, HEMPHILL75).

BIERWISCH67 unterteilt die Klasse der relativen Adjektive weiter in positiv und negativ polarisierte Adjektive (Abk.: +POL, -POL). Positiv polarisierte Adjektive drücken aus, daß ein erwarteter Wert auf einer durch das Adjektiv angesprochenen Skala überschritten wird, negativ polarisierte Adjektive deuten darauf hin, daß ein erwarteter Wert unterschritten wurde.

BEISPIELE: +POL : *lang breit groß alt schnell*
 -POL : *kurz schmal klein jung langsam*

In sprachlichen Konstruktionen, die sich auf eine bestimmte Skala beziehen, wird nur das positiv polarisierte Adjektiv verwendet.

BEISPIEL: *Wie lange ist die Rennstrecke ? (nicht: wie kurz)*
Sie ist 10 km lang. (nicht: kurz)

Eine weitere Klasse von antonymen¹⁶ Adjektiven bilden Paare wie *gesund - krank*, die von EISENBERG76 als Qualitätsadjektive bezeichnet werden. BIERWISCH67 unterscheidet diese Klasse von der oben genannten dadurch, daß hier nicht ein typischer Wert über- oder unterschritten wird sondern eine Norm erfüllt (*gesund*) oder nicht erfüllt (*krank*) ist. Der Unterschied zwischen Qualitätsadjektiven und relativen Adjektiven spiegelt sich auch in der Anwendbarkeit negativer Inferenzen (vgl. CARBONELL/COLLINS73) wieder.

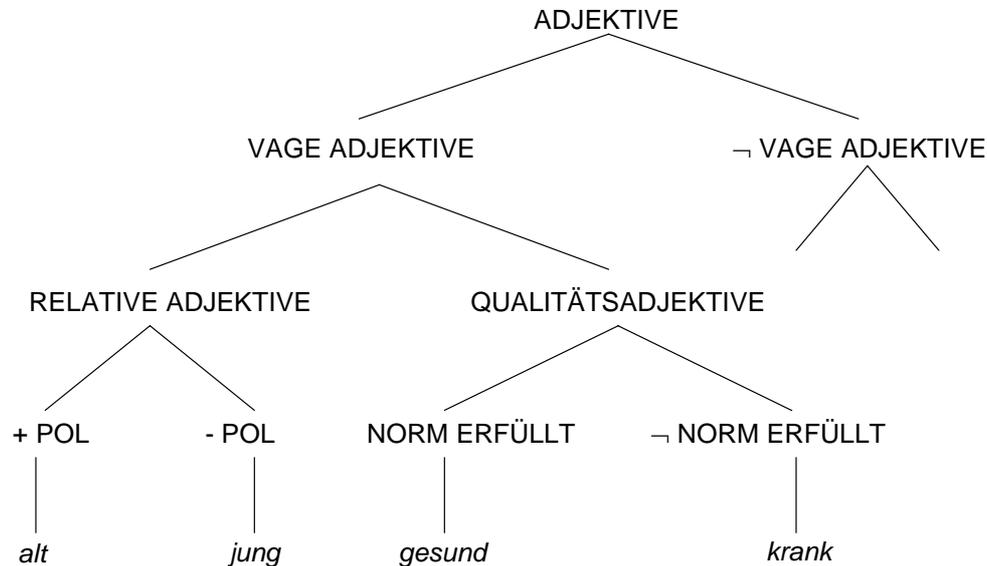
BEISPIEL: $(\neg \text{lang}(x) \Rightarrow \text{kurz}(x))$ ist keine gültige Inferenzregel
 $(\neg \text{gesund}(x) \Rightarrow \text{krank}(x))$ ist eine gültige Inferenzregel

¹⁵vgl. auch KATZ67, syn.: relationelle Adj. (SCHWIND76)

¹⁶Die Relation, in der ein Adjektivpaar wie *lang - kurz* steht, bezeichnet man als Antonymie.

Wir können also zumindest die in Fig. 16 zusammengestellten Klassen von Adjektiven unterscheiden.

Eine besondere Schwierigkeit im Bereich vager Adjektive ergibt sich dadurch, daß bestimmte Adjektive (z.B. *heiß*) abhängig vom Kontext als relatives Adjektiv (z.B. *ein heißer Tag*, höhere Temperatur als ein typischer Tag) oder als Qualitätsadjektiv (z.B. *heißer Kaffee*, Erwartungsnorm erfüllt) interpretiert werden können (vgl. LEISI 53).



FIGUR 16

Für die durch ein relatives Adjektiv angesprochene Wahrnehmungsdimension muß keine normierte, quantitative Skala sondern lediglich eine geordnete Wertemenge vorhanden sein. Bierwisch stellt fest:

... measurability does not depend on any scientific system of measure or on our scientific knowledge of the world.
(BIERWISCH67, S.11)

BARTSCH/VENNEMANN73 weisen darauf hin, daß die Existenz quantitativer Skalen lediglich von der Wissenschaftsgeschichte abhängt. So gibt es gegenwärtig keine normierten Skalen für Schönheit, Fleiß und Hunger aber für Intelligenz, Lautstärke und Temperatur. Vor einigen Jahrzehnten fehlten noch standardisierte Skalen für Lautstärke und Intelligenz. Das Fehlen einer quantitativen, normierten Skala verhindert also nicht den Vergleich. Häufig beziehen sich sprachliche Vergleiche nicht einmal auf die wenigen vorhandenen standardisierten Skalen (vergl. 3.2.).

BEISPIEL: Ein Satz wie *Das Radio ist sehr laut eingestellt* bezieht sich meist nicht auf die in Phon gemessene Lautstärke sondern auf eine an subjektiven Erwartungswerten orientierten Skala.

Die Frage, welche Skalen bei der Repräsentation von relativen Adjektiven im einzelnen sinnvoll sind, wird in dieser Arbeit, in welcher die mit einer formalen Repräsentation solcher Skalen verbundenen Probleme im Mittelpunkt stehen, nicht untersucht, sondern muß den zuständigen Einzeldisziplinen (z.B. Wahrnehmungspsychologie, Psycholinguistik) überlassen bleiben.

Prädikationen, die vage, relative Adjektive enthalten, können auch dann von einem Hörer verstanden werden, wenn der für relative Adjektive charakteristische Vergleich mit einer Erwartungsnorm nicht explizit durchgeführt werden kann, da in einem bestimmten Anwendungskontext keine Erwartungsnorm existiert.

BEISPIEL: Obwohl dem Hörer des Satzes *Fritz ist alt* nicht bekannt ist, wer Fritz ist, was die typische Lebensdauer der Klasse von Individuen ist, zu der Fritz gehört, und wer diesen Satz wann und wo geäußert hat, kann dieser Satz verstanden werden als *Das Alter, das ein als Fritz bezeichnetes*

Individuum hat, überschreitet das von dem Sprecher des Satzes für durchschnittlich gehaltene und auf die Referenzmenge von Fritz bezogene Lebensalter.

Diese für die Repräsentation von vagem Wissen wichtige Beobachtung führt uns zu der Unterscheidung zwischen Intension und Extension des vagen Prädikats *alt* und zur Trennung von referenzsemantischen und inhaltsemantischen Aspekten der Bedeutung (vgl. auch 3.2.1., 4.1.4., 4.4.3., 4.4.4.).

2.3.3. INHALTSSEMANTISCHE UND REFERENZSEMANTISCHE ASPEKTE DER VAGHEIT

Wir trennen zunächst die sprachexterne Bedeutung eines sprachlichen Ausdrucks, die seinen Bezug zu den im Weltmodell dargestellten Objekten, Sachverhalten und Prozessen charakterisiert und in einer Referenzsemantik zu behandeln ist, von der sprachinternen Bedeutung eines sprachlichen Ausdrucks, die als Relation des Ausdrucks zu anderen Ausdrücken aufgefaßt wird und im Rahmen einer Inhaltssemantik zu beschreiben ist.

Innerhalb der Referenzsemantik unterscheidet man zwischen der Intension und Extension eines sprachlichen Ausdrucks. Die Intension eines Satzes legt - vereinfacht dargestellt - die Bedingungen fest, unter denen ein Satz in verschiedenen Situationen, zu verschiedenen Zeitpunkten, an verschiedenen Orten usw. wahr ist, während die Extension eines Satzes dem Wahrheitswert eines Satzes entspricht. Als Extension eines Allgemeinnamens wie *Hund* wird in der Referenzsemantik die Menge von Individuen betrachtet, die in einem bestimmten Kontext als Hund bezeichnet werden. Als Extension eines n-stelligen Prädikats wird in einer einfachen Form der Referenzsemantik die Menge der n-Tupel betrachtet, die in einem gegebenen Kontext das Prädikat erfüllen.

Die Intension wird als eine Funktion aufgefaßt, die als Argument ein Bündel kontextbestimmender Angaben hat und als Wert die passende Extension liefert. Die kontextbestimmenden Angaben, wie Zeit, Ort, Sprecher, Hörer usw. der Äußerung werden oft als Index oder Referenzpunkt bezeichnet (vgl. LEWIS74, WUNDERLICH74). Die Extension ist also der Wert der Intension bei einem bestimmten Index.

Das letzte Beispiel in Abschnitt 2.3.2. kann jetzt so interpretiert werden, daß ein Verstehen des Beispielsatzes durch Wissen über die Intension von *alt* ermöglicht wurde. Die Extension und damit die Antwort auf die Frage, bei welcher Anzahl von Jahren das Prädikat *alt* zu welchem Grad zutrifft, kann erst angegeben werden, wenn die kontextbestimmenden Angaben, d.h. der Index, welcher das Argument für die Funktion 'Intension von *alt*' darstellt, zur Verfügung stehen. Falls aus dem Kontext hervorgeht, daß sich *alt* auf die durchschnittliche Lebensdauer von Menschen in Europa bezieht, können wir die Extension von *alt* z.B. als 'fuzzy set' darstellen (vgl. 3.2.1.).

Wir können also drei Bereiche einer semantischen Beschreibung vager Ausdrücke unterscheiden:

- Vagheit im Bereich der Inhaltssemantik (vgl. z.B. RIEGER76b).
Z.B. kann die intralexikalische Beziehung $Synonym(x,y)$ als vage Relation im Sinne der Theorie der 'fuzzy sets' aufgefaßt werden.
- Vagheit im Bereich der Referenzsemantik (vgl. LEWIS74, S.188ff, WUNDERLICH, S.267ff):
 - intensionale Vagheit (syn.: partielle Intension): für bestimmte entlegene oder neue Anwendungskontexte ist die Intension eines Ausdrucks nicht definiert.
 - referentielle Vagheit: der Wert der Intension bei einem bestimmten Index ist nicht genau bestimmt. Die Extension des Ausdrucks kann nur als 'fuzzy set' beschrieben werden. (vgl. 3.2.1.)

2.3.4. ZUSAMMENFASSUNG

Aus den linguistischen Untersuchungen vager sprachlicher Ausdrücke lassen sich folgende Forderungen an eine formale Repräsentation von vagem Wissen ableiten:

- Die Wirkung linguistischer Hecken, welche die Vagheit eines Ausdrucks verstärken oder abschwächen, muß in einer RS formal dargestellt werden. Eine semantische Analyse und

Generierung von linguistischen Hecken und anderen Gradpartikeln sollte ermöglicht werden. Auch die Bedeutung einer Kombination von Hecken muß formal dargestellt werden (vgl. 3.2.2., 4.6.2.2.).

- Es sind RK für limitierende Adverbiale, die eine Einschränkung des Gültigkeitsbereichs vager Prädikationen bewirken, vorzusehen (vgl. 4.6.2.3.).
- In den RK muß der mit relativen Adjektiven verbundene kontextabhängige Vergleich mit einer Referenzmenge berücksichtigt werden (3.2.1., 4.4.3.2., 4.4.3.3., 4.4.4.).
- Die Repräsentation von vagem Wissen in NSS setzt voraus, daß in einer Wissensquelle des Systems die Bedeutung vager sprachlicher Ausdrücke abgespeichert ist. Dabei sind die intensionalen und extensionalen Aspekte der Bedeutung zu berücksichtigen.
 - Bei der Darstellung der Bedeutung vager Ausdrücke muß zunächst der inhaltssemantische Aspekt, d.h. die Relation des Ausdrucks zu anderen Ausdrücken des Systems, beachtet werden. In der entsprechenden Wissensquelle, die meist als 'Lexikon' bezeichnet wird, werden relative, absolute und Qualitätsadjektive als solche klassifiziert, Antonymrelationen dargestellt und zwischen positiv und negativ polarisierten Adjektiven unterschieden.
 - Auch der intensionale Aspekt der Bedeutung vager Ausdrücke muß in einer RS dargestellt werden können (vgl. WOODS75, S.49ff). Es sind RK vorzusehen, in denen z.B. das Wissen codiert ist, daß *alt* als überschreiten eines für die jeweilige Referenzmenge festgelegten Altersbereich interpretiert werden kann (vgl. 4.1.4.)
 - Aus der RK für die extensionale Bedeutung des vagen Prädikats *alt* muß hervorgehen, bei welchem Alter von Fritz die Prädikation *alt (Fritz)* zu welchem Grad zutrifft (vgl. 3.2.1., 4.4.3.).

3 DIE THEORIE DER 'FUZZY SETS' UND IHRE ANWENDUNG AUF DIE REPRÄSENTATION VON VAGEM WISSEN

3.1. ZUR THEORIE DER 'FUZZY SETS'

3.1.1. EINFÜHRENDE ÜBERSICHT

Im folgenden werden einige grundlegende Begriffe der Theorie der 'fuzzy sets' eingeführt. Die Entwicklung dieser Theorie wurde 1965 von Lofti Zadeh eingeleitet (vgl. ZADEH65). Es werden dabei nur die Begriffe eingeführt, die für eine formale Darstellung von vagem Wissen relevant sind. Für eine umfassende Einführung in das Gebiet der 'fuzzy sets' sei auf die Lehrbücher von KAUFMANN75 und NEGOITA/RALESCU75 hingewiesen. Unter den zahlreichen Überblicksartikeln ist ZADEH73 einer der kompaktesten. Es sei hier mit Nachdruck auf die kommentierte Bibliographie von GAINES/KOHOUT77 hingewiesen, in der 1164 Titel zur Theorie und Anwendung der 'fuzzy sets' und eng benachbarter Gebiete zu finden sind.

Zunächst führen wir den Begriff 'fuzzy set' ein, für den in der deutschsprachigen Literatur die Bezeichnungen 'F-Menge' (DEMANT71), 'gefaserte Menge' (RIEGER74), 'unscharfe Menge' (RIEGER74/76a), 'gefrante Menge' (WUNDERLICH74) und 'verschwommene Menge' (DILGER76) zu finden sind. Dabei gehen wir von einer speziellen Darstellung einer klassischen Menge aus und verallgemeinern diese anschließend.

Sei U eine beliebige klassische Menge. Wir können jede Menge $A \subset U$ durch eine Zugehörigkeitsfunktion $\delta_A : U \rightarrow \{0,1\}$ darstellen, indem wir definieren:

$$\delta_A(x) \triangleq \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \notin A \end{cases}$$

Eine natürliche Erweiterung einer solchen Zugehörigkeitsfunktion besteht darin, als Wertebereich nicht nur $\{0, 1\}$ sondern auch z.B. $[0, 1] \subset \mathbb{R}$ zuzulassen.

DEFINITION: Sei U eine klassische Menge. Jede Zugehörigkeitsfunktion

$$\delta_{\underline{A}} : U \rightarrow [0, 1] \subset \mathbb{R}$$

definiert eine F-Menge (engl. fuzzy set) $\underline{A} \subset U$. U bezeichnen wir als Basismenge (engl. universe of discourse).

$\delta_{\underline{A}}$ ordnet jedem Element x der Basismenge¹⁷ U ein Zugehörigkeitswert $\delta_{\underline{A}}(x)$ zu, der den Grad angibt, mit dem x als Element der F-Menge \underline{A} zu gelten hat. Obwohl Zadeh die Semantik des Wertebereichs der Zugehörigkeitsfunktionen nicht formal definiert hat (vgl. GAINES76, S.628), ist der Gebrauch stets konsistent mit der natürlichen Ordnungsstruktur auf dem Einheitsintervall, d.h. $\delta_{\underline{A}}(x) = 0.4$ gibt einen geringeren Zugehörigkeitsgrad von x zu \underline{A} an, als das für y bei $\delta_{\underline{A}}(y) = 0.8$ der Fall ist.

Der einfache und plausible Grundgedanke der Theorie der F-Mengen besteht also darin, die Zugehörigkeit eines Individuums zu einer F-Menge graduell anzugeben. Das elementare mathematische Konzept der F-Menge ist oft besser geeignet zur formalen Modellbildung im Bereich menschlicher Kognition und Kommunikation als das Konzept der Menge, da, wie wir aus philosophischen, psychologischen und linguistischen Untersuchungen erkennen konnten, daß die Vagheit und die damit verbundene graduelle Abstufung der Kategorienzugehörigkeit für den genannten Bereich konstitutiv sind.

Die Theorie der F-Mengen ist allerdings nur als erster Schritt bei der Entwicklung formaler Beschreibungsmethoden für vage Konzepte zu betrachten. Außerdem kann natürlich jede Struktur, die durch F-Mengen beschrieben werden kann, auch durch zahlreiche andere mathematische Konzepte beschrieben werden. F-Mengen sind also lediglich eine neue Sichtweise, ein 'Kuhn'sches Paradigma', aus dem heraus sich allerdings ähnlich wie z.B. in der Graphentheorie durchaus neue Problemlösungen ergeben können.

Ausgehend von den Ideen des Systemtheoretikers Zadeh wurden seit 1965 zahlreiche neue Modelle im Bereich der theoretischen Informatik, der Mathematik, der Linguistik und Psychologie entwickelt und viele praktische Probleme der Mustererkennung, der Entscheidungstheorie, des Information Retrieval sowie aus Gebieten der Wirtschafts- und Sozialwissenschaft, der Biologie und der Medizin durch systemtheoretische Analysen unter Verwendung von F-Mengen gelöst (vgl. GAINES/KOHOUT77).

Die jährliche Zuwachsrate an Literatur, die 40% beträgt, unterstreicht das derzeitige Interesse an dieser Theorie und veranlaßte GAINES/KOHOUT77 dazu, das laufende Jahrzehnt als 'the fuzzy decade' zu bezeichnen.

Wir werden in Kapitel 4 sehen, daß die Anzahl der Arbeiten im Bereich der KI, die sich mit einer Anwendung von Konzepten aus der Theorie der F-Mengen auf die Repräsentation von vagem Wissen beschäftigen, noch sehr gering ist (vgl. GAINES/KOHOUT77, S. 20).

Auch zu einer anderen möglichen Anwendung der Theorie der F-Mengen bei der Konstruktion von NSS, dem sog. 'fuzzy parsing', sind in der KI-Literatur erst wenige Hinweise zu finden (vgl. MINSKY75, BROWN/BURTON75, S.330, GRISHMAN76). Chomsky hatte bereits 1965 gefordert, daß in einer formalen Grammatik zur Beschreibung der NS unterschiedliche Grammatikalitätsgrade erfaßt werden sollten (vgl. CHOMSKY73, S. 188ff). Auch LAKOFF73 skizziert eine F-Grammatik, in der jede Produktion nur zu einem gewissen Grad anwendbar ist und Wortklassen als F-Mengen beschrieben werden. Inzwischen wurden in der theoretischen Informatik Modelle formaler Grammatiken untersucht, durch die eine formale Sprache beschrieben werden kann, die aus einer F-

¹⁷ Da jede F-Menge in Bezug auf eine Basismenge definiert ist, findet man auch die Bezeichnung 'fuzzy subset' für F-Mengen (vgl. KAUFMANN75).

Menge von Zeichenketten über einem Grundalphabet besteht (vgl. GAINES/KOHOUT77).

Obwohl in mehreren Arbeiten zur KI gegenwärtig gefordert wird, daß NSS robust¹⁸ sein sollen, d.h. daß kleine Abweichungen z.B. in der Syntax der Eingabe vom System akzeptiert werden sollen (vgl. SHAKET76, GOGUEN75/76), wurde bisher noch kein Syntaxanalysealgorithmus, der auf der Basis einer F-Grammatik arbeitet, in ein NSS integriert.

Nicht weiter untersucht werden in dieser Arbeit formale Konzepte wie F-Grammatiken, F-Automaten und F-Algorithmen (vgl. z.B. NEGOITA/RALESCU75) und ihr Bezug zu Problemen, die sich in NSS für eine syntaktische und semantische Analyse von Eingaben ergeben, falls verschiedene Grade der Akzeptabilität unterschieden werden. Wir müssen uns hier auf die Anwendung der Theorie der F-Mengen bei der Repräsentation von vagem Wissen beschränken.

Bevor wir im folgenden einige Grundbegriffe der Theorie der F-Mengen einführen, soll hier darauf hingewiesen werden, daß trotz einiger formaler Ähnlichkeiten die Theorie der F-Mengen keinesfalls als eine spezielle Form der Wahrscheinlichkeitstheorie aufgefaßt werden kann. Im Gegensatz zur Wahrscheinlichkeitstheorie werden Elemente aus $[0,1] \subset \mathbb{R}$ in der Theorie der F-Mengen als Grade der Zugehörigkeit zu einer F-Menge in einer deterministischen Umgebung aufgefaßt, in der Wahrscheinlichkeitstheorie geht es dagegen um die Unsicherheit in Bezug auf das Auftreten exakt definierter Ereignisse.

3.1.2. GRUNDLEGENDE DEFINITIONEN

Im folgenden Abschnitt werden grundlegende Begriffe der Theorie der F-Mengen formal definiert. Dabei können wir auf eine Veranschaulichung der Definitionen durch Beispiele verzichten, weil im nächsten Abschnitt (vgl. 3.2.) mehrere Beispiele ausführlich diskutiert werden, in denen die hier definierten Begriffe angewendet werden.

Für alle folgenden Definitionen gilt $\underline{A}, \underline{B} \subset U$.

DEFINITION: Eine F-Menge ist vom Typ $n \in \mathbb{N}$, genau dann, wenn der Wertebereich ihrer Zugehörigkeitsfunktion aus F-Mengen vom Typ $n-1$ besteht. Die Zugehörigkeitsfunktionen von F-Mengen vom Typ 1 haben den Wertebereich $[0,1] \subset \mathbb{R}$.

DEFINITION: X ist die Trägermenge $T(\underline{A})$ einer F-Menge \underline{A}
 $:\Leftrightarrow X = T(\underline{A}) = \{y \in U \mid \delta_{\underline{A}}(y) > 0\}$

DEFINITION: \underline{A} ist eine singuläre F-Menge $:\Leftrightarrow |T(\underline{A})| = 1$

DEFINITION: $y \in \underline{A}$ ist ein Übergangspunkt $:\Leftrightarrow \delta_{\underline{A}}(y) = 0.5$

SCHREIBKONVENTIONEN:

Eine F-Menge \underline{A} mit endlicher Trägermenge $Y = \{y_1, y_2, \dots, y_n\}$, $n \in \mathbb{N}$

schreiben wir als $\underline{A} = \{ \delta_{\underline{A}}(y_1) / y_1, \delta_{\underline{A}}(y_2) / y_2, \dots, \delta_{\underline{A}}(y_n) / y_n \}$.

Eine F-Menge \underline{A} , die keine endliche Trägermenge hat, schreiben wir als

$$\underline{A} = \int_U \delta_{\underline{A}}(y) / y$$

wobei \int_U die Vereinigung aller singularen F-Mengen $\delta_{\underline{A}}(y) / y$, $y \in U$ sei.

¹⁸ (engl. robustness im Gegensatz zur 'brittleness', vgl. BOBROWetal.76)

DEFINITION: Als α -Ausschnitt (engl. α -level set) bezeichnen wir die Menge

$$A_\alpha = \{y \in U \mid \delta_{\tilde{A}}(y) \geq \alpha \in [0,1] \subset \mathbb{R}\}, \text{ wobei } \tilde{A} \text{ eine F-Menge ist.}$$

DEFINITION: Die Höhe von \tilde{A} (Abk.: $h(\tilde{A})$) sei das Supremum von $\delta_{\tilde{A}}$ über U ,

d.h.
$$h(\tilde{A}) = \text{SUP}_U \delta_{\tilde{A}}(U).$$

DEFINITION: \tilde{A} ist normalisiert $:\Leftrightarrow h(\tilde{A}) = 1$

DEFINITION: $\tilde{A} = \tilde{B} \quad :\Leftrightarrow \forall x \in U : \delta_{\tilde{A}}(x) = \delta_{\tilde{B}}(x)$

DEFINITION: \tilde{A} ist eine F-Teilmenge von \tilde{B} (kurz: $\tilde{A} \subset \tilde{B}$)

$$:\Leftrightarrow \forall x \in U : \delta_{\tilde{A}}(x) \leq \delta_{\tilde{B}}(x)$$

DEFINITION: Gegeben sei U , ein kartesisches Produkt aus n Basismengen, d.h.

$U = U_1 \times U_2 \times \dots \times U_n$. Eine n -stellige F-Relation ist eine F-Menge $\tilde{R} \subset U_1 \times U_2 \times \dots \times U_n$

$$\tilde{R} = \int_{U_1 \times U_2 \times \dots \times U_n} \delta_{\tilde{R}}(y_1, y_2, \dots, y_n) / (y_1, y_2, \dots, y_n) \\ (y_1, y_2, \dots, y_n) \in U$$

DEFINITION: Sei KN eine endliche Menge von Knoten, KA eine endliche Menge von Kanten. Ein endlicher, gerichteter und markierter F-Graph (Abk.: F-Graph) ist eine F-Relation

$$\tilde{R} = \int_{KN \times KA \times KN} \delta_{\tilde{R}}(x_1, x_2, x_3) / (x_1, x_2, x_3) \\ (x_1, x_2, x_3) \in KN \times KA \times KN$$

DEFINITION: Das Komplement $\overline{\tilde{A}}$ einer F-Menge \tilde{A} sei

$$\overline{\tilde{A}} \hat{=} \int_U (1 - \delta_{\tilde{A}}(y)) / y$$

DEFINITION: Die Vereinigung $\tilde{A} \cup \tilde{B}$ von F-Mengen sei

$$\tilde{A} \cup \tilde{B} \hat{=} \int_U \text{MAX}(\delta_{\tilde{A}}(y), \delta_{\tilde{B}}(y)) / y$$

DEFINITION: Der Durchschnitt $\tilde{A} \cap \tilde{B}$ von F-Mengen sei

$$\tilde{A} \cap \tilde{B} \hat{=} \int_U \text{MIN}(\delta_{\tilde{A}}(y), \delta_{\tilde{B}}(y)) / y$$

DEFINITION: Das Produkt $\underline{A} \bullet \underline{B}$ von F-Mengen sei

$$\underline{A} \bullet \underline{B} \hat{=} \int_U \delta_{\underline{A}}(y) \bullet \delta_{\underline{B}}(y) / y$$

DEFINITION: Sei $\alpha > 0$. Die Potenzierung einer F-Menge durch einen Exponenten α ist definiert durch:

$$\underline{A}^\alpha \hat{=} \int_U (\delta_{\underline{A}}(y))^\alpha / y$$

DEFINITION: Sei $\alpha \in \mathbb{R}^+$, $\alpha \bullet h(\underline{A}) \leq 1$. die Multiplikation einer F-Menge \underline{A} mit einem Skalar α ist definiert durch:

$$\alpha \bullet \underline{A} \hat{=} \int_U \alpha \bullet \delta_{\underline{A}}(y) / y$$

DEFINITION:

$$\underline{A} < \text{oder} > \underline{B} \hat{=} \int_U \delta_{\underline{A}}(y) + \delta_{\underline{B}}(y) - \delta_{\underline{A}}(y) \bullet \delta_{\underline{B}}(y) / y$$

DEFINITION: Seien $\underline{A}_1, \underline{A}_2, \dots, \underline{A}_n$ F-Mengen mit der Basismenge U und w_1, w_2, \dots, w_n

$$0 \leq w_i \leq 1, \sum_{i=1}^n w_i = 1 \text{ Gewichte.}$$

Eine gewichtete Kombination \underline{A} von $\underline{A}_1, \underline{A}_2, \dots, \underline{A}_n$ ist definiert durch:

$$\underline{A} \hat{=} \int_U \left(\sum_{i=1}^n w_i \bullet \delta_{\underline{A}_i}(y) \right) / y$$

3.2. DIE FORMALE DARSTELLUNG VON VAGEM WISSEN DURCH 'FUZZY SETS'

3.2.1. DIE REPRÄSENTATION DER EXTENSIONALEN BEDEUTUNG VAGER SPRACHLICHER AUSDRÜCKE

Im folgenden wird gezeigt, wie die Bedeutung vager sprachlicher Terme mit Hilfe von F-Mengen repräsentiert werden kann und welche Operationen über diesen RK durchgeführt werden können. Dabei gehen wir von Zadehs 'linguistischem Ansatz' (vgl. ZADEH71/73/76) aus, der besagt, daß

words or sentences are used in place of numbers to describe phenomena which are too complex or too ill-defined to be susceptible of characterization in quantitative terms
(ZADEH76, S.250)

the key elements in human thinking are not numbers, but labels of fuzzy sets.
(ZADEH73, S.28)

Es muß darauf hingewiesen werden, daß die im folgenden diskutierten RK jeweils nur bestimmte Aspekte der komplexen Semantik von vagen sprachlichen Ausdrücken (vgl. 2.3.) und den auf ihnen operierenden linguistischen Hecken berücksichtigen. Beim Versuch, die Komplexitätsbarriere, die mit einer formalen Darstellung vager Terme verbunden ist, zu durchbrechen, sind zunächst viele Vereinfachungen notwendig.

Wir werden die intensionalen Aspekte der Bedeutung zunächst unberücksichtigt lassen (vgl. z.B. 4.1.4.) und lediglich eine stark vereinfachte extensionale Semantik skizzieren.

Die Extension eines sprachlichen Ausdrucks t wird im folgenden als $M(t)$ geschrieben. $M(t)$ wird als F-Menge \underline{A} über einer Basismenge U beschrieben.

BEISPIEL: Sei $U = \{\text{FIAT500, GOLF, OPEL-REKORD, MERCEDES, ROLLS-ROYCE}\}$
Die Extension des vagen Ausdrucks *Mittelklassewagen* kann als F-Menge definiert werden (vgl. auch RIEGER76b).

$$M(\text{MITTELKLASSEWAGEN}) = \{0/\text{FIAT500}, 0.7/\text{GOLF}, 1/\text{OPEL-REKORD}, 0.3/\text{MERCEDES}, 0/\text{ROLLS-ROYCE}\}$$

$M(\text{MITTELKLASSEWAGEN})$ ist eine normalisierte F-Menge vom Typ 1. Das Element *Opel-Rekord* kann als Prototyp betrachtet werden (vgl. 2.2.2.).

BEISPIEL: Es sei $U_1 = \{\text{DORIS, HANS, PAUL, INA}\}$ und $U_2 = \{0, 0.1, 0.2, \dots, 1\}$.
 U_2 wird als Intensitätsskala interpretiert. Außerdem seien folgende Ausdrücke als F-Mengen definiert:

$$\begin{aligned} M(\text{WENIG}) &= \{0.5/0.2, 0.7/0.3, 1/0.4, 0.7/0.5, 0.5/0.6\} \\ M(\text{MITTEL}) &= \{0.5/0.4, 0.7/0.5, 1/0.6, 0.7/0.7, 0.5/0.8\} \\ M(\text{HOCH}) &= \{0.5/0.7, 0.7/0.8, 0.9/0.9, 1/1\} \end{aligned}$$

Das vage Konzept *agil* definieren wir als F-Menge vom Typ 2, wobei wir als Extension der Personennamen die Menge des gespeicherten Wissens über die Person auffassen.

$$M(\text{AGIL}) = \{\text{MITTEL/DORIS}, \text{WENIG/HANS}, \text{HOCH/PAUL}, \text{WENIG/INA}\}$$

BEISPIEL: Sei $U_1 = \{\text{UDO, ULF}\}$, $U_2 = \{\text{FRITZ, PETER}\}$. Die Extension von *ähnelt* bei einem bestimmten Index (vgl. 2.3.3.) kann als F-Relation $M(\text{ÄHNELT}) \subset U_1 \times U_2$ dargestellt werden :

$$M(\text{ÄHNELT}) = \{0.8/(\text{UDO, FRITZ}), 0.4/(\text{UDO, PETER}), 0.4/(\text{ULF, FRITZ}), 0.24/(\text{ULF, PETER})\}$$

Durch folgende Operation kann diese nicht-normalisierte F-Menge normalisiert werden:

$$\text{NORM}(\underline{A}) \triangleq h(\underline{A})^{-1} \bullet \underline{A}, \quad h(\underline{A}) \neq 0$$

$$\text{NORM}(M(\text{ÄHNELT})) = \{1/(\text{UDO, FRITZ}), 0.5/(\text{UDO, PETER}), 0.5/(\text{ULF, FRITZ}), 0.3/(\text{ULF, PETER})\}$$

In den bisherigen Beispielen wurde die Zugehörigkeitsfunktion ohne Bezug auf Eigenschaften der Elemente der Basismenge definiert. Im folgenden werden Beispiele betrachtet, in denen die Zugehörigkeitsfunktion in Abhängigkeit von Eigenschaften der betrachteten Elemente definiert wird.

Um eine möglichst einfache Darstellung von Zugehörigkeitsfunktionen zu erreichen, bedienen wir uns standardisierter Funktionen, die als Definitionsbereich \mathbb{R} und als Wertebereich $[0,1] \subset \mathbb{R}$ haben. Die beiden Funktionen SFUNK (monotone Funktion, S-förmige Kurve) und PFUNK (puls- bzw.

glockenförmige Kurve) können über Parameter so eingerichtet werden, daß viele unterschiedliche Zugehörigkeitsfunktionen durch sie repräsentiert werden können (vgl. ZADEH76, S.256).

DEFINITION: Es seien $NULL, EINS \in \mathbb{R}$, und $HALB = \frac{NULL + EINS}{2}$; dann definieren wir SFUNK durch :

$$SFUNK(x, NULL, EINS) \triangleq \begin{cases} 0 & \text{falls } x \leq NULL \\ 2 \cdot \left(\frac{x - NULL}{EINS - NULL} \right)^2 & \text{falls } NULL \leq x \leq HALB \\ 1 - 2 \cdot \left(\frac{x - EINS}{EINS - NULL} \right)^2 & \text{falls } HALB \leq x \leq EINS \\ 1 & \text{falls } x \geq EINS \end{cases}$$

Durch HALB wird der Übergangspunkt einer durch SFUNK beschriebenen F-Menge festgelegt.

$\{x \in U \mid x \geq NULL\}$ ist die Trägermenge und $\{x \in U \mid x \geq EINS\}$ der 1-Ausschnitt einer durch SFUNK definierten F-Menge.

DEFINITION: Es seien BREITE, SCHEITEL $\in \mathbb{R}$

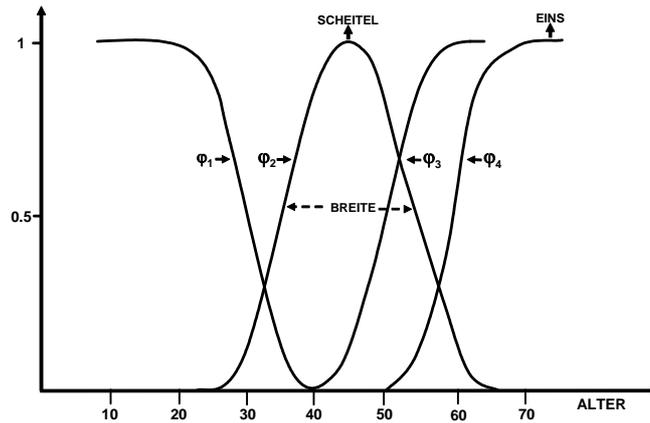
$$PFUNK(x, BREITE, SCHEITEL) \triangleq \begin{cases} SFUNK(x, SCHEITEL - BREITE, SCHEITEL) & \text{falls } x \leq SCHEITEL \\ 1 - SFUNK(x, SCHEITEL, SCHEITEL + BREITE) & \text{falls } x \geq SCHEITEL \end{cases}$$

Durch BREITE wird der Abstand zwischen den beiden Übergangspunkten und durch SCHEITEL der 1-Ausschnitt der durch PFUNK definierten F-Menge festgelegt.

BEISPIEL: Die extensionale Bedeutung von *jung*, *mittel alt*, *älter* und *alt* bei einem bestimmten Index kann als F-Menge in Abhängigkeit vom Alter der Lebewesen, auf die ein solches vages Prädikat zutrifft, definiert werden (vgl. Fig. 17).

M(X:Y) deutet an, daß Y die durch den Kontext bestimmte Referenzmenge zu X ist (vgl. 2.3.2.).

$$\begin{array}{ll} M(JUNG:MANN) = 1 - SFUNK(ALTER, 20, 40) & [\varphi_1] \\ M(MITTEL-ALT:MANN) = PFUNK(ALTER, 20, 45) & [\varphi_2] \\ M(ÄLTER:MANN) = SFUNK(ALTER, 40, 60) & [\varphi_3] \\ M(ALT:MANN) = SFUNK(ALTER, 50, 70) & [\varphi_4] \end{array}$$



FIGUR 17

Für ein Alter von 53 Jahren ergeben sich dabei folgende Werte:

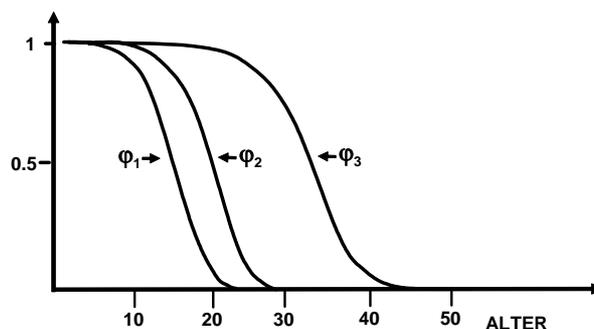
$$\begin{aligned} \delta_{M(JUNG:MANN)}(53) &= 0 \\ \delta_{M(MITTEL-ALT:MANN)}(53) &= 0.68 \\ \delta_{M(ÄLTER:MANN)}(53) &= 0.75 \\ \delta_{M(ALT:MANN)}(53) &= 0.05 \end{aligned}$$

Wenn in einem NSS die Aussage *Fritz ist 53 Jahre alt* in der Wissensbasis abgespeichert ist, so kann auf die Frage *Wie alt ist Fritz?* z.B. *Fritz ist ein Mann von mittlerem Alter* oder *Fritz ist ein älterer Mann* als vage Antwort ausgegeben werden.

Die Bedeutungsüberlappung vager sprachlicher Terme wie *mittel alt* und *älter*, die oft als Charakteristikum vager Ausdrücke angesehen wird (vgl. KOOIJ71, S.119ff) kann von einem NSS als einfache 'stilistische' Variationsmöglichkeit benutzt werden.

In Abschnitt 2.3.2. haben wir festgestellt, daß die extensionale Bedeutung von relativen Adjektiven wie *jung* von der durch den Kontext bestimmten Referenzmenge abhängt. Dieser Aspekt der Bedeutung vager Adjektive bleibt in den Arbeiten von LAKOFF75 und ZADEH71/72/74/76 unberücksichtigt. Im einfachsten Fall kann diese Abhängigkeit dadurch dargestellt werden, daß Wissen über die Referenzmenge, das in Form von typischen Beispielen in einer Wissensquelle codiert ist, die Parameter *NULL* und *EINS* der Funktion SFUNK bzw. BREITE und SCHEITEL der Funktion PFUNK beeinflusst (vgl. 4.4.3.3.).

BEISPIEL: $M(JUNG:MÄDCHEN) = 1 - SFUNK(ALTER, 5, 25)$ [φ₁]
 $M(JUNG:SPORTLER) = 1 - SFUNK(ALTER, 10, 30)$ [φ₂]
 $M(JUNG:PARLAMANTARIER) = 1 - SFUNK(ALTER, 21, 45)$ [φ₃]



FIGUR 18

Die Semantik natürlichsprachlicher Junktoren wie *nicht*, *und*, *oder* wird durch F-Mengen-theoretische Operationen über den F-Mengen repräsentiert, die der Semantik der durch die Junktoren verknüpften vagen Terme entsprechen.

In der Literatur (vgl. GAINES/KOHOUT77) wurden sehr unterschiedliche Operationen zur Explikation der Semantik von NS-Junktoren vorgeschlagen. Fig. 19 enthält eine Auswahl alternativer Definitionen, zwischen denen abhängig vom jeweiligen Anwendungskontext gewählt werden kann.

Durch diese Art der Darstellung von Junktoren gelingt es, die Bedeutung eines zusammengesetzten Terms aus der Bedeutung seiner Komponenten herzuleiten und damit eine der Anforderungen an eine formale Semantik zu erfüllen.

BEISPIEL:

$$\begin{aligned} \delta_{M(JUNG:MANN)}(35) &= 0.13 \\ \delta_{M(nicht\ JUNG:MANN)}(35) &= 0.87 \\ \delta_{M(JUNG\ und_1\ MITTEL-ALT:MANN)}(35) &= 0.13 \\ \delta_{M(MITTEL-ALT:MANN)}(35) &= 0.5 \\ \delta_{M(JUNG\ oder_1\ MITTEL-ALT:MANN)}(35) &= 0.5 \end{aligned}$$

NS-JUNKTOR	OPERATION AUF F-MENGE
<i>nicht</i> x	$\bar{M}(x)$
x <i>und</i> ₁ y	$M(x) \cap M(y)$
x <i>und</i> ₂ y	$M(x) \bullet M(y)$
x <i>oder</i> ₁ y	$M(x) \cup M(y)$
x <i>oder</i> ₂ y	$M(x) < oder > M(y)$

FIGUR 19

Oft ist die Semantik eines komplexen sprachlichen Terms t aus der Semantik mehrerer elementarer Terme t_1, t_2, \dots, t_n zusammengesetzt:

$$\delta_{M(t)} = f(\delta_{M(t_1)}, \delta_{M(t_2)}, \dots, \delta_{M(t_n)})$$

Die F-Mengen $M(t_1), M(t_2), \dots, M(t_n)$ werden dann als Bedeutungskomponenten bezeichnet.

BEISPIEL: $M(KONZILIANZ)$ kann als gewichtete Kombination dargestellt werden.

$$\begin{aligned} \delta_{M(KONZILIANZ)} &= w_1 \cdot \delta_{M(ENTGEGENKOMMEND)} + w_2 \cdot \delta_{M(FREUNDLICH)} + \\ &w_3 \cdot \delta_{M(VERSÖHNLICH)} + w_4 \cdot \delta_{M(VERBINDLICH)} \quad 0 \leq w_i \leq 1 \end{aligned}$$

3.2.2. DIE DARSTELLUNG VON LINGUISTISCHEN HECKEN

Hecken (vgl. 2.3.1.) können im Rahmen einer extensionalen Semantik als unäre Operatoren dargestellt werden, welche die Bedeutung der als Operanden fungierenden F-Mengen verändern. Wir unterscheiden zwei Klassen von Hecken:

DEFINITION: Eine Hecke ist vom Typ 1 genau dann, wenn sie sich als Operator über der Zugehörigkeitsfunktion des Operanden darstellen läßt.

DEFINITION: Eine Hecke ist vom Typ 2 genau dann, wenn sie sich nur als Operator über den einzelnen Bedeutungskomponenten des Operanden darstellen läßt.

BEISPIELE: Hecken vom Typ 1: *sehr, mehr oder weniger, ziemlich*
 Hecken vom Typ 2: *grob gesprochen, im Grunde genommen, typisch*

3.2.2.1. Linguistische Hecken vom Typ 1

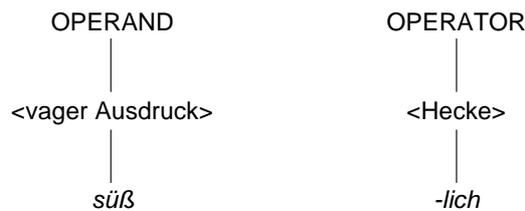
Linguistische Hecken vom Typ 1 treten meist als Präfix-Operatoren auf und werden als Präzisions- bzw. Unschärfeoperatoren interpretiert.

BEISPIEL:



In einigen Fällen können diese Hecken auch durch Synonyme ersetzt werden, die in Postfix-Form auftreten (vgl. auch NEUHAUS71, S.86). Beispielsweise kann die Hecke *mehr oder weniger* durch das Adjektivsuffix *-lich* ersetzt werden (u.a. in *rötlich, länglich, säuerlich, süßlich*).

BEISPIEL:



Eine semantische Analyse und Generierung von Hecken in Postfix-Form setzt also die Integration morphologischer Algorithmen (vgl. z.B. HOEPPNER77) in NSS voraus. Meist treten Hecken jedoch als Lexeme oder lexikalisierte Wortgruppen wie *mehr oder weniger* (vgl. auch BOLINGER72) auf und können daher in einem NSS durch einen Zugriff auf eine Wissensquelle, in der RK für die extensionale Bedeutung von Hecken abgespeichert sind, auf einfache Weise analysiert oder generiert werden.

Eine Hecke h vom Typ 1 kann als Abbildung $h: \Phi(U) \rightarrow \Phi(U)$ aufgefaßt werden, wobei $\Phi(U)$ die Menge aller F-Mengen über einer Basismenge U bezeichnet. Bei der formalen Darstellung von Hecken bedienen wir uns folgender elementarer Präzisions- und Unschärfeoperatoren (vgl. auch ZADEH72, S.21ff, LAKOFF75, S.243ff, SHAKET76, S.76ff):

Präzisionsoperatoren: Sei $A \subset U$

$CON(A) \triangleq \tilde{A}^2$; verringert kleine Zugehörigkeitswerte relativ stärker als große
 (Konzentration)

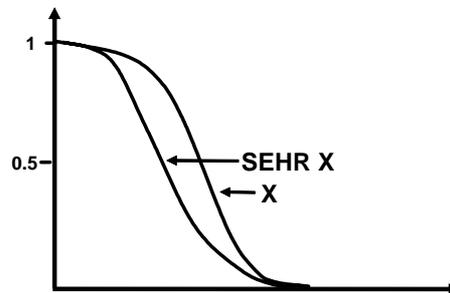
$INT(\delta_A) \triangleq SFUNK(\delta_A, 0,1)$; vergrößert Zugehörigkeitswerte über 0.5
 (Kontrast Intensivierung) und verkleinert diejenigen unter 0.5

Unschärfeoperator:

$DIL(A) \triangleq \sqrt{\tilde{A}}$; Gegenteil des Effekts von CON
 (Erweiterung, engl. dilation)

Ausgehend von einer kleinen Zahl elementarer, nichtlinearer Operatoren kann durch Kombination die Wirkung vieler linguistischer Hecken in bestimmten Anwendungskontexten beschrieben werden:

$$M(\text{sehr } x) \hat{=} \text{CON}(M(x)) \text{ (vgl. Fig. 20)}$$



FIGUR 20

$$M(\text{mehr – oder – weniger } x) \hat{=} \text{DIL}(M(x))$$

BEISPIEL:

$$\delta_{M(\text{SÜSS:OBST})}(\text{BANANE}) = 0.25, \delta_{M(\text{mehr-oder-weniger SÜSS:OBST})}(\text{BANANE}) = 0.5$$

$$\delta_{M(\text{SÜSS-lich:OBST})}(\text{BANANE}) = 0.5$$

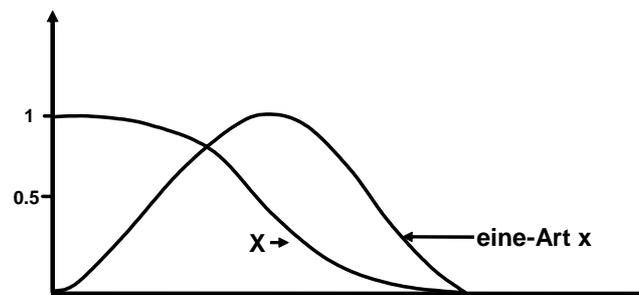
Zusammengesetzte Hecken vom Typ 1 sind z.B.:

$$M(\text{ziemlich } x) \hat{=} \text{INT}(\text{CON}(M(x)))$$

$$M(\text{nahezu } x) \hat{=} \text{INT}(M(x)) \cap \text{INT}(\text{CON}(M(x)))$$

$$M(\text{eine – Art } x) \hat{=} \text{NORM}(\text{CON}^2(M(x))) \cap \text{DIL}(M(x))$$

(vgl. Fig. 21)



FIGUR 21

Die Wirkung der Hecke *eine Art* besteht darin, den Zugehörigkeitswert für zentrale Elemente der Kategorie, auf die sie angewandt wird, zu senken und für periphere Elemente zu erhöhen (vgl. Fig. 21).

Aus der Darstellung von Hecken als unäre Operatoren ergibt sich, daß Hecken auch miteinander kombiniert werden können (vgl. auch 2.3.1.)

BEISPIEL: $M(\text{sehr sehr } x) \hat{=} M(x)^4$

An dieser Stelle soll nochmals darauf hingewiesen werden, daß dies nur annähernde Beschreibungen sind und daß

all these functions may well be fuzzy themselves, so talk of an exact multiple or power in the equations may make no sense.
(LAKOFF75, S.244)

Oft wird auch eine Dekomposition von linguistischen Hecken, d.h. eine Zerlegung komplexer Hecken in mehrere elementare Hecken, vorgeschlagen. Als elementare Hecken werden sog. künstliche Hecken, die keine direkte Entsprechung in der NS haben, häufig zur Dekomposition verwendet.

BEISPIEL: Die künstliche Hecke 'plus' wird definiert als $M(\text{plus } x) \hat{=} M(x)^{1.25}$

Zadeh schlägt folgende Dekomposition der Hecke *höchst* vor:

$$M(\text{höchst } x) \hat{=} M(\text{plus sehr } x) \quad (\text{vgl. ZADEH72, S.26})$$

In Abschnitt 4.3.2.2. wird gezeigt, wie Hecken vom Typ 1 durch prozedurale RK in der mehrwertigen PLANNER-artigen Programmiersprache FUZZY dargestellt werden können.

Eine F-Mengen-theoretische Behandlung der Hecken vom Typ 1 wird ansatzweise bereits in einigen NSS vorgenommen (vgl. SHAKET76, WAHLSTER/ V.HAHN76, RHODES/KLINGER77).

3.2.2.2. Linguistische Hecken vom Typ 2

Bei der Beschreibung der linguistischen Hecken vom Typ 2 im Rahmen der Theorie der F-Mengen sind zwei Ansätze zu unterscheiden, welche von ZADEH72 bzw. LAKOFF75 entwickelt wurden.

ZADEH72 schlägt vor, Hecken vom Typ 2 als Operatoren über den Gewichten der einzelnen Bedeutungskomponenten eines vagen Konzeptes (vgl. 3.2.1.) darzustellen. Nach diesem Verfahren kann z.B. die Hecke *im Grunde genommen* als Operator beschrieben werden, der diejenigen Bedeutungskomponenten, die hohe Gewichte haben, stärker bewertet und andere weniger wichtige Bedeutungskomponenten weitgehend unberücksichtigt läßt.

BEISPIEL: Sei

$$M(\text{ENTGEGENKOMMEND}) = \underset{\sim_1}{A}, \quad M(\text{FREUNDLICH}) = \underset{\sim_2}{A}, \quad M(\text{VERSÖHNLICH}) = \underset{\sim_3}{A},$$

$$M(\text{VERBINDLICH}) = \underset{\sim_4}{A}.$$

Die Bedeutung des vagen Ausdrucks *konzilient* sei dargestellt als

$$\delta_{M(\text{KONZILIENT})} = 0.4 \cdot \delta_{\underset{\sim_1}{A}} + 0.3 \cdot \delta_{\underset{\sim_2}{A}} + 0.2 \cdot \delta_{\underset{\sim_3}{A}} + 0.1 \cdot \delta_{\underset{\sim_4}{A}}$$

$$\text{Sei } \delta_{\underset{\sim_1}{A}}(\text{FRITZ}) = 0.9, \quad \delta_{\underset{\sim_2}{A}}(\text{FRITZ}) = 0.8, \quad \delta_{\underset{\sim_3}{A}}(\text{FRITZ}) = 0.9, \quad \delta_{\underset{\sim_4}{A}}(\text{FRITZ}) = 0.2$$

$$\text{Dann gilt } \delta_{M(\text{KONZILIENT})}(\text{FRITZ}) = 0.8$$

Sei *h* die Hecke *im Grunde genommen*. Die Bedeutung von *im Grunde genommen konzilient* wird dann repräsentiert als:

$$\delta_{M(h \text{ KONZILIENT})}(x) = \sum_{i=1}^4 w_i^* \cdot \delta_{\underset{\sim_i}{A}}, \quad \text{wobei}$$

$$w_i^* = \frac{\text{NORM}(w_i)^2}{\sum_{i=1}^4 \text{NORM}(w_i)^2} \quad \text{gilt.}$$

$$(w_i \in W, \text{NORM}(w_i) \hat{=} (\text{SUP}(W))^{-1} \cdot w_i)$$

Wir erhalten also:

$$\delta_{M(h \text{ KONZILIENT})}(\text{FRITZ}) = 0.53 \cdot 0.9 + 0.3 \cdot 0.8 + 0.13 \cdot 0.9 + 0.04 \cdot 0.2 = 0.84$$

Die Tatsache, daß FRITZ einen geringeren Zugehörigkeitswert in Bezug auf M(VERBINDLICH) hat, spielt bei Anwendung der Hecke *im Grunde genommen* also keine Rolle, da die

Bedeutungskomponente M(VERBINDLICH) eine geringe Gewichtung innerhalb von M(KONZILIENT) hat. Dies führt dazu, daß die Prädikation 'im-Grunde-genommen-konzilient(FRITZ)' einen höheren Wahrheitswert hat als die Prädikation 'konzilient(FRITZ)'.

LAKOFF75 schlägt ein weniger aufwendiges Verfahren zur formalen Darstellung der Hecken vom Typ 2 vor, das von einer Klassifikation der Bedeutungskomponenten eines vagen Ausdrucks ausgeht. Einige Bedeutungskomponenten werden dabei z.B. als primär und andere als sekundär angesehen. Die Wirkung einer Hecke vom Typ 2 besteht darin, daß sie Bedeutungskomponenten einer bestimmten Klasse suspendiert. So werden z.B. durch die Hecke *genaugenommen* Bedeutungskomponenten suspendiert, die als sekundär klassifiziert sind, während die Hecke *grob gesprochen* die primären Bedeutungskomponenten aus der Gesamtbedeutung des als Operand dienenden vagen Ausdrucks herausfiltert.

BEISPIEL: In dem Satz *Grob gesprochen ist ein Wal ein Fisch* wird nur die sekundäre Bedeutungskomponente 'lebt-im-Wasser' berücksichtigt. Die primäre Bedeutungskomponente von Fisch, in der das Wissen codiert ist, daß Fische nicht zur Klasse der Säugetiere gehören, bleibt unberücksichtigt.

Die auf einer Klassifikation von Bedeutungskomponenten beruhende Beschreibung der Hecken vom Typ 2 ziehen wir einer Beschreibung, die Operationen über Gewichten durchführt, vor, da wir soweit wie möglich von symbolischen und diskreten RK Gebrauch machen wollen (vgl. 5.1.).

In Abschnitt 4.6.2.2. wird gezeigt, wie Hecken vom Typ 2 ausgehend von Lakoffs Ansatz in der Programmiersprache KRL, die speziell für die Repräsentation von Wissen entwickelt wurde, dargestellt werden können.

3.2.3. DAS KONZEPT DER LINGUISTISCHEN VARIABLEN

Das Konzept der linguistischen Variablen, d.h. einer Variablen, deren Werte keine Zahlen sondern meist vage Wörter der NS sind, baut auf die Darstellung der Semantik vager Ausdrücke durch F-Mengen und die Beschreibung von Hecken als Operatoren über solchen F-Mengen auf. Der vage Ausdruck *jung* wird dabei z.B. als Wert der linguistischen Variablen *Alter* interpretiert (vgl. z.B. ZADEH75, S. 201ff, S. 313ff).

Die Verwendung von linguistischen Variablen bei der Repräsentation von Wissen führt zu einer Repräsentationsebene, in der numerische Werte nicht mehr zu finden sind (vgl. 3.2.1.). Zunächst soll der Begriff der linguistischen Variablen formal definiert werden.

DEFINITION: Eine linguistische Variable ist ein Quadrupel (N, U, G, M). N sei ein eindeutiger Bezeichner für die linguistische Variable. G sei eine formale Grammatik. Die Wörter der Sprache L(G) sind die möglichen Werte der linguistischen Variablen.

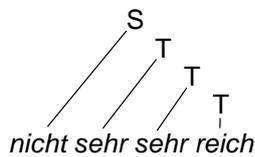
M sei eine Menge von semantischen Interpretationsregeln, die jedem Wort $w \in L(G)$ eine semantische Interpretation in Form einer F-Menge über der Basismenge U zuordnen.

BEISPIEL: Die linguistische Variable (VERMÖGEN, KAPITAL, G, M) hat als Werte verschiedene vage sprachliche Ausdrücke, deren Bedeutung als F-Mengen über einer Basismenge dargestellt werden, die als KAPITAL bezeichnet wird und den positiven, rationalen Zahlen entspricht.

Jede semantische Regel aus M wird als Affix zu einer entsprechenden Produktionsregel der formalen Grammatik G aufgefaßt. Auf diese Weise ergibt sich eine attribuierte Grammatik (vgl. z.B. KNUTH68).

BEISPIEL: Sei $G \in (V_N, V_T, S, P)$, $T \in V_N$, *nicht, sehr, reich* $\in V_T$, $P_1, \dots, P_4 \in P$ $M_1, \dots, M_4 \in M$ L und R kennzeichnen die linke bzw. rechte Seite.

$$\begin{array}{ll}
P_1: S \rightarrow T & M_1: M(S) = M(T) \\
P_2: S \rightarrow \text{nicht} + T & M_2: M(S) = \overline{M(T)} \\
P_3: S \rightarrow \text{sehr} + T & M_3: M(T_L) = M(T_R)^2 \\
P_4: S \rightarrow \text{reich} & M_4: M(T) = SFUNK(x, 10^5, 10^6) = M(\text{reich})
\end{array}$$

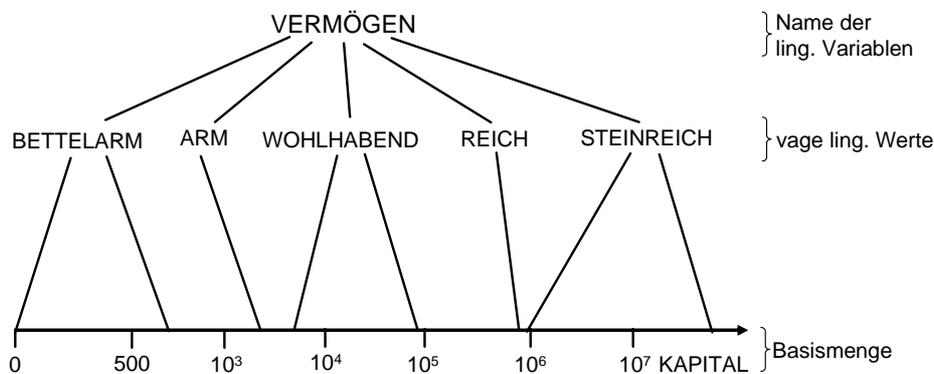


FIGUR 22

Eine Syntaxanalyse des Ausdrucks *nicht sehr sehr reich* mit der Grammatik G ergibt den in Fig. 22 dargestellten Strukturbaum. Eine semantische Interpretation des Strukturbaums ergibt:

$$M(S) = \overline{M(\text{reich})}^4$$

Die beiden unterschiedlichen Ebenen der Repräsentation einer linguistischen Variablen werden in Fig. 23 verdeutlicht.



FIGUR 23

Auf die Notwendigkeit einer kontextabhängigen semantischen Interpretation der Adjektive, die als Werte einer linguistischen Variablen fungieren, wurde bereits in Abschnitt 3.2.1. hingewiesen.

Meist besteht die Basismenge einer linguistischen Variablen nicht aus numerischen Werten, sondern aus einer Menge von Bezeichnern für Wissensseinheiten (vgl. auch 3.2.1. und Fig. 24). So kann z.B. die linguistische Variable *Erscheinungsbild* mit Werten wie *hübsch*, *häßlich* usw. über einer Basismenge definiert werden, die aus Namen von Personen besteht, auf welche die vagen Werte beispielhaft bezogen werden.

Vage Ausdrücke wie *ziemlich wahr*, *völlig falsch*, *sehr wahr*, *nahezu falsch* usw. können als Werte einer linguistischen Variablen *Wahrheit* aufgefaßt werden, für die als Basismenge das reelle Intervall $[0,1]$ definiert ist (vgl. auch ZADEH75, S.332ff). Die Anwendung dieser linguistischen Variablen auf vage Propositionen führt zu einer F-Logik, die eine spezielle Form der mehrwertigen Logik ist (vgl. 4.1.2.).

Wenn wir mit $T(A(x)) \in [0,1] \subset \mathbb{R}$ den Wahrheitswert der Prädikation $A(x)$ bezeichnen, so können wir $T(A(x))$ von einer Prädikation $A(x)$, in der die Zugehörigkeit des Individuums x zur F-Menge $M(A)$ ausgedrückt wird, einfach als den Zugehörigkeitswert von x zur F-Menge $M(A)$ definieren:

$$T(A(x)) \hat{=} \delta_{M(A)}(x)$$

BEISPIEL: $T(\text{mittleres - Alter}(35)) = \delta_{M(\text{mittleres - Alter})}(35) = 0.5$

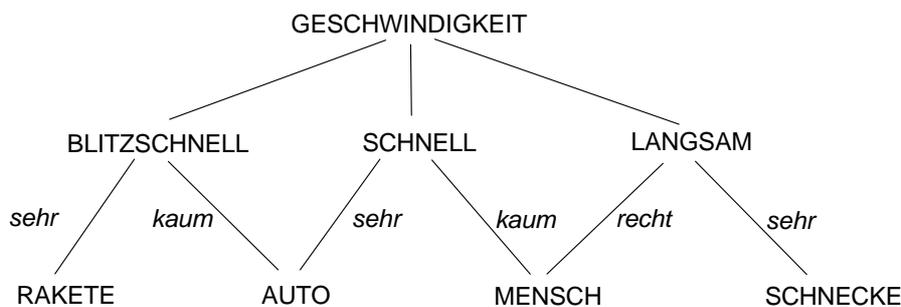
Wird 0.5 nun als Element aus der Basismenge der linguistischen Variablen *Wahrheit* interpretiert und $M_{(mehr-oder-weniger WAHR)} = \{1/0.5, 0.9/0.4, 0.8/0.6\}$ als ein Wert dieser Variablen definiert, so kann die Bedeutung des Wertes *mittleres-Alter*, der zur linguistischen Variablen *Alter* gehört, als F-Menge vom Typ 2 dargestellt werden.

BEISPIEL: $\delta_{M(mittleres-Alter)}(35) = mehr - oder - weniger WAHR$

Zadeh bezeichnet das Verfahren, den Zugehörigkeitswert 0.5 durch den vagen sprachlichen Ausdruck *mehr-oder-weniger WAHR* zu ersetzen, als linguistische Approximation.

Mit der linguistischen Approximation ist ein Verfahren gegeben, das es gestattet ausgehend von den Werten einer linguistischen Variablen *Wahrheit*, die als F-Mengen vom Typ 1 über $[0,1] \subset \mathbb{R}$ definiert sind, die Werte aller anderen linguistischen Variablen als F-Mengen vom Typ 2 zu charakterisieren.

BEISPIEL: Die linguistische Variable *Geschwindigkeit* mit vagen Werten wie *langsam*, *schnell*, *blitzschnell* kann über einer Menge von Beispielen, die als Basismenge fungiert, definiert werden (vgl. dagegen Fig. 23). Obwohl eine quantitative Skala zum Vergleich von Geschwindigkeiten existiert, ist im Rahmen von Alltagswissen z.B. eher die in Fig. 24 dargestellte Struktur anzusetzen. Dabei sind *sehr*, *kaum*, *recht* Abkürzungen für die entsprechenden Werte der linguistischen Variablen *Wahrheit*.



FIGUR 24

Der Nachteil einer solchen Repräsentation besteht darin, daß Operationen auf den RK (besonders: Inferenzen) aufwendiger werden. Als Vorteile einer linguistischen Approximation sind zu nennen:

- An die Stelle von numerischen Werten der Zugehörigkeitsfunktion treten in den RK diskrete, linguistische Werte. Damit wird die heute in der KI bevorzugte Darstellungsweise benutzt (vgl. 5.1).
- Es wird der Forderung entsprochen, daß die Funktionen, durch die der Zugehörigkeitswert eines Elements zu einer F-Menge in Abhängigkeit von Eigenschaften des Elements zu ermitteln ist, vage definiert sind (vgl. 3.2.2.).

Unabhängig davon ist in einem FAS das Verfahren der linguistischen Approximation bei der Generierung von NS-Antworten vorzusehen, da z.B. die Frage *Wie alt ist Fritz?* nicht etwa mit *0.8/jung* sondern beispielsweise mit *recht jung* beantwortet werden sollte.

In Abschnitt 4.4.3.1 wird gezeigt, wie linguistische Variable in erweiterten semantischen Netzen dargestellt werden können.

3.3. PSYCHOLINGUISTISCHE ASPEKTE DER 'FUZZY SETS'

In den vorangegangenen Abschnitten dieses Kapitels haben wir gezeigt, welche Möglichkeiten zur Darstellung von vagen sprachlichen Ausdrücken und von Operatoren über solchen Ausdrücken es

innerhalb der Theorie der F-Mengen gibt. Im folgenden Abschnitt setzen wir uns mit empirischen Untersuchungen im Bereich der kognitiven Psychologie und der Psycholinguistik auseinander, durch die überprüft wird, inwieweit eine auf der Theorie der F-Mengen basierende Darstellung von sprachlichen Ausdrücken und Hecken als adäquates, psychologisches Modell der Repräsentation und Verarbeitung von vagem Wissen betrachtet werden kann (vgl. auch 2.2.).

Diese Fragestellung wird in einigen neueren psychologischen Untersuchungen¹⁹ behandelt, die folgende Forschungsziele haben:

Our aim is to establish the psychological reality of fuzzy sets.

(KOCHEN/BADRE74, S.53)

...the theoretical position that runs through this line of research is that people comprehend vague concepts (i.e. all natural language concepts) as if the concepts are represented as fuzzy sets. Moreover, people manipulate vague concepts as if they are processing according to the rules of fuzzy logic.

(HERSH/CARAMAZZA76, S.263)

In den angesprochenen Untersuchungen²⁰ wurde zunächst versucht, experimentell den Verlauf der Zugehörigkeitsfunktionen für einige F-Mengen zu bestimmen, welche die Bedeutung vager sprachliche Ausdrücke beschreiben. Solche F-Mengen wurden ermittelt z.B. für die Adjektive klein und groß bezogen auf die Größe von Quadraten (vgl. HERSH/CARAMAZZA76), groß bezogen auf die Körpergröße einer Person (vgl. MACVICAR WHELAN74), kalt bezogen auf Temperaturangaben, weit bezogen auf räumliche Entfernungen (vgl. DREYFUSSetal.75) und viel größer als 5 bezogen auf Zahlenwerte (vgl. KOCHEN/BADRE74).

Ausgehend von diesen F-Mengen wurde der Verlauf der Zugehörigkeitsfunktionen für Ausdrücke wie nicht groß, sehr groß usw. in weiteren Experimenten mit Testpersonen bestimmt. Die Resultate dieser Tests wurden mit den Werten verglichen, die sich durch Anwendung der formalen Definitionen von Operationen wie die Komplementbildung für nicht oder die Konzentration für sehr auf die ursprünglichen F-Mengen ergaben. Dabei zeigte es sich, daß die formale Beschreibung des Junktors nicht und der Hecke eine Art durch die empirischen Ergebnisse bestätigt werden kann (vgl. HERSH/CARAMAZZA76, S.263ff).

Zadehs einfache Darstellung der Hecke sehr (vgl. 3.2.2.1. und SEHR₁, X in Fig. 25) konnte allerdings experimentell nicht belegt werden. Hersh und Caramazza stellten fest, daß sehr eine Translation der Zugehörigkeitsfunktion entlang der Abszisse bewirkt (vgl. HERSH/CARAMAZZA76, S. 265). Die Ergebnisse von KOCHEN/BADRE74 sprechen für eine Beschreibung von sehr als Kombination der elementaren Operationen Konzentration und Translation. Diese empirischen Untersuchungen sprechen für die von Lakoff vorgeschlagene Beschreibung von sehr x als:

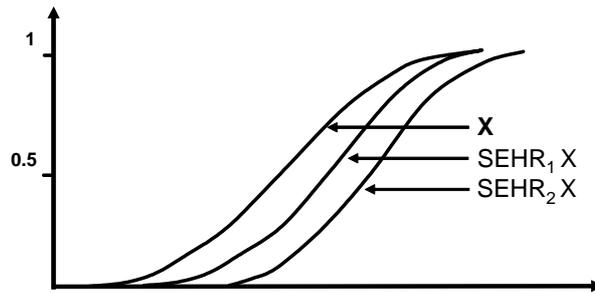
$$\delta_{M(\text{sehr } x)} = \text{CON}(\delta_{M(x)} - c)$$

wobei c eine von der Referenzmenge abhängige Konstante ist (vgl. LAKOFF 75, S.251ff und SEHR₂, X in Fig. 25).

Nach Zadehs Analyse von sehr gilt $\delta_{M(\text{GROSS})}(x) = 1 \Rightarrow \delta_{M(\text{sehr GROSS})}(x) = 1$, was kontraintuitiv zu sein scheint.

¹⁹ vgl. z.B. MACVICAR WHELAN74/76, KOCHEN/BADRE74, KOCHEN75, DREYFUSS etal.75, HERSH/CARAMAZZA76

²⁰ Diese Untersuchungen unterscheiden sich deutlich von den bekannten Arbeiten zum semantischen Differential (vgl. z.B. OSGOODetal.57), da Osgood die konnotative Bedeutung von Wörtern, d.h. emotionale Begleitvorstellungen und andere Nebeninhalte untersuchte, während sich die hier betrachteten Arbeiten ausschließlich auf die referentielle Bedeutung beziehen.



FIGUR 25

Außerdem ergab sich aus den psychologischen Tests, daß die Wirkung einer Hecke sich bei mehrfachem Auftreten dieser Hecke innerhalb eines Ausdrucks ändern kann²¹ und daß bei einigen Ausdrücken, die idiomatischen Charakter haben²², Zadehs einfaches Verfahren, das die Wirkung einer zusammengesetzten Hecke als lineare Kombination der Wirkungen ihrer Konstituenten beschreibt, zu Ergebnissen führt, die mit den empirischen Daten nicht übereinstimmen.

Zusammenfassend stellen wir fest:

- Einige psychologische Untersuchungen sprechen dafür, vage Konzepte durch F-Mengen darzustellen. Eine solche Repräsentation ist nicht nur formal befriedigend sondern auch als psychologisches Modell brauchbar.
- Schon die formale Beschreibung einer in Zadehs System so einfachen Hecke wie *sehr* erfordert bei einer genaueren Analyse einen erheblichen Aufwand. Obwohl für komplexe Hecken noch keine empirischen Aussagen vorliegen, ist zu erwarten, daß diese oft durch komplizierte, kontextabhängige Operationen über Zugehörigkeitsfunktionen repräsentiert werden müssen. Alle genannten Psychologen sehen allerdings den allgemeinen Ansatz, Hecken als Operationen über F-Mengen zu repräsentieren, durch ihre Untersuchungen bestätigt. Auch die Beschreibung der Hecke *sehr* als Präzisionsoperator, der aus einer F-Menge eine F-Teilmenge auswählt, wird durch empirische Daten gestützt.

²¹ z.B. bei Konstruktionen wie *sehr*, *sehr klein* nimmt die Wirkung der Hecke *sehr* ab.

²² z.B. wird *nicht sehr groß* als *ziemlich klein* und nicht als Komplement von *sehr groß* interpretiert.

4. ZUR FORMALEN DARSTELLUNG VON VAGEM WISSEN IN VERSCHIEDENEN REPRÄSENTATIONSSPRACHEN DER KI

4.1. DIE DARSTELLUNG VON VAGEM WISSEN IN DEN AN DER PRÄDIKATENLOGIK ORIENTIERTEN REPRÄSENTATIONSSPRACHEN

4.1.1. EINFÜHRENDE ÜBERSICHT

Der Prädikatenkalkül (Abk.: PC), der sich in der Mathematik als Beschreibungssprache bewährt hat, wird auch in der KI-Forschung häufig als Repräsentationssprache vorgeschlagen und benutzt. Dies hängt u.a. damit zusammen, daß die Syntax und Semantik des PC seit langem bekannt und die Eigenschaften des PC bereits intensiv untersucht sind.

Für NSS hat die Repräsentationssprache PC zwei besondere Vorteile:

- In der Linguistik wurden semantische Modelle entwickelt, in denen die Bedeutung sprachlicher Ausdrücke durch Formeln des PC repräsentiert wird. Die Ergebnisse dieser Untersuchungen können beim Aufbau einer Wissensbasis eines NSS direkt verwendet werden.
- Das Problem, eine Frage durch Inferenzoperationen über einer Wissensbasis zu beantworten, kann auf die allgemeinere Fragestellung zurückgeführt werden, ob und wie eine bestimmte Formel des PC aus den in der Wissensbasis enthaltenen Axiomen abgeleitet werden kann. Dabei wird sowohl die zu beantwortende Frage als auch das gesamte Wissen in Form von prädikatenlogischen Formeln repräsentiert (vgl. z.B. GREEN/RAPHAEL68, CHANG/LEE73). Auf diese Weise lassen sich die zahlreichen seit der Entwicklung des Resolutionsprinzips im Jahre 1965 (vgl. ROBINSON65) gewonnenen Ergebnisse des Forschungsschwerpunktes 'Automatisches Beweisen' (vgl. 1.1.1.) auf die Konstruktion von NSS, insbesondere von FAS, anwenden.

Als weitere Vorteile dieser deklarativen Repräsentationssprache (vgl. 1.1.3.) gelten die große Allgemeinheit und die Möglichkeit einer modularen Darstellung von Wissen (vgl. z.B. WINOGRAD74, S.51ff). Als wesentlicher Nachteil für die Realisation von NSS, die auf dem PC aufbauen und Deduktionen durch einen automatischen Theorembeweiser ausführen, wird oft die Ineffizienz von Inferenzprozessen im PC und das damit verbundene Problem der kombinatorischen Explosion genannt (vgl. CHARNIAK76). Diese Beurteilung muß allerdings relativiert werden in bezug auf neuere Arbeiten aus dem Forschungsgebiet 'Automatisches Beweisen', in denen z.B. das Deduktionsverfahren selbst in starkem Maße durch Wissen gesteuert wird (vgl. z.B. BLEDSOE75) oder besondere Verfahren wie die Verwendung von Verbindungsgraphen (vgl. KOWALSKI75) die Effektivität des Gesamtverfahrens erheblich steigern.

Da bereits viele Wissensstrukturen gefunden wurden, die innerhalb des PC erster Stufe nur durch sehr aufwendige und unnatürliche RK dargestellt werden können, wurden in der KI verschiedene Erweiterungen des PC erster Stufe vorgeschlagen, die eine einfachere Abbildung der NS in Formeln eines PC ermöglichen sollen.²³

Diese Bestrebungen in der KI stehen im Einklang mit linguistischen Arbeiten, in denen auf die Beschränkungen des PC erster Stufe bei der Beschreibung der Semantik der NS aufmerksam gemacht wird. In solchen Arbeiten werden meist Formen der Modallogik (vgl. HUGHES/CRESSWELL68) und vereinzelt auch mehrwertige Logiken vorgeschlagen (vgl. KEENAN72, BLAU73, LAKOFF75).

Im Rahmen dieser Arbeit werden nur diejenigen Erweiterungen des PC berücksichtigt, die eine Repräsentation von vagem Wissen ermöglichen sollen.

Während im PC erster Stufe ein Syllogismus wie

BEISPIEL: Fritz ist reich.
Alle reichen Leute bezahlen viele Steuern.
Fritz zahlt viele Steuern.

²³ vgl. McCARTHY/HAYES69, HAYES69/71/74, BRUCE71, ZIFONUN75, SCHWIND76

formalisiert werden kann, ohne die Vagheit der sprachlichen Ausdrücke *reich* und *viele* zu berücksichtigen, versagt die direkte Repräsentation im PC erster Stufe bereits bei so alltäglichen Inferenzen wie

BEISPIEL: Fritz ist sehr reich.
Alle reichen Leute bezahlen viele Steuern.
Fritz bezahlt sehr viele Steuern.

Im folgenden werden u.a. RK diskutiert, die solche Inferenzen über vagem Wissen ermöglichen.

Dabei werden beim Leser Grundkenntnisse der Terminologie und der Verfahren im Bereich 'Automatisches Beweisen' vorausgesetzt.²⁴ Wir gehen zunächst auf Fuzzy Logiken (kurz: F-Logiken) ein, die eine spezielle semantische Interpretation von mehrwertigen Logiken sind, und zeigen dann, wie die intensionale Bedeutung relativer Adjektive und Vergleiche durch logische Formalismen dargestellt werden können.

Bereits im Jahr 1969 wurde in der KI vorgeschlagen, eine mehrwertige Logik bei der Darstellung von Weltmodellen (vgl. 1.1.3.) einzusetzen, wie sie Roboter (z.B. Shakey, Freddy) bei der Planung ihrer Aktionen benutzen (vgl. GREEN69, S.229). Die interne Repräsentation der äußeren Umgebung eines Roboters ist zwangsläufig vage und unvollständig. Wird eine aufgrund eines solchen internen Modells geplante Handlung durch den Roboter ausgeführt und das reale mit dem vorausberechneten Ergebnis der Aktion verglichen, so ergeben sich aus folgenden Gründen meist Abweichungen, welche durch Sensoren des Roboters ermittelt und in mehreren Rückkopplungsschleifen ausgeglichen werden:

- die reale Aktion des Roboters entspricht nicht genau der geplanten Handlung.
- das Weltmodell enthält zu starke Vergrößerungen (z.B. wurden beim Verschieben einer Kiste die Gewichtsverteilung der Kiste und Unebenheiten auf der Fläche, welche die Kiste trägt, nicht berücksichtigt).
- die Meßergebnisse der Sensoren sind ungenau.

Hier bietet sich eine mehrwertige Logik zur Darstellung des internen Weltmodells an, wenn z.B. das Ziel einer Aktion, nämlich die Gültigkeit einer Prädikation $p(a)$ zu bewirken, nur zu einem gewissen Grad z.B. $p(a) = 0.8$ erreicht wird.

Dieser Grundgedanke durchzieht eine Reihe von Veröffentlichungen (z. B. MUNSON71, SIMON73) und wurde von LEE72 als Motivation dafür genannt, eine F-Logik als Basis für einen automatischen Theorembeweiser zu wählen.

Ein Theorembeweiser, der auf der Basis einer F-Logik arbeitet, wurde bisher noch nicht implementiert²⁵.

4.1.2. MEHRWERTIGE LOGIK UND 'FUZZY LOGIC'

Mehrwertige Logiken²⁶ sind gegenüber den üblichen zweiwertigen Logiken dadurch erweitert, daß in ihnen mit mehr als zwei Wahrheitswerten, die Elemente aus einer geordneten Menge Π sind, gearbeitet wird. Wir beschäftigen uns hier nur mit unendlichen Logiken, d.h. mit Logiken, in denen Π eine unendliche Menge ist. Obwohl GOGUEN69/74 zeigte, daß Π im Rahmen von F-Logiken allgemein als ein beliebiger boolscher Verband charakterisiert werden kann, wird im folgenden lediglich der Standardfall $\Pi = [0,1] \subset \mathbb{R}$ betrachtet. ZADEH75 definiert die F-Logik als eine auf der Theorie der F-Mengen beruhende semantische Interpretation des von Lukasiewicz bereits 1930 vorgeschlagenen Kalküls L_{Aleph_i} (vgl. BERKOWSKI70, RESCHER69, S. 336).

Die Evaluation $T(S) \in [0,1]$ einer wohlgeformten Formel, die nach derselben Syntax wie Formeln der

²⁴ Einführungen findet man z.B. in NILSSON71, CHANG/LEE73.

²⁵ vgl. auch den Überblick zu Theorembeweisern für nicht-klassische Logiken in MORGAN76

²⁶ vgl. SINOWJEW68, RESCHER69, EPSTEINetal.74, WOLF75

zweiwertigen Logik (vgl. z.B. CHANG/LEE73) aufgebaut ist, wird in L_{Aleph_1} , folgendermaßen definiert:

- (1) $T(\neg P) \triangleq 1 - T(P)$
- (2) $T(P \wedge Q) \triangleq \text{MIN}(T(P), T(Q))$
- (3) $T(P \vee Q) \triangleq \text{MAX}(T(P), T(Q))$
- (4) $T(P \Rightarrow Q) \triangleq \text{MIN}(1, 1 - T(P) + T(Q))$

Die einzelnen F-Logiken unterscheiden sich von L_{Aleph_1} lediglich in der Definition der Implikation (vgl. auch GAINES76, S. 640ff).

LEE72 schlägt für eine F-Logik eine Form der Implikation vor, wie sie schon von DIENES49 in einem mehrwertigen Kalkül benutzt wurde:

$$(4') \quad T(P \Rightarrow Q) \triangleq \text{MAX}(1 - T(P), T(Q))$$

In GOGUEN89 ist schließlich eine dritte Variante einer F-Logik zu finden, in der die Implikation folgendermaßen definiert ist:

$$(4'') \quad T(P \Rightarrow Q) \triangleq \text{MIN}\left(1, \frac{T(Q)}{T(P)}\right)$$

Quantoren führen wir durch folgende Definitionen (vgl. auch LEE72, GAINES 76) ein, wobei SUP und INF als Supremum bzw. Infimum zu verstehen sind:

- (5) $T(\forall x: p(x)) \triangleq \text{INF}(T(p(x)) | x \in D)$
- (6) $T(\exists x: p(x)) \triangleq \text{SUP}(T(p(x)) | x \in D)$

Gaines legte die erste vollständige Axiomatisierung einer F-Logik vor, die Quantoren enthält und auch Hecken vom Typ 1 axiomatisch als unäre Operatoren einführt (vgl. GAINES76, S.643ff). Dadurch wurde gezeigt, daß eine F-Logik ohne Rückgriff auf Zahlenwerte, d.h. außerhalb eines arithmetischen Kontextes, allein aufgrund von metamathematischen Deduktionsregeln eingeführt werden kann. Die Tatsache, daß die logischen Junktoren auch als Operatoren über Zahlenwerten charakterisiert werden können, ist nur von sekundärer Bedeutung und hat höchstens den praktischen Vorteil, daß eine aufwendige Verarbeitung von Zeichenketten z.T. durch einfache numerische Operationen ersetzt werden kann.

4.1.3. DIE REPRÄSENTATION VAGER INFERENZREGELN UND IHRE VERARBEITUNG DURCH 'FUZZY RESOLUTION'

Wir beschränken uns im folgenden auf die Betrachtung des modus ponens als einer der bei einer Axiomatisierung der F-Logik benutzten metamathematischen Deduktionsregeln, da der modus ponens die Grundlage für die von automatischen Theorembeweisern benutzte Resolutionsregel bildet. Dabei gehen wir von folgendem praktischen Repräsentationsproblem aus:

BEISPIEL: Folgende in NS gegebenen Fakten sollen in einer Wissensbasis dargestellt werden:

- (1) Delmenhorst liegt recht nahe bei Bremen.
- (2) Bremen liegt ziemlich nahe bei Hamburg.

Außerdem soll auch eine RK für folgende vage Inferenzregel angegeben werden:

- (3) Wenn x nahe bei y liegt und y nahe bei z liegt, dann liegt x recht nahe bei z.

Auf die Frage *Liegt Delmenhorst nahe bei Hamburg?* hin soll ein NSS, dem RK für (1)-(3) in der Wissensbasis zur Verfügung stehen, eine Inferenzoperation durchführen und natürlichsprachlich antworten.

Zunächst stellen wir fest, daß in (1)-(3) das vage Prädikat nahe und die Hecken *recht* und *ziemlich* enthalten sind. Die Transitivität der Relation nahe gilt, wie WINOGRAD75 feststellt, nur beschränkt, d.h. die mehrfache Anwendung innerhalb einer Deduktionskette führt zu kontraintuitiven Ergebnissen (z.B. Hamburg liegt nahe bei München, da Hamburg nahe bei Hannover, Hannover nahe bei Kasselliegt).

Sei nahe(x, y) ein zweistelliges F-Prädikat, S die Trägermenge der F-Menge aller Städte und DEL, HB, HH Individuenkonstanten, welche die Städte Delmenhorst, Bremen und Hamburg repräsentieren.

Um eine möglichst einfache Darstellung der Hecken vom Typ 1 zu erreichen, wird angenommen, daß jede Hecke in der Objektsprache eine Hecke in der Metasprache induziert und umgekehrt. Nach diesem Prinzip gilt z.B.

$$Fritz\ ist\ sehr\ reich \Leftrightarrow Fritz\ ist\ reich\ ist\ sehr\ wahr.$$

Außerdem nehmen wir an, daß *recht wahr* und *ziemlich wahr* zwei Werte der linguistischen Variablen *Wahrheit* (vgl. 3.2.3.) mit der Basismenge $[0,1] \subset \mathbb{R}$ sind. $M_{(ziemlich\ WAHR)}$ und $M_{(recht\ WAHR)}$ werden stark vereinfacht als die singulären F-Mengen vom Typ 1 $\{0.7/1\}$ bzw. $\{0.9/1\}$ dargestellt. Wenn wir die Hecken in (1)-(3) von der Objekt- in die Metasprache transformieren, so können wir (1)-(3) durch folgende RK in der Wissensbasis darstellen:

$$(1') \quad T(nahe(DEL, HB)) = 0.9$$

$$(2') \quad T(nahe(HB, HH)) = 0.7$$

$$(3') \quad T(\forall x : \forall y : \forall z : \underset{x \in S}{nahe(x, y)} \wedge \underset{y \in S}{nahe(y, z)} \Rightarrow \underset{z \in S}{nahe(x, z)} = 0.9$$

Um eine Beantwortung der Frage zu ermöglichen, muß nun der modus ponens auf diese RK angewendet werden. Der modus ponens in seiner allgemeinen Form besagt:

$$(P \quad \langle \text{Abtrennungsoperator} \rangle \quad (P \Rightarrow Q)) \Rightarrow Q$$

In der zweiwertigen Logik wird die Konjunktion als Abtrennungsoperator benutzt. In der F-Logik gibt es mehrere Möglichkeiten, einen Abtrennungsoperator zu definieren, wobei als Randbedingung stets die Ungleichung

$$T(P) \quad \langle \text{Abtrennungsoperator} \rangle \quad T(P \Rightarrow Q) \leq T(Q)$$

gilt, die sicherstellt, daß der Wahrheitswert der Konklusion nicht T(Q) übersteigt.

Wir diskutieren im folgenden zwei unterschiedliche Abtrennungsoperatoren, die auf die Definition der Implikation in (4') und (4'') aufbauen. Ergebnisse von Lee weisen darauf hin, daß das Minimum ein geeigneter Abtrennungsoperator ist (vgl. LEE72, Lemma4 und Theorem9). GOGUEN69 schlägt dagegen die Multiplikation als Abtrennungsoperator für eine F-Logik vor. Die Eigenschaften der beiden Abtrennungsoperatoren sind in Fig. 26 dargestellt.

In einem Resolutionsverfahren kann bei Anwendung der Resolutionsregel der Wahrheitswert einer Resolvente $R(C_1, C_2)$ von zwei Klauseln C_1 und C_2 abhängig davon, für welchen Abtrennungsoperator man sich bei einer gegebenen Problemstellung entscheidet, als $\text{MIN}(T(C_1), T(C_2))$ oder als $T(C_1) \cdot T(C_2)$ berechnet werden.

MINIMUM ALS ABTRENNUNGSOPERATOR (LEE)	MULTIPLIKATIVER ABTRENNUNGSOPERATOR (GOGUEN)
$T(B) = \min(T(A), T(A \Rightarrow B))$	$T(B) = T(A) \cdot T(A \Rightarrow B)$
Der Wahrheitswert der Konklusion hängt vom schwächsten Glied der Kette ab.	Der Wahrheitswert der Konklusion nimmt mit der Länge der Deduktionskette ab.

FIGUR 26

Um ein F-Resolutionsverfahren (engl. fuzzy resolution) bei unserem Beispiel einsetzen zu können, formen wir zunächst (1')-(3') in Klauselform um und ergänzen die so erhaltenen Klauseln um die negierte Form derjenigen Formel, welche die Frage repräsentiert. Auf diese Klauselmengen wenden wir die Resolutionsregel und einen Unifikationsalgorithmus solange an, bis wir die leere Klausel \square erhalten (vgl. Fig. 27).

$C_1 : \text{nahe}(\text{DEL}, \text{HB})$	$T(C_1) = 0.9$
$C_2 : \text{nahe}(\text{HB}, \text{HH})$	$T(C_2) = 0.7$
$C_3 : \neg \text{nahe}(x, y) \vee \neg \text{nahe}(y, z) \vee \text{nahe}(x, z)$	$T(C_3) = 0.9$
$C_4 : \neg \text{nahe}(\text{DEL}, \text{HH})$	$T(C_4) = 1$

	UNIFIKATOR	ABTRENNUNGSOPERATOR	
		MINIMUM	MULTIPLIKATION
$C_5 = R(C_4, C_3) : \neg \text{nahe}(\text{DEL}, y) \vee \neg \text{nahe}(y, \text{HH})$	{DEL/x, HH/z}	$T(C_5) = 0.9$	$T(C_5) = 0.9$
$C_6 = R(C_5, C_2) : \neg \text{nahe}(\text{DEL}, \text{HB})$	{HB/y}	$T(C_6) = 0.7$	$T(C_6) = 0.63$
$C_7 = R(C_6, C_1) : \square$		$T(C_7) = 0.7$	$T(C_7) = 0.567$

FIGUR 27

Auf das Ergebnis der mechanischen, Beweisprozedur $T(\square) = 0.7$ bzw. $T(\square) = 0.567$ wird zunächst das Verfahren der linguistischen Approximation (vgl. 3.2.3.) angewendet. Nach einer sich anschließenden Transformation der Hecken von der Metasprache in die Objektsprache erhält man z.B. folgende natürlichsprachliche Formulierung als Antwort: *Delmenhorst liegt ziemlich nahe bei Hamburg*

Im vorliegenden Beispiel entspricht der multiplikative Abtrennungsoperator der 'beschränkten Transitivität' von *nahe*. Falls auch die Anfrage eine Hecke enthält, so muß auf das Ergebnis des Resolutionsverfahrens der entsprechende Heckenoperator angewandt werden.

BEISPIEL: An ein NSS, das in der Wissensbasis die im vorangegangenen Beispiel benutzten RK enthält, wird die Frage *Liegt Delmenhorst sehr, sehr nahe bei Hamburg* gestellt. Wie oben sei das Ergebnis der Resolution $T(\square) = 0.7$. Verwendet man die einfache Zadehsche Definition der Hecke *sehr*, so erhält man 0.24 als Wahrheitswert der Konklusion. Eine linguistische Approximation ergibt dann z.B. die Antwort *Nein, Delmenhorst liegt ziemlich nahe bei Hamburg*.

Auch W-Fragen lassen sich durch das skizzierte Verfahren bearbeiten, wenn die Frage durch eine RK ersetzt wird, die das Greensche Antwortprädikat (vgl. CHANG/LEE73, Kap. 11) enthält.

Das hier angedeutete, stark vereinfachte Verfahren, bei dem die unterschiedlichen Wahrheitswerte der einzelnen Klauseln nicht in die Resolutionsstrategie eingehen, ist lediglich eine von vielen Möglichkeiten zur Entwicklung eines Theorembeweisens auf der Basis einer F-Logik.

Die oben definierte F-Logik wird der Fülle von linguistischen und psychologischen Aspekten der Vagheit von NS-Ausdrücken nur in sehr begrenztem Maße gerecht (z.B. wurden Hecken vom Typ 2 nicht berücksichtigt).

Die wenigen Ansätze zum Ausbau der F-Logik zu einer F-Modallogik (vgl. SCHOTSCH75, LAKOFF75) oder F-Präsuppositionslogik (vgl. LAKOFF75) sind, obwohl sie eine weitergehende

Darstellung der Semantik vager Ausdrücke ermöglichen, von einer Mechanisierung noch weit entfernt und daher für die formale Repräsentation von vagem Wissen in NSS ungeeignet.

REITER75 beurteilt generell die Möglichkeiten, komplizierte vage Inferenzen durch die an der Prädikatenlogik orientierten Formalismen zu repräsentieren, skeptisch.

4.1.4. DIE PRÄDIKATENLOGISCHE DARSTELLUNG DER INTENSIONALEN BEDEUTUNG RELATIVER ADJEKTIVE UND VERGLEICHE

Nachdem bisher lediglich die Repräsentation von vagen Inferenzregeln und Hecken, in den am Prädikatenkalkül orientierten Repräsentationssprachen erörtert wurde, soll im folgenden geprüft werden, inwieweit eine solche Repräsentationssprache zur Darstellung der intensionalen Bedeutung relativer Adjektive und Vergleiche geeignet ist (vgl. auch 2.3.2.).

Es wurde bereits gezeigt, auf welche Weise die extensionale Bedeutung relativer Adjektive wie *jung*, *arm*, *groß* usw. im Rahmen der Theorie der F-Mengen formal repräsentiert werden kann (vgl. 3.2.1). Ein NSS, dessen Wissensbasis entsprechende F-Mengen-theoretische RK enthält, kann eine Frage wie *Ist Fritz ein großer Jockey?* dadurch beantworten, daß es, falls es Wissen über die Körpergröße von Fritz gespeichert hat, den Wert von $\delta_{M(\text{GROSS:JOCKEY})}(\text{FRITZ})$ berechnet und nach einer linguistischen Approximation z.B. das Ergebnis *Ja, er ist ein sehr großer Jockey* ausgibt.

Um aber eine Frage wie *Was bedeutet 'Fritz ist ein großer Jockey'?* z.B. mit *Die Körpergröße von Fritz übersteigt die Körpergröße eines typischen Jockeys* beantworten oder den Widerspruch in der Eingabe *Fritz ist ein kleiner Basketballspieler und größer als ein typischer Basketballspieler* erkennen zu können, muß das NSS auch auf Wissen über die intensionale Bedeutung (vgl. 2.3.3) von *klein* und *groß* zurückgreifen können.

Woods hat mit Nachdruck auf die Notwendigkeit zur Repräsentation der intensionalen Bedeutung in NSS hingewiesen²⁷ (vgl. WOODS75, S.49).

Eine Darstellung des Satzes (1) *Fritz ist ein großer Jockey* durch die prädikatenlogische RK $\text{jockey}(\text{Fritz}) \wedge \text{groß}(\text{Fritz})$ ist nicht adäquat, da das bei relativen Adjektiven im Positiv auftretende Vergleichsmoment darin nicht berücksichtigt wird. Erweitert man den Satz (1) zu (2) *Fritz ist ein großer Jockey und kein großer Basketballspieler* so ergibt sich durch die inadäquate Formalisierung folgender 'künstliche Widerspruch':

$\text{jockey}(\text{Fritz}) \wedge \neg \text{basketballspieler}(\text{Fritz}) \wedge \text{groß}(\text{Fritz}) \wedge \neg \text{groß}(\text{Fritz})$.

Mit dem in BARTSCH72, BARTSCH/VENNEMANN735, VENNEMANN73 entwickelten Formalismus kann (1) folgendermaßen dargestellt werden:

BEISPIEL: $\text{jockey}(\text{Fritz}) \wedge (f^M(\text{Fritz}, \text{Körpergröße}) > N_{\text{Körpergröße, Jockey}})$

Dabei wird f^M als allgemeine Maßfunktion (vgl. auch ROSS70, WUNDERLICH73) interpretiert, die auf verschiedene Skalen (hier: Körpergröße, aber auch Schönheit, Faulheit usw.) anwendbar ist. $N_{\text{Körpergröße, Jockey}}$ ist eine Konstante, welche die typische Körpergröße von Jockeys darstellt.

DAMERAU75 schlägt eine vergleichbare RK in LISP-ähnlicher Notation vor:

BEISPIEL: (AND (JOCKEY(FRITZ) GREATER (KÖRPERGRÖSSE(FRITZ)) (NORM(KÖRPERGRÖSSE(JOCKEY))))))

Als Verfeinerungen dieser beiden Vorschläge sind die von EISENBERG76 und SCHWIND76 entwickelten Formalismen zu betrachten, in denen für jedes Adjektiv eines Antonympaares wie *groß* - *klein* ein spezieller Vergleichswert auf einer entsprechenden Skala vorgesehen ist, um negative Inferenzen wie $\neg \text{groß}(x) \Rightarrow \text{klein}(x)$ zu verhindern (vgl. auch 2.3.2).

EISENBERG76 erweitert Bartsch/Vennemanns Notation um zwei Funktionssymbole f_{V1} und f_{V2} , die typische Vergleichswerte für positiv bzw. negativ polarisierte Adjektive liefern (vgl. EISENBERG76, S.117ff):

²⁷ vgl. auch die Kritik von SCHWIND76, S.202 an Modellen, die wie Winograds SHRDLU-System nur auf einer extensionalen Semantik basieren.

BEISPIEL: $\text{jockey}(\text{Fritz}) \wedge \text{Üb}(f^M(\text{Fritz}, \text{Körpergröße}), f_{V_1}(\text{Jockey}))^{28}$; Ü**Ü** steht bei Eisenberg als Abkürzung für das Prädikatensymbol 'Überschreitet'.

Schwind führt eine Konstante $\text{VG}_{\text{groß}, \text{Jockey}}$ in ihren erweiterten Prädikatenkalkül ein, die wie Eisenbergs $f_{V_1}(\text{Fritz})$ die Körpergröße angibt, die ein großer Jockey überschreitet (vgl. SCHWIND76, S.199):

BEISPIEL: $\text{jockey}(\text{Fritz}) \wedge (\varphi_{\text{groß}}(\text{Fritz}) \geq_{\{\text{cm}\}} \text{VG}_{\text{groß}, \text{Jockey}})$
 $\varphi_{\text{groß}}(\text{Fritz})$ steht wie $f^M(\text{Fritz}, \text{Körpergröße})$ für die Körpergröße von Fritz.

Nach dieser detaillierten Betrachtung des Positivs ergeben sich folgende RK für die expliziten Vergleichsformen (vgl. auch SANDEWALL71):

BEISPIELE:

Original: *Fritz ist so groß wie Maria.*

RK: $\varphi_{\text{groß}}(\text{Fritz}) \approx_{\{\text{cm}\}} \varphi_{\text{groß}}(\text{Maria})^{29}$

Original: *Fritz ist größer als Maria.*

RK: $\varphi_{\text{groß}}(\text{Fritz}) >_{\{\text{cm}\}} \varphi_{\text{groß}}(\text{Maria})$

Original: *Fritz ist der größte von Dieters Söhnen.*

RK: $\forall x (\text{sohn}(x, \text{Dieter}) \wedge x \neq \text{Fritz}) \Rightarrow \varphi_{\text{groß}}(\text{Fritz}) >_{\{\text{cm}\}} \varphi_{\text{groß}}(x)$

Um das in prädikatenlogischen Formeln dargestellte Wissen in der Datenbasis eines NSS abspeichern zu können, müssen die oben angegebenen RK noch in eine lineare Schreibweise transformiert werden.

Zusammenfassend stellen wir fest, daß für die Repräsentation der intensionalen Bedeutung vager Adjektive Formeln der Prädikatenlogik geeignet sind, während für die Darstellung der extensionalen Bedeutung F-Mengen-theoretische RK vorzusehen sind.

Dieses Ergebnis deckt sich mit der Feststellung von v. Held, daß einer extensionalen Vagheit von Begriffen keineswegs eine intensionale entsprechen oder vorausgehen muß (vgl. v.HELD76 S. 174).

4.2. DIE REPRÄSENTATION VON VAGEM WISSEN IN PRODUKTIONEN-SYSTEMEN

4.2.1. EINFÜHRENDE ÜBERSICHT

Produktionensysteme (Abk.: PS), die in die KI von Newell und Simon eingeführt wurden (vgl. z.B. NEWELL/SIMON72, NEWELL73), gehören zur Klasse der deklarativen Repräsentationssprachen (vgl. 1.1.3.).

Ein klassisches Produktionensystem³⁰ besteht aus 3 Komponenten:

- einer geordneten Menge von Produktionen (syn.: Regeln)
- einer Datenbasis³¹
- einem Interpreter

²⁸ Nach EISENBERG76 dagegen: $f_{V_1}(\text{Fritz})$

²⁹ Das Prädikatensymbol \approx wird als F-Prädikat 'ungefähr gleich' interpretiert.

³⁰ vgl. die Übersichten in DAVIS/KING75, WATERMAN77, oder auch: WINOGRAD74, S. 55ff, BOLEY76a, S.75ff

³¹ speziellere Verwendung des Begriffs als in Informationssystemen

Jede Produktion hat die Form $\langle \text{Bedingung} \rangle \rightarrow \langle \text{Aktion} \rangle$. Sowohl die linke als auch die rechte Seite (Abk.: LS bzw. RS) einer Produktion kann Variablen enthalten (vgl. POST43). Wenn man LS als einen ggf. komplexen bedingten Ausdruck auffaßt, der sich auf den Inhalt der Datenbasis bezieht und der, falls er vom Interpreter zu wahr ausgewertet wird, die in RS codierten Aktionen auf der Datenbasis auslöst, so entspricht jede Produktion einer Inferenzregel. Da das zu repräsentierende Wissen oft aus vagen Inferenzregeln besteht (vgl. 4.1.3.), sind in der RS u.a. Möglichkeiten zur Darstellung der unterschiedlichen Aussagekraft der einzelnen Inferenzregeln vorzusehen.³²

In der Datenbasis (Abk.: DB) eines PS ist die Menge der in einem bestimmten Anwendungskontext gültigen Aussagen abgespeichert. Dazu gehören auch alle Angaben über den Zustand des Problemlösungsprozesses, der sich aus einer Anwendung der Produktionen auf die in der DB gespeicherten Fakten ergibt. Die RK, aus denen die DB besteht, sind allerdings in den einzelnen KI-Systemen syntaktisch sehr unterschiedlich aufgebaut (z.B. beliebige Symbolketten bei einer Anwendung in der kognitiven Psychologie, Quadrupel im MYCIN-System, Linearisierte komplexe Graphen im DENDRAL-System). Wenn das in der DB kurzfristig gespeicherte Wissen u.a. auch aus vagen Angaben und Mutmaßungen besteht, dann sind RK vorzusehen, durch die eine Modifikation von Prädikationen dargestellt werden kann.

Der Interpreter führt solange Selektions- und Exekutionszyklen über den Produktionen und der DB aus, bis der zu Beginn des Problemlösungsprozesses in der DB z.B. in Form von (GOAL x) abgelegte Zielzustand x (z.B. die Beantwortung einer Anfrage x in einem FAS) erreicht ist, was von dem Interpreter z.B. dadurch erkannt wird, daß die DB den Ausdruck (GOAL-COMPLETED x) enthält (vgl. 4.5.). Es sind zwei Arten von Interpretern zu unterscheiden, je nachdem, ob die Produktionen von links nach rechts (z.B. im META-DENDRAL-System) oder von rechts nach links (z.B. im MYCIN-System) gelesen werden, was einer Vorwärtsverkettung³³ bzw. einer Rückwärtsverkettung³⁴ entspricht.

Falls in der Menge der Produktionen und in der DB Repräsentationskonstruktionen für vages Wissen enthalten sind, besteht der Interpreter nicht nur aus Mechanismen zum Vergleich der entsprechenden Regelseiten mit dem gesamten Inhalt der DB, zur Variablenbindung und zur Ausführung des Aktionsteils einer Regel sondern auch aus Algorithmen, die abhängig vom Grad der Zuverlässigkeit des innerhalb einer Schlußkette verwendeten Wissens den Zuverlässigkeitsgrad des Inferenzergebnisses bestimmen.

Produktionensysteme in ihrer oben beschriebenen strengen Form werden hauptsächlich in der kognitiven Psychologie verwendet. In der KI wurden sog. Expertensysteme (vgl. 1.1.1.) entwickelt, in denen das durch Befragen von Experten oder Auswerten von Fachtexten gewonnene, oft vage Wissen über ein Sachgebiet u.a. in Form von Schlußregeln in der Wissensbasis repräsentiert ist. Da in diesen Systemen nur ein bestimmter Teil des Wissens in Form von Regeln dargestellt wird und bei der Codierung einiger Wissensquellen auch auf andere RS zurückgegriffen wird, spricht man hier oft von regel-orientierten Systemen (engl. rule-based systems), um diese Systeme mit heterogener Wissensbasis (vgl. 1.1.3) von anderen abzugrenzen, in denen reine Produktionensysteme benutzt werden.

Im folgenden beschäftigen wir uns mit Problemen der Repräsentation von vagem Wissen in regel-orientierten Systemen. Dabei verwenden wir Beispiele von Wissen aus dem medizinischen Bereich, wie sie von SHORTLIFFE74 im Zusammenhang mit dem MYCIN-System erörtert werden. Zunächst gehen wir von folgendem Beispiel einer vagen Inferenzregel aus (vgl. auch 1.3.3.):

BEISPIEL: Regel q: Falls S_1 : die Färbung des Organismus grampositiv ist,
 und S_2 : das Erscheinungsbild des Organismus kokkenförmig ist,
 und S_3 : die Wachstumsform des Organismus kettenförmig ist
 dann H: gibt es Anzeichen dafür (0.7), daß der Typ des Organismus Streptokokke ist.

³² vgl. SHORTLIFFE74,Kap.5, SHORTLIFFE/BUCHANAN75, DUDAetal.76

³³ vgl. auch Antezedens-Theoreme (PLANNER), 'if-added-methods' (CONNIVER), Dämonen (CHARNIAK76), 'condition-driven' PS (WATERMAN77)

³⁴ vgl. auch Konsequenz-Theoreme (PLANNER), 'if-needed-methods' (CONNIVER), Inferenzen zur Fragezeit (CHARNIAK76), 'action-driven' PS (WATERMAN77)

Die vage sprachliche Formulierung *gibt es Anzeichen dafür, daß* weist dabei auf die eingeschränkte Gültigkeit dieser Schlußregel hin.

Wir quantifizieren diese vage Formulierung und interpretieren sie in einem ersten, einfachen Ansatz als subjektive Wahrscheinlichkeit, d.h. als Wahrscheinlichkeit, die als Beschreibung der subjektiven Überzeugung einer Person von der Richtigkeit eines bestimmten Sachverhaltes und nicht wie bei empirischen Wahrscheinlichkeiten als Beschreibung zufälliger Experimente interpretiert wird.

Die Inferenzregel besagt dann folgendes: Falls für einen Organismus x $S_1(x) \wedge S_2(x) \wedge S_3(x)$ gilt, ist der Experte sich z.B. zu 70% sicher, daß H gilt. Transformieren wir die vereinfachte formale Darstellung der Regel $S_1 \wedge S_2 \wedge S_3 \xrightarrow{0.7} H$, in der die in einer prädikatenlogischen Schreibweise der Regel auftretende gebundene Variable 'Organismus' und der Allquantor nicht berücksichtigt werden, in eine bedingte Wahrscheinlichkeit, so können wir schreiben : (1) $P(H / S_1 \wedge S_2 \wedge S_3) = 0.7$. Aus (1) folgt (2) $P(\neg H / S_1 \wedge S_2 \wedge S_3) = 0.3$, da für bedingte Wahrscheinlichkeiten stets $P(\neg A/B) = 1 - P(A/B)$ gilt.

Übersetzt man (2) in die NS-Formulierung *Falls S_1, S_2 und S_3 gilt, ist man sich zu 30% sicher, daß H falsch ist* und legt diese Regel einem Experten zur Beurteilung vor, so lehnt der Experte diese Inferenzregel mit der Begründung ab, S_1, S_2 und S_3 sagten lediglich etwas über die Gültigkeit von H aber keinesfalls etwas über die Negation von H aus. Diese Beobachtung zeigt, daß die klassische Wahrscheinlichkeitstheorie und die auf ihr aufbauenden probabilistischen Logiken zur Darstellung von vagen Inferenzen kaum geeignet sind (vgl. auch SHORTLIFFE74).

Im folgenden werden zwei Formalismen dargestellt und verglichen, die eine Repräsentation und Verarbeitung von vagem Wissen in regel-orientierten Systemen ermöglichen: das Modell von SHORTLIFFE74, in dem mit Zuverlässigkeitsfaktoren (engl. certainty factors, kurz: CF-Modell) gearbeitet wird und der Ansatz von DUDAetal.76 (vgl. auch NILSSON75, S.219ff), der auf einer Erweiterung des Bayesschen Theorems für inkonsistente, subjektive Wahrscheinlichkeiten basiert.

4.2.2. DIE DARSTELLUNG VAGER INFERENZREGELN IM CF-MODELL

Im Gegensatz zu den klassischen, probabilistischen Logiken kann im CF-Modell jede vage Inferenzregel $S \rightarrow H$ durch zwei unterschiedliche Werte in bezug auf ihre Gültigkeit charakterisiert werden.

DEFINITION: $MB(H,S)=X$ $0 \leq MB(H,S) \leq 1$ $:\Leftrightarrow$ Wenn S gilt, steigert sich die Überzeugung des Experten, daß H wahr ist, um X .

DEFINITION: $MD(H,S)=X$ $0 \leq MD(H,S) \leq 1$ $:\Leftrightarrow$ Wenn S gilt, steigert sich die Überzeugung des Experten, daß H falsch ist, um X .

Die beiden Angaben $MB(H,S)$ und $MD(H,S)$ werden zu einem Zuverlässigkeitsfaktor kombiniert.

DEFINITION:

Der Zuverlässigkeitsfaktor $CF(H,S)$ ist definiert durch:

$$CF(H,S) \triangleq MB(H,S) - MD(H,S) \quad -1 \leq CF(H,S) \leq 1$$

$CF(H,S) > 0$ bedeutet, daß S stärker für als gegen H spricht, während $CF(H,S) < 0$ besagt, daß S von einem Experten eher als Widerspruch zu H empfunden wird.

Bei der Anwendung des CF-Modells auf ein regel-orientiertes System ergibt sich folgende Darstellung von vagem Wissen:

- Zu jeder Produktion $S \rightarrow H$ existiert in der Wissensbasis eine Angabe $CF(H,S)$. Falls $CF(H,S) > 0$ gilt, wird vom System $MB(H,S) = CF(H,S)$ und $MD(H,S) = 0$ angenommen, d.h. falls die Prämisse S nicht erfüllt ist, hat die Produktion keinen Einfluß auf die Gültigkeit von H .

- Sei $E(t)$ die Menge aller gültigen Aussagen in der DB zum Zeitpunkt t . Dann existieren für jede Proposition H in der DB Angaben über $MB(H, E(t))$ und $MD(H, E(t))$.

Die Verarbeitung des vagen Wissens erfolgt durch den Interpreter, in dem u.a. die im folgenden genannten Funktionen implementiert sind.

(1) Zusatzinformationen und Regelkonflikte

Wenn ein PS innerhalb eines dialogfähigen Systems eingesetzt wird, ist es möglich, daß, während das System die Gültigkeit einer Aussage H überprüft, laufend neue Informationen in der DB abgelegt werden, die z.B. vom Benutzer des Systems als Antwort auf eine Rückfrage eingegeben wurden. Wenn S_2 eine neu eingegebene Information ist, die für H spricht, oder $S_2 \rightarrow H$ eine weitere Regel ist, die H stützt (Regelkonflikt), dann definieren wir:

DEFINITION:

$$MB(H, S_1 \wedge S_2) \hat{=} \begin{cases} 0 & \text{falls } MD(H, S_1 \wedge S_2) = 1 \\ MB(H, S_1) + MB(H, S_2) \cdot (1 - MB(H, S_1)) & \text{sonst} \end{cases}$$

Falls die Zweifel, daß H gilt ($1 - MB(H, S_1)$) noch groß waren, als nur S_1 bekannt war, wird die neue Information S_2 durch die Multiplikation stärker berücksichtigt als bei geringerem Zweifel in H .

Entsprechendes gilt auch für $MD(H, S_1 \wedge S_2)$:

DEFINITION:

$$MD(H, S_1 \wedge S_2) \hat{=} \begin{cases} 0 & \text{falls } MB(H, S_1 \wedge S_2) = 1 \\ MD(H, S_1) + MD(H, S_2) \cdot (1 - MD(H, S_1)) & \text{sonst} \end{cases}$$

(2) Konjunktion von vagen Aussagen

Die Stärke der Überzeugung, daß die Konjunktion zweier vager Aussagen wahr ist, hängt von der am wenigsten überzeugenden Aussage ab:

DEFINITION:

$$MB(H_1 \wedge H_2, E) \hat{=} \text{MIN}(MB(H_1, E), MB(H_2, E))$$

Die Stärke der Überzeugung, daß die Konjunktion zweier vager Aussagen falsch ist, hängt von der am stärksten bezweifelten Aussage ab:

DEFINITION:

$$MD(H_1 \wedge H_2, E) \hat{=} \text{MAX}(MD(H_1, E), MD(H_2, E))$$

(3) Disjunktion von vagen Aussagen

Die Disjunktion wird komplementär zur Konjunktion definiert:

DEFINITION:

$$MB(H_1 \vee H_2, E) \hat{=} \text{MAX}(MB(H_1, E), MB(H_2, E))$$

$$MD(H_1 \vee H_2, E) \hat{=} \text{MIN}(MD(H_1, E), MD(H_2, E))$$

(4) Partiiell erfüllte Prämissen einer vagen Inferenzregel

$MD'(H, S_1)$ und $MB'(H, S_1)$ seien die charakteristischen Werte der Produktion $S_1 \rightarrow H$. Falls die Prämisse S_1 nur partiell erfüllt ist, d.h. $CF(S_1, E) < 1$, wird der Wert der Konklusion folgendermaßen berechnet:

DEFINITION:

$$MB(H, S_1) \triangleq MB'(H, S_1) \cdot MAX(0, CF(S_1, E))$$

$$MD(H, S_1) \triangleq MD'(H, S_1) \cdot MAX(0, CF(S_1, E))$$

Diese Definition entspricht dem multiplikativen Abtrennungsoperator in einer F-Logik (vgl. 4.1.3.).

Die Arbeit des Interpreters soll an folgendem Beispiel erklärt werden :

BEISPIEL:

Sei $DB(t)$ der Inhalt der DB zum Zeitpunkt t . Wir gehen davon aus, daß der Interpreter wie im MYCIN-System eine Rückwärtsverkettung vornimmt und daß die Gültigkeit einer Aussage H überprüft werden soll.

Zunächst sei $DB(t_1) = \{ (MB(S_1, E) = 1), (MB(S_2, E) = 1), (MB(S_3, E) = 0.6), (MD(S_1, E) = 0), (MD(S_2, E) = 0), (MD(S_3, E) = 0) \}$.

Wir nehmen an, daß der Interpreter als erste passende Regel R_1 findet.

$$R_1: \quad S_1 \wedge S_2 \wedge S_3 \xrightarrow{0.7} H$$

Die Prämisse S_3 von R_1 ist nur partiell erfüllt, da $MB(S_3, E) = 0.6$ in der DB enthalten ist. Nach (4) gilt:

$$MB(H, S_1 \wedge S_2 \wedge S_3) = MB'(H, S_1 \wedge S_2 \wedge S_3) \cdot MAX(0, CF(S_1 \wedge S_2 \wedge S_3, E))$$

$$\begin{aligned} CF(S_1 \wedge S_2 \wedge S_3, E) &= MB(H, S_1 \wedge S_2 \wedge S_3) - MD(S_1 \wedge S_2 \wedge S_3, E) \\ &= MIN(1, 1, 0.6) - MAX(0, 0, 0) = 0.6 \end{aligned} \quad \text{(nach (2))}$$

$$MD'(H, S_1 \wedge S_2 \wedge S_3) = 0 \quad \wedge \quad MD(H, S_1 \wedge S_2 \wedge S_3) = 0$$

$$MB(H, S_1 \wedge S_2 \wedge S_3) = 0.7 \cdot 0.6 = 0.42 \quad \text{(nach 4)}$$

$$CF(H, S_1 \wedge S_2 \wedge S_3) = 0.42$$

Der Zuverlässigkeitswert von H ist also zunächst 0.42.

Zum Zeitpunkt t_2 werden neue Fakten bekannt:

$$DB(t_2) = DB(t_1) \cup \{ (MB(S_4, E) = 1), (MD(S_4, E) = 0), (MB(S_5, E) = 1), (MD(S_5, E) = 0) \}$$

Der Interpreter findet eine weitere Regel, die H stützt:

$$R_2: \quad S_2 \wedge S_4 \xrightarrow{0.3} H$$

Der Zuverlässigkeitsfaktor von H steigt³⁵:

$$\begin{aligned} MB(H, DB(t_2)) &= MB(H, DB(t_1)) + MB(H, S_2 \wedge S_4) \cdot (1 - MB(H, DB(t_1))) \\ &= 0.42 + 0.3 \cdot 0.58 = 0.59 \quad (\text{nach (1)}) \end{aligned}$$

$$MD(H, DB(t_2)) = 0 \rightsquigarrow CF(H, S_1 \wedge S_2 \wedge S_3 \wedge S_4) = 0.59$$

Beim Durchsuchen der Menge von Produktionen stößt der Interpreter auf eine weitere Inferenzregel, die sich auf H bezieht:

$$R_3: \quad S_5 \xrightarrow{-0.1} H, \quad \text{d.h. } MD(H, S_5) = 0.1, \quad MB(H, S_5) = 0$$

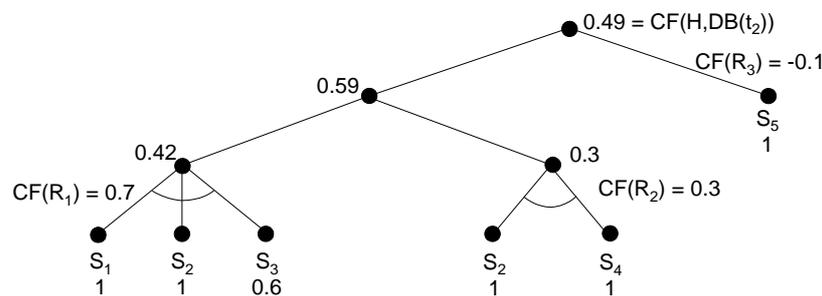
Da S₅ gilt, wird gemäß R₃ der Zuverlässigkeitsfaktor von H vermindert:

$$\begin{aligned} MD(H, DB(t_2)) &= D(H, DB(t_1)) + MD(H, S_5) \cdot (1 - MD(H, DB(t_1))) \quad (\text{nach (1)}) \\ &= 0 + 0.1 \cdot 1 = 0.1 \end{aligned}$$

$$\left. \begin{aligned} MB(H, S_1 \wedge S_2 \wedge S_3 \wedge S_4 \wedge S_5) &= 0.59 \\ MD(H, S_1 \wedge S_2 \wedge S_3 \wedge S_4 \wedge S_5) &= 0.1 \end{aligned} \right\} CF(H, S_1 \wedge S_2 \wedge S_3 \wedge S_4 \wedge S_5) = 0.49$$

Sind alle passenden Regeln angewandt worden und kommt in der DB keine neue, für H relevante Information mehr hinzu, so ergibt sich 0.49 als Zuverlässigkeitsfaktor von H.

Zusammenfassend läßt sich die Rückwärtsverkettung und Berechnung des CF von H für das behandelte Beispiel als Tiefensuche in einem Und/Oder-Graphen darstellen (vgl. Fig. 28).



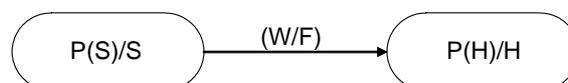
FIGUR 28

4.2.3. INFERENZNETZE ZUR REPRÄSENTATION VAGER SCHLUSSREGELN

DUDAetal.76 entwickeln ein formales Modell zur Repräsentation und Verarbeitung vager Inferenzregeln, das sich im Gegensatz zum CF-Modell eng an die traditionelle Wahrscheinlichkeitstheorie anlehnt.

Wir gehen hier lediglich auf eine der Varianten dieses Modells ein, die von SUTHERLAND76 in einem experimentellen System implementiert wurde. Dabei wird auf eine formale Herleitung des Modells aus dem Bayesschen Theorem verzichtet (vgl. DUDAetal.76).

Jede Produktion $S \rightarrow H$ wird als Teil eines bewerteten, gerichteten Graphen, der keine Schleifen und Kreise enthält und als Inferenznetz bezeichnet wird, repräsentiert:

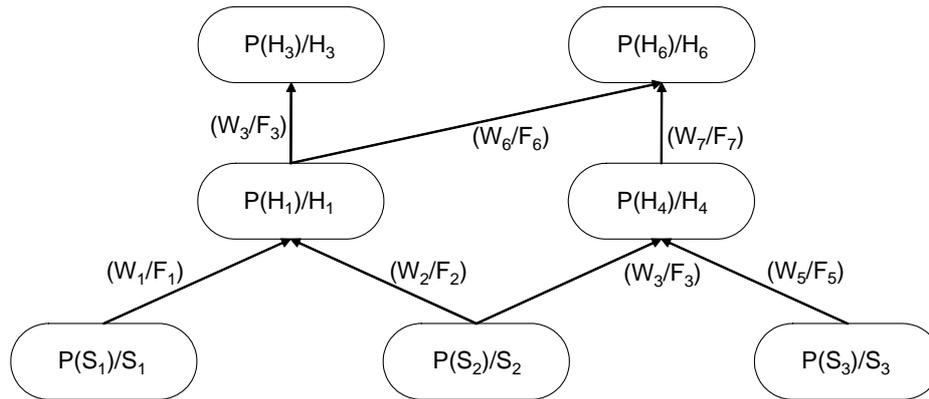


³⁵ Alle in diesem Kapitel angegebenen Zahlenwerte sind auf zwei Stellen hinter dem Komma gerundet.

Der Experte macht zu jeder eingegebenen Produktion $S \rightarrow H$ folgende Angaben:

- Die unbedingte Wahrscheinlichkeit $P(S)$ dafür, daß die Prämisse S wahr ist
- Die unbedingte Wahrscheinlichkeit $P(H)$ dafür, daß die Konklusion H wahr ist (d.h. die a-priori Wahrscheinlichkeit)
- Die Stärke W der Regel, die angibt, welchen Einfluß die Gültigkeit von S auf die Gültigkeit der Aussage H hat
- Die Stärke F der Regel, die angibt, welchen Einfluß die Gültigkeit der Negation von S auf die Gültigkeit von H hat.

Ein Inferenznetz hat z.B. die in Fig. 29 dargestellte Form:



FIGUR 29

In dem durch Fig. 29 wiedergegebenen Inferenznetz ist z.B. das Wissen repräsentiert, daß die Aussage H_3 zu einem gewissen Grad durch die Aussage H_1 gestützt wird, die wiederum durch S_1 und S_2 plausibel wird.

Die Angaben W und F werden formal so interpretiert, daß durch sie für jede Produktion die bedingten Wahrscheinlichkeiten $P(H/S)$ und $P(H/\bar{S})$ berechnet werden können:

$$(1) P(H/S) = \frac{W \cdot P(H)}{(W-1) \cdot P(H) + 1} \quad (2) P(H/\bar{S}) = \frac{F \cdot P(H)}{(F-1) \cdot P(H) + 1}$$

Da DUDAetal.76 wie auch SHORTLIFFE74 annehmen, daß alle Produktionen und alle Prämissen in den Produktionen voneinander unabhängig sind, können sie die für das CF-Modell besprochenen Probleme der Verarbeitung von Zusatzinformationen und Regelkonflikten sowie von Konjunktionen bzw. Disjunktionen vager Aussagen im Rahmen der traditionellen Wahrscheinlichkeitstheorie lösen. Lediglich das Problem der partiell erfüllten Prämissen einer vagen Inferenzregel bedarf einer besonderen Lösung.

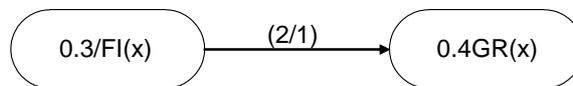
Der Benutzer eines auf einem Inferenznetz basierenden Systems gibt zunächst den Wert der bedingten Wahrscheinlichkeit $P(S/E)$ in das System ein. E seien alle Beobachtungen des Benutzers, die für den Wahrheitswert von S relevant sind. Der Inferenzmechanismus berechnet dann den Wahrheitswert der Konklusion als $P(H/E)$, wobei auf den Gültigkeitsgrad $P(S/E)$ der Prämisse und die Angaben zur Produktion $S \rightarrow H$ in der Wissensbasis zurückgegriffen wird:

DEFINITION: (3)

$$P(H/E) \triangleq \begin{cases} P(H/\bar{S}) + \frac{P(S/E)}{P(S)} \cdot P(H) - P(H/\bar{S}) & \text{falls } 0 \leq P(S/E) \leq P(S) \\ \frac{P(H) - P(H/S) \cdot P(S)}{1 - P(S)} + P(S/E) \cdot \frac{P(H/S) - P(H)}{1 - P(H)} & \text{falls } P(S) \leq P(S/E) \\ & P(S/E) \leq 1 \end{cases}$$

An folgendem einfachen Beispiel wird das Inferenzverfahren verdeutlicht:

BEISPIEL: Seien $FI(x)$ und $GR(x)$ Abkürzungen für die einstelligen Prädikate $fieber(x)$ bzw. $grippe(x)$. In der Wissensbasis sei folgende vage Inferenzregel enthalten:



$F=1$ besagt, daß diese Inferenzregel irrelevant ist, falls die Person x kein Fieber hat, da dann nach (2) gilt: $P(GR(x)/FI(x)) = P(GR(x))$.

Falls der Benutzer des Systems der Aussage $fieber(Fritz)$ aufgrund von Beobachtungen E den Wahrheitswert $P(FI(Fritz)/E) = 0.8$ zuordnet, berechnet der Interpretier $P(GR(Fritz)/E)$ nach (1), (2) und (3):

$$\begin{array}{ll}
 P(GR(Fritz)/FI(Fritz)) = 0.57 & \text{(nach (1))} \\
 P(GR(Fritz)/FI(Fritz)) = 0.4 & \text{(nach (2))} \\
 P(GR(Fritz)/E) = 0.56 & \text{(nach (3))}
 \end{array}$$

Obwohl die Prämisse der Regel, daß Fieber auf Grippe hinweist, für Fritz nicht voll erfüllt ist, ergibt die Anwendung der Inferenzregel für die These, daß Fritz Grippe hat, einen höheren Wahrheitswert als für die Grundwahrscheinlichkeit angenommen wurde, d.h. $P(GR(Fritz)/E) > P(GR(Fritz))$.

Die beiden betrachteten Ansätze werden von ihren Verfassern als stark vereinfachte Modelle der menschlichen Fähigkeit betrachtet, auch aus vagem Wissen durch plausibles Schließen (vgl. auch POLYA75) nützliche Aussagen abzuleiten. In beiden Modellen können auch Angaben eines Experten berücksichtigt werden, die in klassischen, wahrscheinlichkeitstheoretischen Inferenzmodellen zu Inkonsistenzen in der Wissensbasis führen würden.

Ein Vorteil des Ansatzes von DUDAetal.76 gegenüber dem CF-Modell besteht darin, daß er sich enger an die klassische Wahrscheinlichkeitstheorie anlehnt und damit theoretisch besser fundiert ist als das CF-Modell (vgl. auch WINSTON77, S. 243ff).

Als Nachteil der Darstellung von vagen Inferenzregeln nach dem Modell von DUDAetal.76 muß die Angabe der unbedingten Wahrscheinlichkeit für LS und RS einer Produktion gewertet werden. Diese Angaben sind für einen Experten nur schwer zu ermitteln.

Die Angabe eines CF zur Bewertung der Aussagekraft einer vagen Inferenzregel scheint natürlicher und für den Aufbau einer großen Wissensbasis praktikabler zu sein.

Sowohl der Grundinterpretier für PS (vgl. auch McDERMOTTetal.76) als auch die dargestellten Erweiterungen zur Verarbeitung von vagem Wissen können in einem NSS einfach und effizient implementiert werden.

4.3. DIE REPRÄSENTATION VON VAGEM WISSEN IN KI-PROGRAMMIERSPRACHEN

4.3.1. EINFÜHRENDE ÜBERSICHT

Ausgehend von der Vorstellung, daß auch Programmiersprachen als Repräsentationssprachen (vgl. 1.1.3.) eingesetzt werden können und daß '*Programming is expressing knowledge*' (vgl. HAYES75, S.566), wurden in der KI eine Reihe von Programmiersprachen entwickelt, deren Daten- und Kontrollstrukturen eine flexible Darstellung, eine modulare Organisation und eine effiziente Verarbeitung komplexer Wissensstrukturen fördern³⁶.

Diese Zielsetzung spiegelt sich z.B. auch in der Bezeichnung 'Knowledge Representation Language' (Abk.: KRL, vgl. BOBROW/WINOGRAD76) für eine der neueren KI-Programmiersprachen wieder (vgl. 4.6.2.).

Da PLANNER³⁷ (vgl. z.B. HEWITT71) die erste Programmiersprache war, die speziell für eine sog. prozedurale Einbettung von Wissen in KI-Systeme entwickelt wurde, werden KI-Programmiersprachen wie CONNIVER (vgl. z.B. McDERMOTT/SUSSMAN72), POPLER (vgl. DAVIES73), QA4 (vgl. z.B. RULIFSONetal.72), QLISP (vgl. z.B. WILBER76), 2. PAK (vgl. MELLI74) und PLASMA (vgl. z.B. HEWITT75) häufig auch als PLANNER-ähnliche Sprachen bezeichnet.

Da diese Sprachen eine Vielzahl von für die Konstruktion experimenteller KI-Systeme nützlichen Konzepten wie automatische Deduktionsmechanismen, 'Backtracking', Zugriff auf eine assoziative Datenbasis, komplexe Datenstrukturen und Kontextmechanismen bereits enthalten, befreien sie den Programmierer von der für ihn lästigen Beschäftigung mit Details, die ihn von seiner eigentlichen Problemstellung, der systematischen Repräsentation, Organisation und Verarbeitung von Wissen, zu weit wegführen würden.

Um die hier angesprochene Klasse von Sprachen von der Klasse der höheren Programmiersprachen (z.B. ALGOL68, PASCAL, SIMULA) abzugrenzen, spricht man u.a. auch von 'sehr hohen Programmiersprachen', die es dem Programmierer ermöglichen sollen, sich stärker auf das 'Was' als auf das 'Wie' bei der Implementation komplexer Systeme zu konzentrieren.

Im folgenden betrachten wir drei KI-Programmiersprachen, in denen spezielle RK für vages Wissen vorgesehen sind:

- FUZZY (vgl. LeFAIVRE74/75/76/77ab)
- DEDUCE (vgl. CHANG76)
- eine QA4-Erweiterung (vgl. WECHSLER75)

Dabei werden beim Leser Grundkenntnisse über Planner-artige Sprachen vorausgesetzt.³⁸ Da alle drei genannten Sprachen auf LISP basieren, erleichtern Kenntnisse über diese listenverarbeitende Programmiersprache das Verständnis der im folgenden diskutierten RK³⁹.

³⁶ Umgekehrt können auch die in 4.1. und 4.2. untersuchten RS Prädikatenkalkül und Produktionensysteme als Programmiersprachen angesehen werden (vgl. KOWALSKI74 bzw. WINOGRAD75b).

³⁷ Als Mikro-Planner implementiert (vgl. SUSSMANetal.70, BAUMGART72).

³⁸ vgl. die einführenden Übersichten von BOBROW/RAPHAEL74, LEAVENWORTH/SAMMET74, FOSTER77 oder auch: HAYES75, WINOGRAD74, Kap. 5, WINOGRAD75a, HEWITT75, SANDEWALL77

³⁹ vgl. die Einführungen in BÖCKERetal.75, SIKLOSSY76, KING/HAYES76, WINSTON77

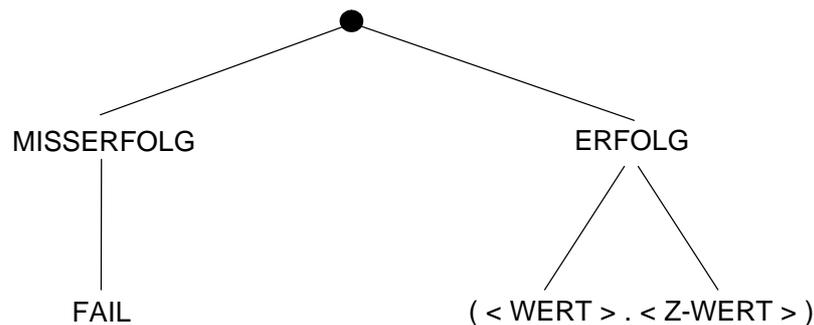
4.3.2. DIE MEHRWERTIGE PLANNER - ARTIGE PROGRAMMIERSPRACHE FUZZY

Die Programmiersprache FUZZY, die von LeFaivre auf den Rechenanlagen UNIVAC1110 und DECsystem10 (PDP-10) implementiert wurde, baut auf der Konzeption von FUZZY-Planner (vgl. KLING73) auf. FUZZY-Planner war ein Vorschlag zur Erweiterung von Mikro-Planner durch Konzepte einer F-Logik (vgl. 4.1.2.). FUZZY wurde bei der Implementation mehrerer kleiner KI-Systeme benutzt (vgl. LeFAIVRE74) und wird z.Z. auch in einem der größten Projekte zur Simulation von Überzeugungssystemen (dem BELIEVER-Projekt, vgl. SRIDHARAN76) erfolgreich eingesetzt (vgl. auch 1.3.2.).

Außerdem hat FUZZY die Konzeption eines 'activities system' (vgl. HARWOOD/HANNA76) beeinflusst, das in einer Erweiterung des 'Actor'-Prinzips (vgl. HEWITT75) u.a. den Austausch von in F-logischen Formeln codierten Nachrichten zwischen 'Actors' zuläßt. Auch Skuce schlägt eine Erweiterung der von ihm entwickelten Sprache KAL (für: Knowledge Aquisition Language) durch Konzepte aus FUZZY vor, um eine Darstellung von vagem Wissen zu ermöglichen (vgl. SKUCE75, S. 598).

4.3.2.1. KI-Programmiersprachen mit mehrwertiger Logik

Bei der Auswertung eines Ausdrucks in FUZZY lassen sich die beiden in Fig. 30 dargestellten Fälle unterscheiden. Das bei einer erfolgreichen Auswertung eines Ausdrucks entstehende Wertepaar besteht aus einem Wert und einem sog. Z-Wert, der den Wert modifiziert.



FIGUR 30

Der numerische Z-Wert kann u.a. als Wahrheitswert im Sinne einer F-Logik oder als Zugehörigkeitswert eines Elements einer F-Menge interpretiert werden (vgl. 3.1.2., 4.1.2.). FUZZY wird daher als eine mehrwertige KI-Programmiersprache bezeichnet⁴⁰.

BEISPIEL: Eine Auswertung des Ausdrucks (1) (GOAL (JUNG FRITZ)) durch den FUZZY-Interpreter ergibt z.B. das Wertepaar ((JUNG FRITZ). 0.8), wobei (JUNG FRITZ) als Wert und 0.8 als Z-Wert bezeichnet wird.

Dieses Ergebnis der Auswertung entspricht in einer F-Logik dem Ausdruck $T(jung(Fritz)) = 0.8$ oder in der Theorie der F-Mengen $\delta_{M(JUNG)}(FRITZ) = 0.8$.

Ist in einem NSS *jung* als Wert der linguistischen Variablen Alter (vgl. 3.2.3.) definiert und liegt dem System keine Information über das Alter von Fritz vor, so ergibt sich FAIL als Wert des Ausdrucks (1). Abhängig von der Umgebung, in der das FAIL auftritt, werden vom System automatisch bestimmte Prozesse (z.B. 'Backtracking') ausgelöst.

Der Wert FAIL ist von dem Auswertungsergebnis ((JUNG FRITZ).0) zu unterscheiden, welches besagt, daß Fritz nicht zur Trägermenge von $M(JUNG)$ gehört (vgl. 3.1.2.).

⁴⁰ Eine weitere mehrwertige Sprache, die u.a. in der Mustererkennung eingesetzt wurde, ist VL/1 (vgl. MICHALSKI 77).

Der Z-Wert liegt stets in dem Intervall [ZLOW,ZHIGH], das mit $[0,1] \subset \mathbb{R}$ als Standardwert initialisiert ist und durch entsprechende Elementaranweisungen undefiniert werden kann (vgl. LeFAIVRE74, S.44ff). Wenn in einem Ausdruck kein Z-Wert spezifiziert ist, wird der Wert von ZHIGH eingesetzt. Es stehen auch Sprachelemente zur Verfügung, die einen getrennten Zugriff auf die erste (VAL) und zweite Komponente (ZVAL) eines Wertepaares ermöglichen.

Auch in anderen KI-Programmiersprachen (z.B. QLISP, KRL) können Einträge in der assoziativen Datenbasis abgestufte Wahrheitswerte im Sinne einer F-Logik annehmen. Da die Eigenschaft TRUTHVALUE eines Eintrags in der Datenbasis von QLISP beliebige Ausdrücke als Wert haben kann, sind in QLISP mehrwertige Logiken realisierbar. Fikes implementierte z.B. ein KI-System mit einer 3-wertigen Logik in QLISP (vgl. FIKES75, S. 100). Im Gegensatz zu FUZZY stehen dem Programmierer in den genannten Sprachen allerdings keine elementaren Zugriffs-, Verknüpfungs- und Verarbeitungsprozeduren für die Wahrheitswerte einer F-Logik zur Verfügung.

Den wesentlichen Unterschied zwischen FUZZY und anderen KI-Programmiersprachen charakterisiert LeFAivre in folgenden Aussagen:

The user need not explicitly manipulate Z-values as he would have to in a traditional language - the system does all the dirty work of adding, removing, comparing, and combining Z-values.

(LeFAIVRE74, S.46)

Indeed, in many cases the user need not even be aware that he is working with fuzzy information.

(LeFAIVRE76, S.1073)

4.3.2.2. Die Repräsentation und Verarbeitung von vagen sprachlichen Ausdrücken in FUZZY

Eine einfache, propositions-zentrierte und deklarative Darstellung von vagem Wissen wird ermöglicht durch die in FUZZY integrierte assoziative Datenbasis, die aus einer Menge von beliebig verschachtelten und optional durch Z-Werte modifizierten Listen besteht.

So kann eine extensionale Darstellung, der in Abschnitt 3.2.1. betrachteten F-Menge *Mittelklassewagen* durch folgende Ausdrücke in der Datenbasis aufgebaut werden. Dabei dient ADD als Grundoperation für das Einfügen von Listen in die assoziative Datenbasis.

BEISPIEL: DB₁ (ADD (MITTELKLASSEWAGEN FIAT-500) . 0)
 (ADD (MITTELKLASSEWAGEN GOLF) . 0.7)
 (ADD (MITTELKLASSEWAGEN OPEL-REKORD))
 (ADD (MITTELKLASSEWAGEN MERCEDES) . 0.3)
 (ADD (MITTELKLASSEWAGEN ROLLS-ROYCE) . 0)

Der Zugriff auf die Datenbasis erfolgt durch die Angabe eines Patterns in einem FETCH-Ausdruck. Wertaufnehmende Variablen werden durch '?' und wertabgebende Variablen durch '!' präfigiert⁴¹.

BEISPIEL: (DO (FETCH (MITTELKLASSEWAGEN ?AUTO)) (PRINT !(DER !AUTO IST EIN TYPISCHER MITTELKLASSEWAGEN))) $\xrightarrow{DB_1}$ (DER OPEL-REKORD IST EIN TYPISCHER MITTELKLASSEWAGEN)

Dabei bedeutet $X \xrightarrow{DB_n} Y$ daß X bei dem Zustand der Datenbasis, der in den vorangegangenen Beispielen durch DB_n gekennzeichnet wurde, zu Y ausgewertet wird.

⁴¹ Listensegmente werden mit '??' bzw. '!!' präfigiert.

In diesem Beispiel wird die nach absteigenden Z-Werten geordnete Datenbasis durchsucht. Es wird das Element einer F-Menge mit dem höchsten Zugehörigkeitswert an die Variable ?AUTO gebunden. Durch die allgemeine Form von FETCH-Ausdrücken (FETCH <Pattern> (<Von>,<Bis>)) können komplizierte Suchprozesse ausgelöst werden, in denen die Datenbasis nach Listen, die ein komplexes Pattern⁴² erfüllen und Z-Werte zwischen < Von > und < Bis > haben, durchsucht wird.

BEISPIEL: (FOR FETCH: (MITTELKLASSEWAGEN ?AUTO) ZVAL: (0.2, 0.9) (PRINT !AUTO)) $\xrightarrow{DB_1}$ MERCEDES GOLF FAIL⁴³

(FETCH (MITTELKLASSEWAGEN ?AUTO) (0.2 , 0.1)) $\xrightarrow{DB_1}$ FAIL

Im einfachsten Fall erfolgt eine FETCH-Operation mit konstantem Pattern:

(FETCH (MITTELKLASSEWAGEN GOLF)) $\xrightarrow{DB_1}$ ((MITTELKLASSEWAGEN GOLF) 0.7)

Durch REMOVE werden Einträge in der Datenbasis gelöscht.

Meist wird eine F-Menge nicht extensional definiert, sondern ihre Zugehörigkeitsfunktion wird in Abhängigkeit von Eigenschaften der Elemente ihrer Trägermenge definiert (vgl. 3.2.1.). Eine solche implizite Definition einer F-Menge ermöglicht FUZZY durch sog. DEDUCE-Prozeduren, die z.B. den Konsequenz-Theoremen von Mikro-PLANNER entsprechen.

BEISPIEL: Sei SFUNK wie in Abschnitt 3.2.1. als FUZZY-Funktion definiert. Die referentielle Bedeutung von M(JUNG) kann durch folgenden FUZZY-Ausdruck repräsentiert werden:

- (1) (ADD DEDUCE:
- (2) (PROC (JUNG ?PERSON)
- (3) (FETCH (ALTER !PERSON ?JAHRE))
- (k) (SUCCEED (JUNG !PERSON) (*DIF 1 (SFUNK !JAHRE 20 40))))

Der Teilausdruck (1) bewirkt, daß der in (2)-(4) definierte Teilausdruck als prozedurales Wissen in der Datenbasis abgelegt wird. In Zeile (2) wird das für die DEDUCE-Prozedur charakteristische Pattern (JUNG ?PERSON) spezifiziert. Durch (3) und (4) wird der Prozedurrumpf definiert. In (3) wird zunächst das Alter der Person in Jahren durch Zugriff auf die assoziative Datenbasis bestimmt. In (4) wird das Wertepaar, welches das Ergebnis der Prozedur darstellt, konstruiert. Dabei wird der Z-Wert als zweite Komponente des Wertepaares durch die Auswertung der Funktion SFUNK berechnet.

Folgt nun auf die Anweisung DB2: (ADD (ALTER UDO 30)) der Ausdruck (GOAL (JUNG UDO)) $\xrightarrow{DB_2}$ ((JUNG UDO) . 0.5), so führt ein Pattern-gesteuerter Prozeduraufruf (engl.: pattern directed invocation) zu einer Suche nach DEDUCE-Prozeduren, deren Rumpf evaluiert wird, falls ihr charakteristisches Pattern die Liste (JUNG UDO) matcht.

Neben der Darstellung durch extensional oder implizit definierte F-Mengen kann auch eine partiell extensionale Definition einer F-Menge zur Repräsentation der Bedeutung eines vagen sprachlichen Ausdrucks benutzt werden. Dabei können auch mehrere DEDUCE-Prozeduren zur Repräsentation eines vagen Konzeptes herangezogen werden.

BEISPIEL: Folgende DEDUCE-Prozedur repräsentiert das zusätzliche Wissen, daß die Eigenschaft 'hat-gelichtetes-Haar' auf einen geringen Zugehörigkeitswert zur F-Menge der jungen Männer hindeutet.

⁴² Auf eine Darstellung des komfortablen Pattern-Matchers und anderer Sprachelemente von FUZZY, die sich nicht unmittelbar auf die Repräsentation von vagem Wissen beziehen, wird hier verzichtet (vgl. (LeFAIVRE74. S.42ff).

⁴³ FOR ist ein Repetitionsoperator.

4.3.2.3 Die Darstellung vager Inferenzregeln in FUZZY

Nachdem wir gezeigt haben, wie vage Propositionen, die extensionale Bedeutung vager sprachlicher Ausdrücke und Operatoren über vagen Konzepten in FUZZY repräsentiert werden können, gehen wir im folgenden auf die Darstellung vager Inferenzregeln (vgl. auch 4.1.3.) ein. Dabei stehen zwei Probleme im Vordergrund:

- die Darstellung der 'Stärke' einer Implikation
- die Spezifikation eines geeigneten Abtrennungsoperators

In FUZZY kann eine vage Inferenzregel durch eine DEDUCE-Prozedur dargestellt werden, in der zwei optionale Angaben ZVAL: und DEMON: im Prozedur-Kopf zur Spezifikation der Implikationsstärke bzw. des Abtrennungsoperators dienen.

BEISPIEL: Das bereits in Abschnitt 2.1.1. als Beispiel einer vagen Inferenz behandelte Paradoxon von der kleinen Zahl läßt sich mit Hilfe der F-Logik in FUZZY folgendermaßen auflösen (vgl. auch LeFAIVRE74, S.89).

Zunächst wird die erste Prämisse durch (ADD (KLEIN 1)) in der assoziativen Datenbasis abgelegt. Folgende rekursive Prozedur repräsentiert die im Paradoxon von der kleinen Zahl verwendete Inferenzregel. Wir gehen dabei davon aus, daß $T(\text{kleine-Zahl}(n) \Rightarrow \text{kleine-Zahl}(n+1)) = 0.9$ gilt.

- (1) (ADD DEDUCE:
- (2) (PROC DEMON: *ABTRENN ZVAL: 0.9
- (3) (KLEIN ?ZAHL)
- (4) (GOAL (KLEIN &(SUB1 !ZAHL))))))

Der Prozedur-Kopf besteht aus Zeile (1)-(3), der Prozedurrumpf ist der Teilausdruck in Zeile (4). Der Operator '&' erzwingt eine Evaluation des Operanden. SUB1 subtrahiert 1 von !ZAHL. Bei der Auswertung des Ausdrucks

(GOAL (KLEIN 10000))

ergibt sich der geringe Z-Wert von 0.9^{9999} , da für jede Anwendung des modus ponens entsprechend dem multiplikativen Abtrennungsoperator (*ABTRENN) 'gezahlt' werden muß (vgl. auch 4.1.3.). Eine Schlußfolgerung, die in einer zweiwertigen Logik zu einem Paradoxon führt, wird hier also bei einer Formulierung in der F-Logik zur validen Deduktion.

Die im Beispiel demonstrierte, kompakte Darstellung einer vagen Inferenzregel wird in FUZZY durch das Konzept der Prozedur-Dämonen (engl.: procedure demon) ermöglicht. Das Konzept der Prozedur-Dämonen wurde durch FUZZY neu in die KI-Programmiersprachen eingeführt und muß von den bekannten Dämon-Prozeduren (z.B. Antezedens-Theoreme in Mikro-PLANNER) unterschieden werden. Prozedur-Dämonen wie *ABTRENN in Zeile (2) des letzten Beispiels operieren über bei der sequentiellen Auswertung einer Prozedur entstehenden Z-Werten jedes einzelnen in der Prozedur vorkommenden unterbrechbaren Ausdrucks. Dagegen operieren Dämon-Prozeduren stets über der gesamten assoziativen Datenbasis des Systems.

DEFINITION: Ein Prozedur-Dämon ist eine LISP-Funktion X, die an eine FUZZY-Prozedur Y gekoppelt ist. Während der Auswertung von Y wird die Kontrolle nach der Evaluation jedes einzelnen Ausdrucks von Y an X übergeben. Der Prozedur-Dämon führt dann Operationen über folgenden Parametern aus:

- Resultat der Auswertung des Ausdrucks
- an die FUZZY-Prozedur gebundener Schwellwert (ZVAL:)
- akkumulierter Z-Wert, der dynamisch verwaltet wird

Nach diesen Operationen gibt der Prozedur-Dämon die Kontrolle wieder an den Prozess ab, der Y auswertet. Wird Y verlassen, so wird die Kontrolle zum letzten Mal an X mit dem Wert DONE übergeben. Da FUZZY in LISP eingebettet ist, kann der Programmierer beliebige Prozedur-Dämonen formulieren. Er hat damit die Möglichkeit, unterschiedliche, dem jeweiligen Inferenztyp angemessene Abtrennungsoperatoren (z.B. Minimum, Multiplikation, Mittelwert) zu definieren.

BEISPIEL: Der im letzten Beispiel benutzte multiplikative Abtrennungsoperator kann z.B. als folgender Prozedur-Dämon definiert werden:

```
(DEFPROP *ABTRENN
  (LAMBDA (WERT SCHWELLWERT AKKUM-Z-WERT)
    (COND ((EQ WERT FAIL) (FAIL))
          ((EQ WERT DONE) AKKUM-Z-WERT)
          (T (*TIMES SCHWELLWERT (ZVAL WERT)))))) EXPR)
```

Bei der Auswertung des Ausdrucks (GOAL (KLEIN 2)) wird zunächst der Prozedur-Dämon *ABTRENN aktiviert. Sobald der Ausdruck (GOAL (KLEIN 1)) im Rumpf der DEDUCE-Prozedur ausgewertet wurde, wird die Kontrolle an *ABTRENN übergeben. Der Prozedur-Dämon liefert als Ergebnis das Produkt aus dem Schwellwert 0.9 und dem Z-Wert 1, der bei der Evaluation von (GOAL (KLEIN 1)) durch Zugriff auf die Datenbasis entstanden ist. Der Prozedur-Dämon gibt dann die Kontrolle wieder an den die DEDUCE-Prozedur auswertenden Prozeß zurück, der in unserem einfachen Beispiel mit DONE die Evaluation abschließt. *ABTRENN wird erneut ausgewertet. Es ergibt sich der Z-Wert 0.9, der in das Ergebnis ((KLEIN 2) .0.9) einfließt.

Das Konzept der Prozedur-Dämonen führt dazu, daß in FUZZY die allgemeine Forderung nach einem einheitlichen Niveau der Repräsentation verletzt wird, da die Definition von Prozedur-Dämonen auf eine Repräsentationsebene führt (normaler LISP-Code), die eine Stufe tiefer als die übrigen Sprachkonzepte liegt.

4.3.3. DEDUCE: VAGE PRÄFERENZBEDINGUNGEN IN EINER ANFRAGESPRACHE FÜR RELATIONALE DATENBASEN

Eine weitere KI-Programmiersprache, bei deren Entwurf, wenn auch in viel geringerem Maße als in FUZZY, die Repräsentation von vagem Wissen berücksichtigt wurde, ist die von Chang entwickelte Sprache DEDUCE (vgl. CHANG 76). Ein wesentlicher Unterschied zwischen DEDUCE und allen anderen bisher erwähnten KI-Programmiersprachen besteht darin, daß DEDUCE nicht auf einer kleinen assoziativen Datenbasis, die vollständig im Hauptspeicher gehalten werden kann, sondern auf einer großen, auf Sekundärspeichern gelagerten Datenbasis arbeitet (vgl. auch SANDEWALL75, McDERMOTT75).

DEDUCE kann u.a. als eine in LISP als Gastsprache eingebettete Datensprache für relationale Datenbasen (vgl. z.B. DATE75) eingesetzt werden. Wir betrachten hier lediglich den Teil der Sprache, in dem die Repräsentation von vagen sprachlichen Formulierungen möglich ist.

In DEDUCE treten F-Prädikate in Form von sog. Präferenz-Bedingungen auf. Präferenz-Bedingungen sind Teile von

- Anfragen an die relationale Datenbasis
- sog. Axiomen, durch die Relationen implizit definiert werden.

Dabei können vage Präferenz-Bedingungen zwei unterschiedliche Funktionen haben:

- als Sortierkriterium bei der Ausgabe der auf eine Anfrage hin gefundenen Daten
- als Wissen, das der Benutzer in eine Anfrage oder ein Axiom als Heuristik einbettet, um dem System eine effizientere Suche auf der Datenbasis zu ermöglichen.

Zunächst machen wir uns an einem Beispiel die Funktion von Präferenz-Bedingungen als Sortierkriterium innerhalb einer Anfrageformulierung klar.

BEISPIEL: Die in Fig. 31 dargestellte Relation VATER/SOHN sei in der Datenbasis abgespeichert.

VATER/SOHN	VATER	ALTER1	SOHN	ALTER2
	Fritz	20	Ulf	2
	Karl	61	Emil	35
	Udo	37	Bernd	11
	Udo	37	Peter	14
	Lutz	41	Uwe	16
	Karl	61	Edi	32
	Udo	37	Uwe	16

FIGUR 31

Die Anfrage *Nenne vor allem die Namen der Väter, die um die 40 Jahre alt sind* enthält die vage Präferenz um die 40 Jahre alt. In DEDUCE kann diese Anfrage folgendermaßen dargestellt werden:

(QUERY (VATER/SOHN *VATER=x ALTER1 = y) \$(UM y 40))

Das terminale Symbol '*' markiert die Spalte der Relation, deren Werte auszugeben sind. '\$' kennzeichnet Präferenz-Bedingungen, die aus beliebig komplexen Formeln einer F-Logik bestehen können. 'UM' ist als F-Prädikat in LISP mit Hilfe der Funktion PFUNK (vgl. 3.2.1.) definiert.

In der Antwort werden die Daten nach dem Wahrheitswert sortiert, der bei der Auswertung der Präferenz-Bedingungen entsteht. Für das Beispiel ergibt sich die Antwort:

(ANSWER (LUTZ UDO FRITZ KARL)).

Eine bedeutendere Anwendung der Präferenz-Bedingungen in Fragen ist ihr Einsatz zur Steuerung von Suchprozessen über der Datenbasis.

BEISPIEL: Die Frage *Finde den Namen eines Vaters, der mehr als einen Sohn im Alter zwischen 10 und 15 Jahren hat* löst Suchprozesse auf der Datenbasis aus, die durch das heuristische, vage Wissen, daß Väter mittleren Alters mehr Söhne als jüngere Väter und eher einen Sohn unter 15 Jahren als ältere Väter haben, evtl. effizienter gestaltet werden können.

Die Frage kann in DEDUCE nun so formuliert werden, daß nach der Bestimmung des Alters Y für jeden Vater X zunächst für den Vater, der die Präferenz-Bedingung \$(MITTLERES-ALTER Y) am besten erfüllt (vgl. 3.2.1.), die Bedingung überprüft wird, ob er mehr als einen Sohn zwischen 10 und 15 Jahren hat. Die Auswertung der Anfrage entspricht in DEDUCE also einer durch Präferenz-Bedingungen gesteuerten Tiefensuche nach Eintragungen in der relationalen Datenbasis, die den Bedingungen genügen, welche in der Anfrage als prädikatenlogische Formeln spezifiziert sind.

Die Evaluation der Frage ergibt für das Beispiel den Wert (ANSWER (UDO)), wobei für FRITZ, KARL und LUTZ die Anzahl und das Alter der Söhne nicht überprüft wurde, da durch die Präferenz-Bedingung die Suche zunächst auf die Einträge für UDO konzentriert wurde.

Da der Benutzer von DEDUCE meist nicht weiß, welcher Aufwand für die von ihm formulierten Präferenz-Bedingungen erforderlich ist, kann dieses von Chang implementierte Verfahren u.U. allerdings dazu führen, daß die Auswertung der Präferenz-Bedingungen aufwendiger wird als die bei einer Beachtung des heuristischen Wissens durch den Suchprozeß im günstigsten Fall zu erwartende Effizienzsteigerung gegenüber einer 'blinden' Suche.

Die beiden Verwendungsweisen von Präferenz-Bedingungen, die bei der Formulierung von Anfragen auftreten, lassen sich auch bei dem Einsatz von Präferenz-Bedingungen in sog. Axiomen unterscheiden, die z.B. den Konsequenz-Theorem in Mikro-PLANNER entsprechen und durch die Relationen, die nicht explizit in der Datenbasis enthalten sind, implizit durch einen Rückgriff auf

bestehende Relationen definiert werden können. Axiome haben in DEDUCE die Form von Inferenzregeln:

$$A_1 \wedge \text{\$}P_1 \wedge \dots \wedge \text{\$}P_{n-1} \wedge A_n \wedge \text{\$}P_n \Rightarrow B$$

Die Präferenzen dienen für den Fall, daß mehrere Elemente der angesprochenen Relationen die Prämissen der Regel erfüllen, zur Sortierung der Ergebniswerte. Außerdem können die Präferenz-Bedingungen $P_1 \dots P_n$ die Suchprozesse steuern, die durch die Auswertung der Prämissen $A_1 \dots A_n$ ausgelöst werden.

Eine Möglichkeit zur Repräsentation vager Inferenzregeln ist in DEDUCE nicht gegeben.

Zusammenfassend stellen wir fest, daß DEDUCE zwar eine der wenigen KI-Programmiersprachen ist, die RK zur Darstellung von vagem Wissen vorsieht, aber wegen der arbiträren Einschränkung solcher RK auf Präferenz-Bedingungen für eine allgemeine Darstellung von vagem Wissen kaum geeignet ist.

4.3.4. REPRÄSENTATIONSKONSTRUKTIONEN FÜR VAGES WISSEN IN EINER QA4-ERWEITERUNG

Ausgehend von Klings Vorschlag zur Erweiterung von Mikro-PLANNER zu FUZZY-Planner (vgl. KLING73) schlägt WECHSLER⁷⁵ einige Erweiterungen der KI-Programmiersprache QA4 (vgl. z.B. RULIFSONetal.72) vor, die eine einfache Repräsentation und Verarbeitung relativer Adjektive und linguistischer Hecken ermöglichen sollen. Ähnlich wie Shortliffe in seinem MYCIN-System (vgl. 4.2.2.) so versucht auch Wechsler vages Wissen auf dem Gebiet der Medizin formal zu repräsentieren.

Im Gegensatz zu Shortliffe, der in seinem CF-Modell ausschließlich die Vagheit von Inferenzregeln berücksichtigt hat, beschäftigt sich Wechsler mit der Darstellung der extensionalen Bedeutung vager sprachlicher Ausdrücke und damit mit den von Shortliffe bewußt ausgeklammerten linguistischen Problemen. Die beiden genannten Ansätze ergänzen sich also gegenseitig.

Obwohl Wechslers Arbeit m.E. kein ausgereiftes Konzept darstellt, wird sein Ansatz hier kurz diskutiert, weil er im Gegensatz zu den bisher betrachteten Modellen die Bedeutung vager Ausdrücke nicht als F-Mengen sondern als Intervalle über einer Basismenge repräsentiert.

In der folgenden Darstellung benutzen wir im Gegensatz zu Wechsler keine Beispiele für vage sprachliche Ausdrücke aus dem medizinischen Bereich (z.B. *hoher Blutdruck*) sondern Standardbeispiele wie *kleiner Basketballspieler* und *junger Mensch*, um dem Leser einen einfacheren Vergleich zu den bisher diskutierten Modellen zu ermöglichen.

Die Grundidee des Wechslerschen Modells besteht darin, die Bedeutung $M(x)$ eines relativen Adjektivs, das eine Eigenschaft P eines Individuums in einem bestimmten Kontext i beschreibt, als Intervall innerhalb einer Basismenge U , auf die sich P bezieht, darzustellen.

BEISPIEL: Sei $P = \text{Körpergröße}$, $U = \text{IN}$, $i = \text{Basketballspieler}$, $x = \text{klein}$.

Dann ergibt sich in Wechslers Modell folgende Darstellung der extensionalen Bedeutung $M(\text{KLEIN}:\text{BASKETBALLSPIELER})$:

$$M(\text{KLEIN}:\text{BASKETBALLSPIELER}) = \underset{\text{Basketballspieler}}{\overset{\text{Körpergröße}}{f}} (\text{KLEIN}) = ([170,190], \text{LB})$$

Durch diese Darstellung und den dazugehörigen Interpreter sollen die beiden folgenden Probleme gelöst werden:

- Wenn das System in der Datenbasis über Einträge für *Fritz ist Basketballspieler* und *Die Körpergröße von Fritz beträgt 180 cm* verfügt, kann es die Frage *Ist Fritz ein kleiner Basketballspieler* dadurch beantworten, daß es auf eine Wissensquelle zugreift, in der $M(\text{KLEIN}:\text{BASKETBALLSPIELER})$ als Intervall codiert ist und feststellt, ob der Wert Körpergröße von Fritz innerhalb des entsprechenden Intervalls liegt.

- Sollen die beiden Basketballspieler Fritz und Udo mit den Größen 180cm bzw. 190 cm als Reaktion auf die Frage *Ist Fritz ein größerer Basketballspieler als Udo?* verglichen werden, so wird die Angabe 'LB' so interpretiert, daß eine Ordnung der Werte innerhalb der Basismenge von links nach rechts einer absteigenden Tendenz in bezug auf die Gültigkeit des vagen Prädikats 'klein(x)' entspricht. 180 cm ergibt also einen besseren Wert in bezug auf *klein* als 190 cm. Für die Explikation der Bedeutung von Ausdrücken wie *mittleres Alter* sieht Wechsler anstatt von 'LB' die Angabe 'M' vor, die so interpretiert wird, daß die Mitte des Intervalls den besten Wert in bezug auf den entsprechenden vagen Ausdruck ergibt.

BEISPIEL: $M(\text{MITTLERES-ALTER:MENSCH}) = f_{\text{Mensch}}^{\text{Alter}}(\text{MITTLERES-ALTER}) = ([20,45], M)$

Linguistische Hecken stellt Wechsler nicht als Operatoren über der extensionalen Bedeutung vager Terme dar, sondern gibt die Bedeutung von aus Hecken und vagen Ausdrücken zusammengesetzten sprachlichen Termen jeweils explizit an.

BEISPIEL: $M(\text{sehr KLEIN:MENSCH}) = f_{\text{Mensch}}^{\text{Körpergröße}}(\text{sehr KLEIN}) = ([140,160], LB)$

Dieser bei der großen Zahl linguistischer Hecken und ihrer Kombinationsmöglichkeiten unpraktikable Ansatz, könnte durch folgendes Modell, das auf einer Intervallarithmetik (vgl. z.B. MOORE66) beruht, ersetzt werden.

Sei $M(x) = ([a,b], LB)$, $[a,b] \subset U$. Wir definieren dann:

$$M(\text{sehr } x) \triangleq ([a,b] - [\frac{b-a}{2}, \frac{b-a}{2}], LB)$$

$$M(\text{mehr-oder-weniger } x) \triangleq ([a,b] + [\frac{b-a}{2}, \frac{b-a}{2}], LB)$$

BEISPIEL: $M(\text{sehr sehr KLEIN:BASKETBALLSPIELER}) = ([150,170], LB)$
 $M(\text{mehr-oder-weniger KLEIN:BASKETBALLSPIELER}) = ([180,200], LB)$

Wechsler schlägt eine Erweiterung der QA4-Anweisungen PUT und EXISTS vor, die RK für die Bedeutung vager sprachlicher Ausdrücke in seinem Intervall-Modell ermöglichen sollen.

Die erweiterten PUT und EXISTS Anweisungen haben folgende Form:

(PUT <Syntakt. Form> <Indikator> <Wert> <Funktion>)
 (EXISTS <Syntakt. Form> <Indikator> <Wert> <Funktion>)

<Syntakt. Form> ist im hier betrachteten einfachsten Fall ein Bezeichner oder eine wertaufnehmende Variable ($\leftarrow X$).

<Indikator> gibt eine Eigenschaft an, die durch <Wert> beschrieben ist.

<Wert> ::= <Adjektiv> | <Hecke> <Adjektiv> | <numerischer Wert> <Funktion> gibt die extensionale Bedeutung von <Wert> an.

BEISPIELE: (1) (PUT FRITZ GRÖSSE 180 F_1) ; durch PUT wird Wissen in der Daten-
 (2) (PUT UDO GROSSE 170 F_1) ; basis abgelegt.

F_1 , ist die in einem bestimmten QA4-Kontext i (hier: Basketballspieler) definierte Funktion $f_i^{\text{Größe}}$ welche die extensionale Bedeutung aller aus Hecken und vagen Termen zusammengesetzten Ausdrücke, die sich auf die Eigenschaft 'Größe' beziehen können, darstellt.

Eine Frage wie *Ist Fritz ein sehr kleiner Basketballspieler?* wird von einem semantisch-orientierten Parser in die RK

(3) (EXISTS FRITZ GRÖSSE SEHR KLEIN F_1) übersetzt.

Nach einer Auswertung von (1) und (2) verläuft die Evaluation von (3) erfolgreich und ergibt den Wert (FRITZ GRÖSSE SEHR KLEIN).

Die Frage *Wer ist klein?*, im Kontext 'Basketballspieler' gestellt, wird als

(4) (EXISTS $\leftarrow X$ GRÖSSE KLEIN F_1) repräsentiert.

Eine Auswertung von (4) ergibt den Wert (UDO GRÖSSE KLEIN), da die Variable $\leftarrow X$ an den Wert gebunden wird, der am 'besten match' (vgl. auch 4.6.3.).

Nach der Anweisung (PUT KARL GRÖSSE ZIEMLICH GROSS F_1) kann der Ausdruck (EXISTS KARL GRÖSSE 190 F_1) nicht erfolgreich ausgewertet werden, d.h. es ergibt sich der Wert FAILURE.

Komplexe vage Konzepte (z.B. die Symptome für ein Magengeschwür) formalisiert Wechsler als Funktion von mehreren primitiven vagen Termen, die er, wie oben gezeigt, als Intervalle über den verschiedenen Basismengen darstellt. Die Kontextabhängigkeit der Bedeutung vager Ausdrücke versucht Wechsler über den Kontextmechanismus von QA4 (vgl. RULIFSONetal.72, S.81) zu erfassen, indem er z.B. an den oben benutzten Bezeichner F_1 in verschiedenen Kontexten (z.B. Jockey, Basketballspieler) unterschiedliche Funktionen bindet.

Zusammenfassend stellen wir fest, daß Wechsler Sprachkonstruktionen für eine QA4-Erweiterung vorschlägt, die eine Speicherung von vagen und präzisen Werten in der QA4-Datenbasis vereinfachen und einen direkten Vergleich von Patterns, die vage Formulierungen enthalten, mit entsprechenden präzisen Angaben ermöglichen.

Wechslers Vorschlag zur Darstellung der Hecken vom Typ 1 ist allerdings unpraktikabel, und die Darstellung der Hecken vom Typ 2 und von vagen Konzepten, zu denen keine numerische Basismenge existiert, bleibt in Wechslers System völlig unberücksichtigt.

Die Explikation der extensionalen Semantik vager Terme als Intervall stellt keine Alternative zu einer Repräsentation durch F-Mengen dar, weil die Theorie der F-Mengen theoretisch besser fundiert und allgemeiner ist.

4.4. DIE REPRÄSENTATION VON VAGEM WISSEN IN SEMANTISCHEN NETZEN

4.4.1. EINFÜHRENDE ÜBERSICHT

In den meisten bisher implementierten NSS wird eine der vielen Varianten eines semantischen Netzes (Abk.: SN) als deklarative Repräsentationssprache verwendet. Im einfachsten Fall kann ein SN als bewerteter, gerichteter Graph aufgefaßt werden, in dem semantische Einheiten, die als Knoten des Graphen repräsentiert werden, in verschiedenen semantischen Relationen zueinander stehen, die als bewertete, gerichtete Kanten dargestellt werden. In der Terminologie der Prädikatenlogik können wir auch sagen, daß die Knoten eines SN mit Individuensymbolen und die Kanten mit Prädikaten- oder Funktionssymbolen beschriftet sind.

Nachdem durch RAPHAEL68 und QUILLIAN68 die Grundlagen dieser RS erarbeitet worden waren, entwickelte man in der KI meist durch spezielle Anwendungsbereiche motiviert unterschiedliche Notationsformen, Verallgemeinerungen und Erweiterungen⁴⁴ des ursprünglich sehr einfachen Konzeptes eines SN.

Auch alle auf dem DECsystem10 des Instituts für Informatik in Hamburg bisher entwickelten NSS benutzen zumindest innerhalb einer ihrer Wissensquellen ein SN⁴⁵.

⁴⁴ z.B. konzeptuelle Dependenz-Graphen (SCHANK75), 'aktive' strukturelle Netze (NORMAN/RUMELHART75), Ähnlichkeitsnetze (WINSTON75), Hierarchische Netze (SCHWIND75), gerichtete, rekursive Bewertungsknoten-Hypergraphen (BOLEY76b)

⁴⁵ vgl. SNP (BOLEY74), HANSA (WITTIG75), HASE (SCHEFE76), HAM-RPM

Nach der Integration von Konzepten aus SN und KI-Programmiersprachen in den gegenwärtig intensiv diskutierten 'Frame'-ähnlichen RS (vgl. 4.6.) scheint der Höhepunkt der theoretischen und experimentellen Erforschung von SN bereits überschritten zu sein, obwohl diese RS m.E. für große, anwendungsorientierte NSS besonders im Zusammenhang mit einer Kopplung an relationale Datenbank-Systeme noch für längere Zeit von Bedeutung sein wird.

Da im folgenden die spezielle Fragestellung, welche Möglichkeiten zur Repräsentation von vagem Wissen es in SN gibt, untersucht wird, müssen beim Leser Grundkenntnisse über SN vorausgesetzt werden.⁴⁶

Obwohl bereits QUILLIAN68 in einer der ersten Arbeiten über SN die Bedeutung der Repräsentation von vagem Wissen erkannte (*It is the very vagueness of meaning of most language terms that makes them useful*, QUILLIAN68, S.234) und daher eine *ready facility for representing knowledge having a great deal of vagueness* (ebd.) forderte, wurden für SN bisher nur vereinzelt Repräsentationskonstruktionen für vages Wissen entworfen.

In den meisten Vorschlägen zu einer SN-Notation bleiben Repräsentationsprobleme, die sich aus der Vagheit von NS-Formulierungen ergeben, völlig unberücksichtigt. Dies bedeutet allerdings nicht, daß in diesen SN weder Knoten- noch Kantenbezeichnungen vage sprachliche Ausdrücke sein können, sondern lediglich, daß keine der linguistischen und logischen Struktur vager Terme entsprechende Interpretation dieser Bezeichnungen vorgenommen wird. Beispielsweise sind RK für durch relative Adjektive modifizierte Objekte wie Fritz $\xrightarrow{\text{hat-die-Eigenschaft}}$ groß in SN ohne implizite oder explizite Angabe einer Referenzmenge genauso inadäquat wie die prädikatenlogische Darstellung von *groß* als einstelliges Prädikat in der RK 'groß(Fritz)' (vgl. 4.1.4.).

Aus dem Auftreten von vagen sprachlichen Ausdrücken oder Hecken als Kanten- oder Knotenbezeichnungen in einem SN (z.B. bei NORMAN/RUMELHART75, SCHANK75, WITTIG75) kann also auf keinen Fall schon geschlossen werden, daß Probleme der Darstellung von vagem Wissen beim Entwurf dieser Netze berücksichtigt wurden. Dies hängt mit einer allgemeinen Gefahr der Fehleinschätzung von RK, in denen natürlichsprachliche Elemente auftreten, zusammen (vgl. auch McDERMOTT76): In die natürlichsprachlichen Bezeichner wird leicht eine Semantik 'hineininterpretiert', die den entsprechenden Komponenten eines NSS beim Zugriff auf die Wissenseinheit in dieser Form nicht zur Verfügung steht oder wie Hewitt es formuliert:

It is important not to be misled into reading more into the symbols in the code than is really there. In every case the semantics of the English symbols must be implemented elsewhere in other modules to realize the desired behavior.

(HEWITT75, S.189)

Von einer adäquaten Repräsentation von vagem Wissen kann erst dann gesprochen werden, wenn es möglich ist, aus den RK genau diejenigen Schlußfolgerungen abzuleiten, welche aus den repräsentierten Originalen (z.B. natürlichsprachliche Sätze) hergeleitet werden können.

Der Nachweis, daß eine bestimmte in der Literatur vorgeschlagene RK keine adäquate Darstellung von vagem Wissen erlaubt, kann besonders für SN oft nur unter großem Aufwand (evtl. Studium des gesamten Programms) erbracht werden, da die Autoren meist die Semantik ihrer semantischen Netze (vgl. WOODS75) nicht explizieren.

4.4.2. ELEMENTARE MODIFIKATIONEN DER KNOTEN UND KANTEN EINES SN

Nur in wenigen Arbeiten aus dem Bereich der KI werden Probleme der Darstellung von vagem Wissen in SN diskutiert. Es sei hier zunächst auf die Veröffentlichungen von QUILLIAN68, CERONE/SCHUBERT75, WAHLSTER/v.HAHN76 und vor allem auf die Berichte über die SCHOLAR-Projekte, die im Rahmen der Forschung zum Computer-gestützten Unterricht durchgeführt wurden (vgl. CARBONELL/COLLINS73, COLLINSetal.75), aufmerksam gemacht.

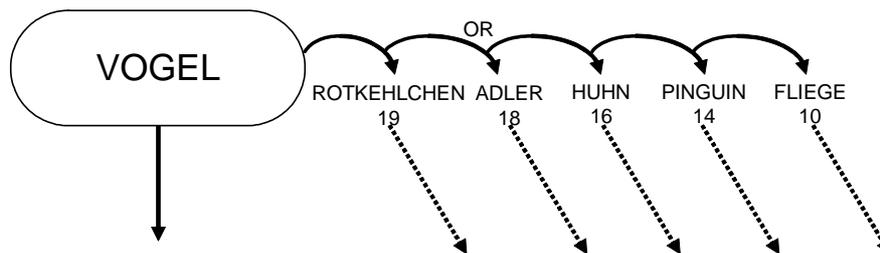
(WAHLSTER/V.HAHN76)

⁴⁶ vgl. die Einführungen in SIMMONS73, WOODS75, LENDERS75, BRACHMAN76, SCRAGG76

4.4.2.1. Gültigkeitsmarkierungen in Quillians SN

QUILLIAN68 benutzte bereits in seinem Semantic Memory-System SN, in denen Knoten des Graphen durch sog. Gültigkeitsmarkierungen (engl. range restriction tags) modifiziert werden können. Die Gültigkeitsmarkierungen I können neun verschiedene Zahlenwerte (0-8) annehmen, wobei 1 bis 8 unterschiedliche Zugehörigkeitswerte zu einem Konzept bzw. verschiedene Abstufungen der Gültigkeit einer Prädikation und 0 die Negation darstellt (vgl. QUILLIAN68, S. 231ff).

BEISPIEL: Ein Teil der Bedeutung des vagen Begriffs *Vogel* (vgl. 2.2.1.) kann durch Gültigkeitsmarkierungen dargestellt werden (vgl. Fig. 32).



FIGUR 32

Linguistische Hecken vom Typ 1 wie *sehr*, *ziemlich* usw. werden in den von Quillian vorgeschlagenen RK nicht explizit als Kanten- oder Knotenbezeichnung verwendet, sondern sie werden durch entsprechende Werte der Gültigkeitsmarkierungen dargestellt, welche sich auf die durch Hecken modifizierten Terme beziehen.

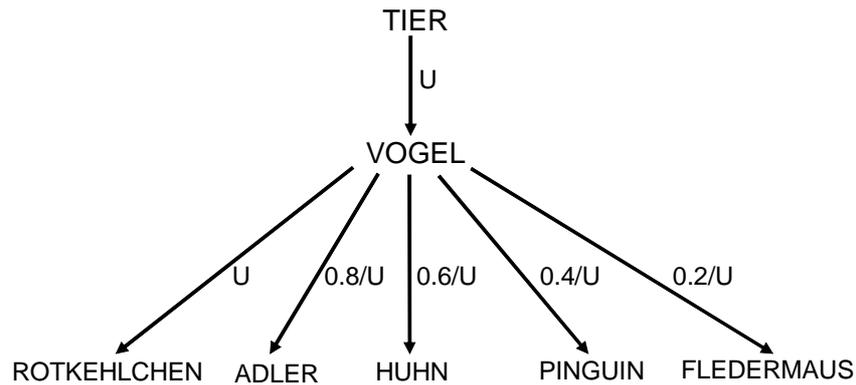
BEISPIEL: *ziemlich reich* kann in der von Quillian vorgeschlagenen RS als 'reich I4' dargestellt werden.

Die Werte der Gültigkeitsmarkierungen können von Sprachgenerierungsalgorithmen, die über den Netzen arbeiten, in linguistische Hecken umgewandelt werden (vgl. QUILLIAN68, S.241). Außerdem sieht Quillian für bestimmte Knoten sog. Relevanzmarkierungen (engl. criteriality tags) vor, durch die eine Klassifikation der einzelnen Bedeutungskomponenten eines vagen Ausdrucks ermöglicht wird, wie sie zur Repräsentation der Hecken vom Typ 2 (vgl. 3.2.2.) notwendig ist.

Während den Gültigkeits- und Relevanzmarkierungen im Semantic Memory-System wegen seiner beschränkten Inferenzkapazität noch eine relativ geringe Bedeutung zukommt, spielen in anderen Varianten von SN zusätzliche Markierungen eine wichtige Rolle bei der Darstellung von vagem Wissen.

4.4.2.2. F-Graphen in propositions- und objektzentrierter Darstellung

Wahlster/v.Hahn haben SN vorgeschlagen, in denen numerische Werte (kurz: F-Werte) aus dem reellen Intervall [0,1] die Kanten des Graphen, welche atomare, semantische Relationen darstellen, zusätzlich markieren (vgl. WAHLSTER/v.HAHN76, S.220). In den dadurch entstehenden F-Graphen (vgl. 3.1.2.) kann z.B. die Zugehörigkeitsrelation zur unscharfen Kategorie *Vogel* auf einfache Weise dargestellt werden (vgl. Fig. 33).



FIGUR 33

Dabei wird $X \xrightarrow{U} Y$ als $Y \subset X$ interpretiert (vgl. 4.4.3.).

In einer propositionszentrierten Darstellung (vgl. 1.1.3) kann dieses Teilnetz durch FUZZY-Anweisungen in der assoziativen Datenbasis aufgebaut werden:

```

(ADD (VOGEL U ROTKEHLCHEN))
(ADD (VOGEL U ADLER) 0.8 )
.
.
.
  
```

Bei einer objektzentrierten Darstellung dieser SN in Form von Eigenschaftslisten (engl. property lists), wie sie in LISP zur Verfügung stehen, wird zu jedem Knoten des Netzes eine Liste aller von ihm wegführenden Kanten abgespeichert. Dabei können zwei unterschiedliche RK benutzt werden:

- (1) (VOGEL ... 0.8-U (ADLER) 0.6-U (HUHN).....)
- (2) (VOGEL ... U ((0.8 ADLER) (0.6 HUHN)....)..)

Nur falls häufig auf α -Abschnitte der F-Relation U (vgl. 3.1.2.) zugegriffen werden soll, ist Darstellung (1) (2) vorzuziehen. Die in (2) verwendete Codierung ist mit Quillians Ansatz vergleichbar, da die zusätzlichen Markierungen mit den Bezeichnungen für die Knoten des Graphen verbunden sind.

4.4.2.3. I- und U-Markierungen im SCHOLAR-System

Neben den von BECKER⁷³ und BRACHMAN⁷⁶ verwendeten SN wird besonders in den SN, die für die SCHOLAR-Systeme (vgl. CARBONELL/COLLINS⁷³, COLLINS/WARNOCK⁷⁴, COLLINSetal.⁷⁵) entwickelt wurden, bei der Darstellung von vagem Wissen intensiv von zusätzlichen Markierungen der Kanten Gebrauch gemacht. Es wird zwischen sog. I- und U-Markierungen unterschieden, die den Wichtigkeits- und Detaillierungsgrad bzw. den Wahrheits- und Sicherheitsgrad einer im SN gespeicherten Information angeben. Die U-Markierungen können als Zugehörigkeitswerte von F-Relationen interpretiert werden.

Die I-Markierungen können sieben verschiedene Werte auf einer Skala von (10) bis (16) annehmen, wobei (10) eine elementare, aber sehr wichtige und (16) eine sehr detaillierte und daher meist unwesentliche Wissensseinheit kennzeichnet. In SCHOLAR werden zusätzlich Eigenschaftslisten von unterschiedlicher Schachtelungstiefe benutzt, um verschiedene Ebenen der Detaillierung in der Wissensbasis zu unterscheiden. Die absolute Relevanz einer Information, die auf der Schachtelungstiefe n einer bestimmten Eigenschaftsliste durch einen Suchprozeß auf der Wissensbasis gefunden wird, ist in SCHOLAR als Summe der n-1 bei einer sequentiellen Suche vor der aufgefundenen Information liegenden I-Markierungen definiert.

BEISPIEL: [PERU (LAGE (10) (IN (10) SÜDAMERIKA))
 (BEVÖLKERUNG (12) (SPRACHE (12)
 (OFFIZIELL (10) SPANISCH)
 (INDIANISCH(12) QUECHUA)))]...

Diese Liste ist ein Ausschnitt aus einer der Wissensbasen von SCHOLAR über die Geographie Südamerikas.

Für die Aussage *In Peru spricht die indianische Bevölkerung die Sprache Quechua* ergibt sich ein I-Wert von 6.

In SCHOLARS SN können sowohl vage als auch präzise Werte einer bestimmten Eigenschaft gespeichert werden (z.B. Angaben wie 'große Bevölkerungsdichte' bzw. die exakte Zahl der Einwohner pro km²), wobei den vagen Angaben geringe I-Werte und den präzisen hohe I-Werte zugeordnet werden. Die I-Werte werden von dialog- und partnertaktischen Prozeduren verarbeitet, die speziell auf die Anwendung von SCHOLAR im Computer-gestützten Unterricht abgestimmt sind. So benutzt das System ein stark vereinfachtes Benutzermodell, wenn es die Höhe der I-Werte, die den eingegebenen Aussagen des mit dem System interagierenden Schülers in SCHOLARS SN entsprechen, als Maß für den Wissensstand des Schülers benutzt und den Detaillierungsgrad seiner Fragen und Antworten wiederum unter Verwendung der I-Werte dem 'Niveau' des Dialogs anpaßt.

4.4.2.4. Unvollständige Mengen

Neben I- und U-Markierungen wird in SCHOLAR die Unterscheidung von vollständigen (engl. closed sets) und unvollständigen Mengen (engl. open sets, Abk.: U-Mengen) bei der Darstellung von vagem Wissen eingesetzt (vgl. CARBONELL/COLLINS73, S.347, COLLINSetal.75, S.385).

DEFINITION: Eine in der Wissensbasis eines NSS dargestellte Menge heißt genau dann unvollständig, wenn Metawissen existiert, welches besagt, daß nicht alle Elemente der Menge in der Wissensbasis repräsentiert sind. Alle nicht unvollständigen Mengen werden als vollständige Mengen bezeichnet.

BEISPIEL: In einer Wissensbasis eines NSS seien z.B. von der Menge *Fußballnationalmannschaft* nur die Elemente *Vogts* und *Maier* abgespeichert. Gleichzeitig sei bekannt, daß eine Fußballmannschaft aus 11 Spielern besteht. { *Vogts*, *Maier* } wird dann als U-Menge interpretiert⁴⁷.

Die meisten für vollständige Mengen gültigen Inferenzregeln müssen für unvollständige Mengen abgewandelt werden.

BEISPIEL: Falls die Namen von Hauptstädten in einem NSS als unvollständige Menge repräsentiert sind, muß die Frage *Ist Alpbach eine Hauptstadt?* für den Fall, daß kein entsprechender Eintrag für Alpbach gefunden wird, nicht mit *Nein* sondern mit *Das weiß ich nicht* beantwortet werden.

Die Autoren von SCHOLAR gehen davon aus, daß zumindest Allgemeinwissen zum größten Teil aus vagem, bruchstückhaftem Halbwissen besteht und folgern daraus, daß sich mehr unvollständige als vollständige Mengen bei der Repräsentation von Wissen ergeben. Bei solchen Randbedingungen ist es vorteilhaft, alle vollständigen Mengen im SN zu markieren.

Für anwendungsorientierte Experten- und Beratungssysteme, in denen möglichst vollständiges Wissen des Systems auf einem kleinen Gebiet angestrebt wird, ist die Markierung der unvollständigen Mengen, die sich selten völlig ausschalten lassen, günstiger.

⁴⁷ Der Begriff 'Teilmenge' trifft hier nicht zu, da keine bestimmte Obermenge definiert ist.

DEFINITION: Mikro- oder Makrowelten (vgl. 1.1.3.), die unvollständige Mengen enthalten, werden als offene Welten bezeichnet. Weltmodelle, in denen alle Objekte und alle Relationen zwischen den Objekten vollständig bekannt sind, werden als geschlossene Welten bezeichnet (vgl. Fig. 34).

GESCHLOSSENE WELT	OFFENE WELT
SHRDLU (WINOGRAD72)	SCHOLAR (CARBONELL/COLLINS73)
LUNAR (WOODSetal.72)	HAM-RPM (WAHLSTER/v.HAHN76)

FIGUR 34

Durch entsprechende zusätzliche Markierungen von Kanten- und Knotenbezeichnungen eines SU können also drei Arten von Mengen unterschieden werden:

- klassische Mengen
- F-Mengen
- U-Mengen

Die zur Darstellung von vagem Wissen geeigneten Konzepte der F-Menge und der U-Menge können auch kombiniert in Form einer unvollständigen F-Menge auftreten, d.h. in Form einer F-Menge, deren Trägermenge eine unvollständige Menge ist.

BEISPIEL: Sei $\underline{A} = \{ 0.8/\text{HAMBURG}, 1/\text{TOKIO}, 1/\text{NEW-YORK}, 0.9/\text{LONDON} \}$ als F-Menge der *Großstädte* in einem NSS definiert. Falls aus dem Metawissen des Systems hervorgeht, daß noch andere Elemente, die dem System nicht bekannt sind, zur Trägermenge von \underline{A} gehören, so ist \underline{A} eine unvollständige F-Menge.

F- und U-Mengen sind immer in bezug auf eine bestimmte Wissensbasis definiert und können interpersonell variieren (*Your closed set may be my open one.* COLLINSetal.75, S.385).

4.4.3. DIE DARSTELLUNG DER EXTENSIONALEN BEDEUTUNG VAGER KONZEPTE IN SN

4.4.3.1. Linguistische Variablen in erweiterten SN

In diesem Abschnitt wird gezeigt, wie die extensionale Bedeutung vager, relativer Adjektive in semantischen Netzen repräsentiert werden kann. Wir beziehen uns dabei auf die in dem NSS HAM-RPM (vgl. v.HAHNetal.76, WAHLSTER/v.HAHN76) verwendeten SN, die eine einfache Darstellung von linguistischen Variablen im Sinne Zadehs (vgl. 3.2.3.) ermöglichen.

In HAM-RPM werden sog. tiefenorientierte SN verwendet, d.h. als Kanten des Graphen treten nicht beliebige zweistellige Oberflächenprädikate auf, wie dies in den sog. oberflächenorientierten SN möglich ist, sondern nur eine gewisse Zahl von atomaren Relationen, die oft auch als 'epistemologisch primitive' Kanten bezeichnet werden.

Der Vorteil tiefenorientierter SN besteht darin, daß für die kleine Zahl von primitiven Kanten effiziente Such- und Inferenzprozeduren einfacher zu realisieren sind (vgl. auch v.HAHNetal.76, S.347). Wir gehen zunächst von folgenden in HAM-RPM definierten primitiven Kanten aus⁴⁸:

⁴⁸ In den NSS von BORGIDA75 und BRACHMAN76 werden primitive Kanten mit einer vergleichbaren Semantik benutzt.

Um eine Frage wie *Ist Fritz reich ?* zu beantworten, wird zunächst ausgehend von dem Knoten FRITZ nach REF-Kanten gesucht, welche zu Knoten führen, die ihrerseits über INST-Kanten mit dem Knoten VERMÖGEN verbunden sind, da REICH durch die entsprechende V-Kante als Wert der linguistischen Variablen *Vermögen* erkannt wird. Im Beispiel wird der Knoten mit der Bezeichnung 'DM225000' gefunden, der das Prädikat DM-NUMBERP erfüllt und daher als präzise Angabe interpretiert wird.

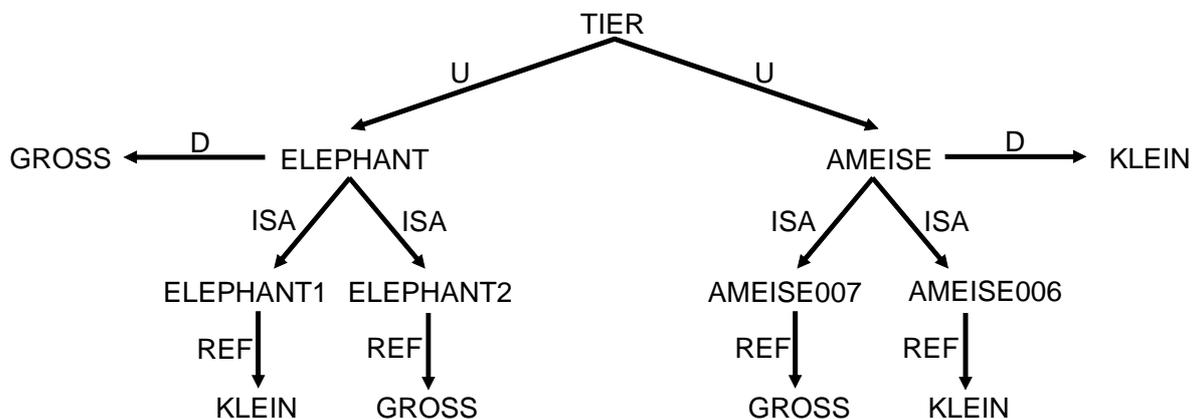
Über die M-Kante wird die Definition der extensionalen Bedeutung von *reich* aufgesucht. Nachdem der Wert, der sich bei der Auswertung des Ausdrucks (SFUNK 2.25 • 10⁴ 10⁴ 10⁶) ergibt, durch eine prozedural definierte Hecke (vgl. 4.3.2.2.) beschrieben wurde, kann z.B. die Antwort *Ja, Fritz ist ziemlich reich* generiert werden.

Auf die Frage *Welches Vermögen hat Fritz* kann abhängig vom Dialogzustand und der jeweiligen Dialog- und Partnertaktik (vgl. WAHLSTER/v.HAHN76, S. 217ff) entweder die vage Antwort *Fritz ist ziemlich reich* oder die präzise Antwort *DM 225000* generiert werden. Die gleiche Frage kann für Udo nur mit *Udo ist reich* beantwortet werden. Auf eine metakommunikative Äußerung des Benutzers wie *Bitte etwas genauer antworten !* reagiert das System in diesem Fall mit der Standardantwort *Genaueres weiß ich nicht*.

4.4.3.2. Die Referenzmengeninferenz in einer einfachen Generalisierungshierarchie

In Fig. 35 wurden für die in dem ESN auftretenden vagen, relativen Adjektive keine Referenzmengen spezifiziert. Wir zeigen im folgenden, durch welche Techniken die Referenzmengen relativer Adjektive in ESN eingeführt werden können. Wir gehen dabei davon aus, daß zwischen relativen und absoluten Adjektiven bereits in einer anderen Wissensquelle des NSS unterschieden wird. Außerdem setzen wir voraus, daß jedes ESN zunächst nur aus sog. einfachen Generalisierungshierarchien (vgl. auch WINOGRAD75b) besteht, d.h. daß es zu jedem Knoten im Netz höchstens eine von ihm wegführende K- oder zu ihm hinführende ISA-Kante gibt. Wir können dann die Referenzmengeninferenz anwenden, die am folgenden Beispiel erläutert werden soll.

BEISPIEL: Das in Fig. 36 dargestellte Teilnetz besteht aus einer einfachen Generalisierungshierarchie.



FIGUR 36

Die Referenzmengeninferenz basiert auf folgenden Regeln:

$$\begin{array}{ll}
 X \in \text{Tokens}^{49} & \Rightarrow \text{Referenzmenge} := \text{Type}(X) \\
 X \in \text{Types} & \Rightarrow \text{Referenzmenge} := \text{Obermenge}(X)
 \end{array}$$

⁴⁹ Tokens sind in den SN durch angehängte Ziffern gekennzeichnet.

Diese Regeln führen zu folgenden Interpretationen der vagen, relativen Adjektive in Fig. 36 :

Elephanten sind groß im Vergleich zu anderen Tieren.
Elephant1 ist klein im Vergleich zu anderen Elephanten.
Ameisen sind klein im Vergleich zu anderen Tieren.
Ameise007 ist groß im Vergleich zu anderen Ameisen.

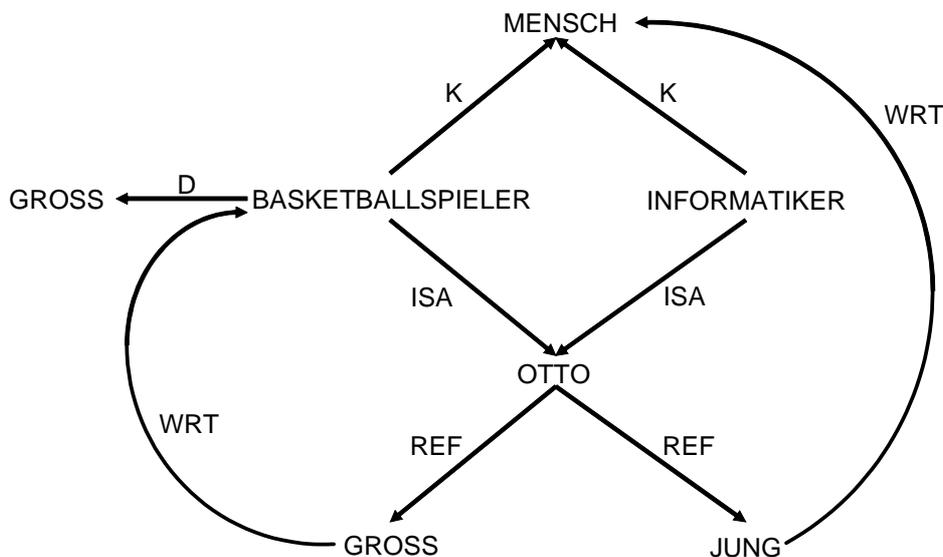
In dem in Fig. 36 dargestellten Teilnetz sind Referenzmengen also nur implizit durch eine spezielle Interpretationskonvention repräsentiert.

Wenn das Prinzip der einfachen Generalisierungshierarchie für Teile eines SN verletzt ist, dann müssen die Referenzmengen explizit gekennzeichnet werden. Zu diesem Zweck führen wir zusätzlich die elementare Kante $X \xrightarrow{\text{WRT}} Y$ ein⁵⁰.

In Fig. 37 ist für den Knoten mit der Bezeichnung OTTO das Prinzip der einfachen Generalisierungshierarchie verletzt, da Otto sowohl Basketballspieler als auch Informatiker ist. Die Referenzmengeninferenz kann nicht angewandt werden. Das NSS erkennt diese Situation an der Existenz von WRT-Kanten. Das in dem dargestellten Teilnetz mit Hilfe der WRT-Kanten repräsentierte Wissen kann folgendermaßen paraphrasiert werden:

- (1) *Otto ist groß für einen Basketballspieler.*
- (2) *Otto ist ein junger Mensch.*

Außerdem ist implizit auch das als *Otto ist ein besonders großer Informatiker* verbalisierbare Wissen in dem Teilnetz enthalten, da diese Aussage aus (1) und der ebenfalls im Netz enthaltenen Information *Ein Basketballspieler ist groß im Vergleich zu anderen Menschen* inferiert werden kann.



FIGUR 37

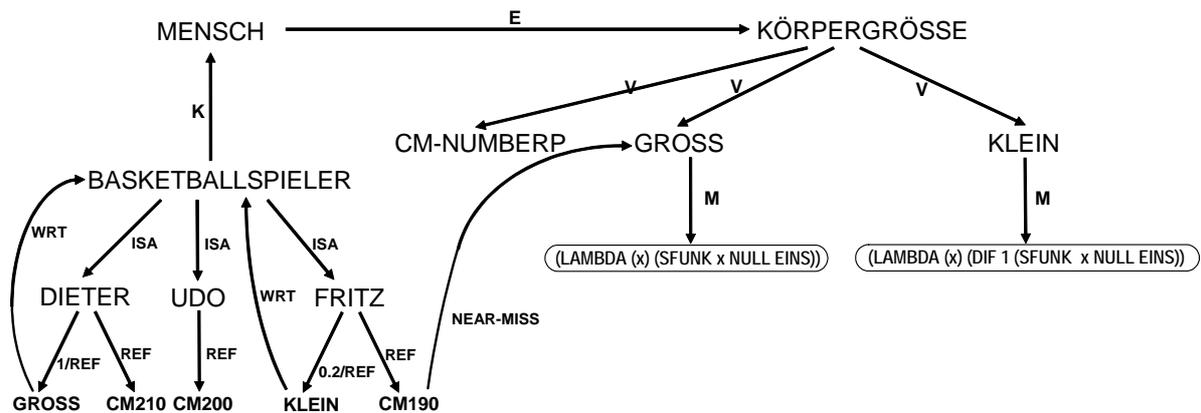
4.4.3.3. Die Repräsentation von Wissen über die Referenzmengen relativer Adjektive

Wie wir in Abschnitt 3.2.1. festgestellt haben, ist die extensionale Bedeutung vager, relativer Adjektive von der Referenzmenge, auf die sie sich beziehen, abhängig. In den SN muß also spezifiziert werden, was z.B. die extensionale Bedeutung von *groß* in bezug auf Basketballspieler, Elephanten, Ameisen usw. ist.

Wir müssen die in Fig. 35 dargestellten Funktionen also parametrisieren, um abhängig von der

⁵⁰ WRT steht für 'with respect to'; (vgl. auch BORGIDA75, S.18, RULIFSONetal.72, S.81)

durch den Kontext festgelegten Referenzmenge eine der jeweiligen Bedeutung entsprechende Funktion wählen zu können.



FIGUR 38

In dem in Fig. 38 dargestellten Teilnetz sind die Funktionen, welche die zu *groß* und *klein* gehörigen F-Mengen spezifizieren, parametrisiert. Die Formalparameter NULL und EINS dieser Funktionen werden während eines Inferenzprozesses durch Aktualparameter ersetzt, die aus dem Wissen über die jeweils angesprochene Referenzmenge extrahiert werden. Soll z.B. die Frage (1) *Ist Udo groß für einen Basketballspieler?* beantwortet werden, so wird zunächst nach einer Größenangabe gesucht, die für einen großen Basketballspieler typisch ist. Dieser Wert wird ermittelt, indem die REF-Kanten aller zu BASKETBALLSPIELER gehörenden Tokens nach einer mit 1/REF bewerteten und zu dem Knoten GROSS führenden Kante abgesucht werden: Im Beispiel wird festgestellt, daß DIETER mit einer Größe von 210 cm ein typisch großer Basketballspieler ist. Der Formalparameter EINS in der Zugehörigkeitsfunktion für *groß* wird durch 210 ersetzt.

In Anlehnung an das von WINSTON75 für die Szenenanalyse entwickelte Konzept der 'near-miss examples' wird in dem in Fig. 38 dargestellten ESN eine Kante $X \xrightarrow{\text{NEAR-MISS}} Y$ benutzt. X wird dabei als ein Wert interpretiert, der knapp unterhalb des Bereichs liegt, von dem ab das Prädikat Y gilt, d.h. X gehört nicht mehr zur Trägermenge von $M(Y)$. Der Formalparameter NULL wird durch den Wert X ersetzt. Für das Beispiel ergibt sich dann $M_{(\text{GROSS: BASKETBALLSPIELER})} = \text{SFUNK}(x, 190, 210)$.

Die mit dem Knoten UDO verbundene Größenangabe wird nun in diese Funktion eingesetzt. Wir erhalten dann den Wert (GROSS 0.5) und können die Frage (1) positiv beantworten.

Das skizzierte Verfahren beruht darauf, daß für jede Referenzmenge nach einem typischen und einem negativ abgrenzenden Beispiel in dem ESN gesucht wird und daß mit Hilfe dieser Angaben die jeweils aktualisierte extensionale Bedeutung des relativen Adjektivs approximiert wird.

Die beiden in den Inferenzprozeß eingehenden ausgezeichneten Elemente der Referenzmenge können als kognitive Orientierungspunkte aufgefaßt werden, auf deren allgemeine Bedeutung für die Verarbeitung vager Informationen bereits in Abschnitt 2.2. hingewiesen wurde. Falls solche Orientierungspunkte für eine bestimmte Referenzmenge nicht vorhanden sind (z.B. wenn jemand keine Vorstellung davon hat, wie schnell ein *schnelles Rennpferd* ist), so ist für diese Referenzmenge die extensionale Bedeutung des entsprechenden vagen Ausdrucks nicht spezifiziert. Es wird dann mit der intensionalen Bedeutung (vgl. 4.4.4.) des Ausdrucks weitergearbeitet.

Schon diese vergleichsweise einfachen Beispiele zeigen die Grenzen der Ausdrucksmöglichkeiten von den hier verwendeten SN: die RK werden unübersichtlich und die Interpretation der SN verliert bei der Darstellung komplexer Originale ihre ursprüngliche Direktheit.

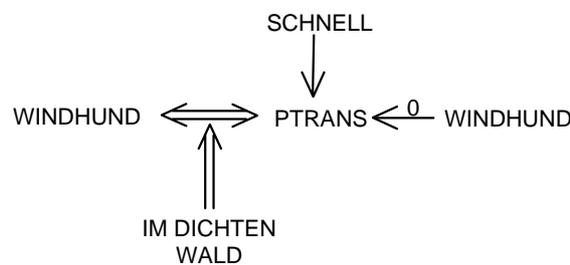
Eine übersichtlichere Darstellung der extensionalen Bedeutung vager Ausdrücke ermöglicht die RS 'Frames', die wir in Abschnitt 4.6. diskutieren werden.

4.4.4. DIE DARSTELLUNG DER INTENSIONALEN BEDEUTUNG VAGER KONZEPTE IN VERSCHIEDENEN FORMEN VON SN

Nachdem in Abschnitt 4.4.3. einige Möglichkeiten zur Repräsentation der extensionalen Bedeutung vager Konzepte diskutiert worden sind, sollen in diesem Abschnitt drei Vorschläge zur Darstellung der intensionalen Bedeutung vager Konzepte in SN erläutert werden.

Die Arbeiten von CERCONE/SCHUBERT⁷⁵ und HEMPHILL⁷⁵ gehen von einer Kritik an RK aus, die Schank im Rahmen seiner Theorie der konzeptuellen Dependenz (engl. conceptual dependency, Abk.: CD, vgl. SCHANK⁷⁵) vorgeschlagen hat. In diesen RK werden u.a. vage Adverben wie *schnell* unanalysiert als Knoten von SN benutzt und die relationale Struktur der Bedeutung vager Adjektive wird nicht berücksichtigt, obwohl Schank für sich in Anspruch nimmt, in diesen RK eine präzise, eindeutige und sprachunabhängige Repräsentation der NS-Formulierungen vorgenommen zu haben.

BEISPIEL: Fig. 39 ist eine Repräsentation des Satzes (1) *Der Windhund rennt schnell im dichten Wald*, die Schanks CD-Graphen benutzt. Dabei bedeutet $\langle \text{Objekt1} \rangle \leftrightarrow \text{PTRANS} \xleftarrow{0} \langle \text{Objekt2} \rangle$, daß $\langle \text{Objekt1} \rangle$ die lokale Position von $\langle \text{Objekt2} \rangle$ ändert.⁵¹



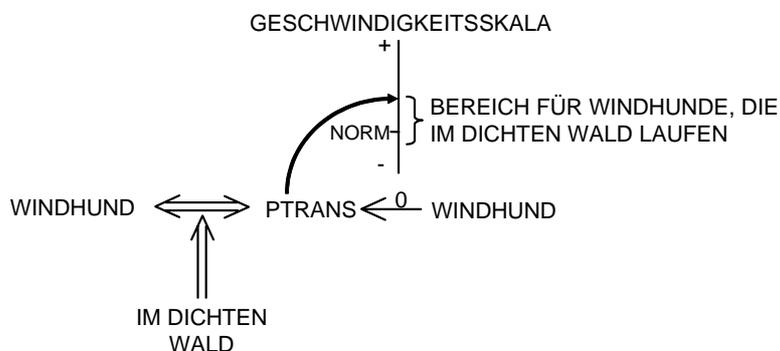
FIGUR 39

Die elementare Aktion PTRANS (engl. physical transport) wird durch das Adverb *schnell* und die gesamte Proposition⁵² durch die Ortsangabe *im dichten Wald* modifiziert.

4.4.4.1. Relative Skalen in Schanks Konzeptuelle-Dependenz-Graphen

Hemphill erweiterte Schanks CD-Graphen u.a. um Möglichkeiten zur expliziten Darstellung der intensionalen Bedeutung vager Adjektive und Adverbien (vgl. HEMPHILL⁷⁵, S.28ff, S.71ff).

BEISPIEL: Fig. 40 ist eine Darstellung von Satz (1) in der von Hemphill vorgeschlagenen RS.



FIGUR 40

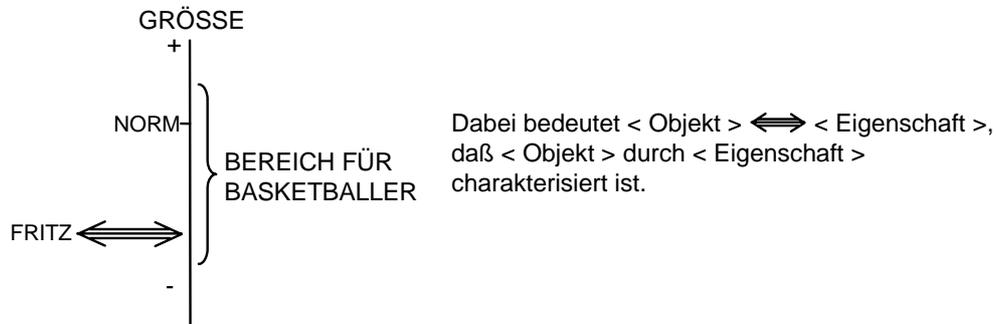
⁵¹ Diese Erklärung ist für unsere Zwecke stark vereinfacht, vgl. SCHANK⁷⁵.

⁵² In Schanks Terminologie 'Konzeptualisierung' genannt.

Auf einer Geschwindigkeitsskala wird ein Bereich möglicher Geschwindigkeiten von Windhunden, die im dichten Wald laufen, festgelegt. Innerhalb dieses Bereichs ist der Wert einer Norm markiert. *Schnell* wird nun als Zeiger auf einen Skalenwert dargestellt, der über dem auf den abgegrenzten Bereich bezogenen Wert der Norm liegt.

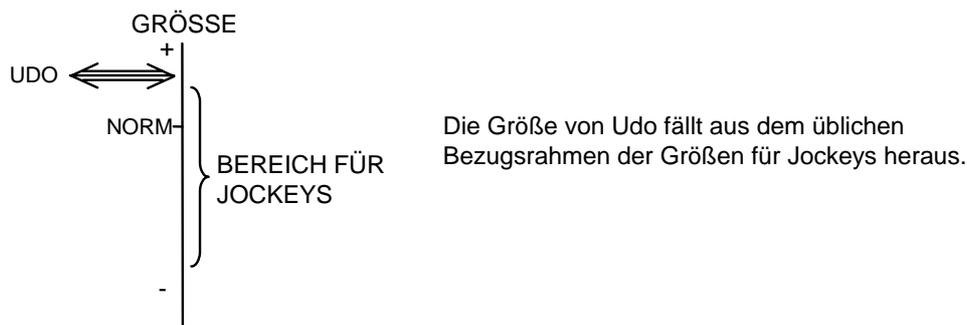
Dasselbe Prinzip gilt auch für die in Fig. 41 und 42 dargestellten Sätze, in denen relative Adjektive vorkommen.

Original: (2) *Fritz ist klein für einen Basketballspieler.*



FIGUR 41

Original: (3) *Udo ist zu groß für einen Jockey.*



FIGUR 42

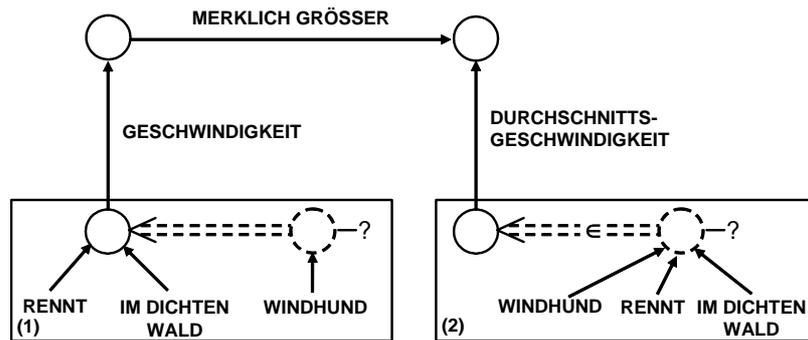
Hemphill erweitert also die von Schank (vgl. SCHANK75, S.47) eingeführten absoluten Skalen (z.B. 'emotionaler Zustand(Marie) =5' bedeutet *Marie ist glücklich*) zu relativen Skalen, auf denen Normwerte fixiert sind.

Ähnliche Vorschläge sind bereits in SIMMONS72 (S.77ff) und WETTLER74 (S.29ff) zu finden. Simmons schlägt vor, relative Adjektive in SN als Abweichungen von Mittelwerten auf bestimmten Skalen zu repräsentieren, und Wettler führt für jede Referenzmenge Angaben über einen Mittelwert und über die mittlere Abweichung von diesem Mittelwert (z.B. in Form einer Standardabweichung) in CD-Graphen ein.

4.4.4.2. Die Repräsentation von Referenzmengen in Schuberts Formalismus

CERCONE/SCHUBERT75 machen einen Vorschlag zur Repräsentation vager Adverbien, in der von SCHUBERT75 entwickelten Form von SN, der von den linguistisch-logischen Analysen vager Adverbiale von BARTSCH/VENNEMANN72 (vgl. 4.1.4.) ausgeht. Fig. 43 ist eine RK für den Satz *Der Windhund läuft schnell im dichten Wald.* Die durch Rechtecke gekennzeichneten Teile des SN stellen folgende Originale dar (vgl. CERCONE/SCHUBERT75, S. 89):

- | | |
|---|-----|
| <i>Der Windhund, der im dichten Wald rennt</i> | [1] |
| <i>Alle im dichten Wald rennenden Windhunde</i> | [2] |



FIGUR 43

Das in Rechteck (1) dargestellte Teilnetz repräsentiert die Referenzmenge für *schnell*⁵³. Die Kanten GESCHWINDIGKEIT und DURCHSCHNITTSGESCHWINDIGKEIT werden als Funktionen und MERKLICH GRÖßER als F-Prädikat interpretiert. Das in Fig. 43 dargestellte Wissen kann als *die Geschwindigkeit des im dichten Wald rennenden Windhundes ist merklich größer als die Durchschnittsgeschwindigkeit von im dichten Wald rennenden Windhunden* verbalisiert werden.

Die Parallelität zwischen den in Abschnitt 4.1.4. als Darstellung der intensionalen Bedeutung vager Adjektive im Prädikatenkalkül diskutierten Beispiele und den hier angegebenen RK ist offensichtlich.

4.4.4.3. Die Modifikation vager Konzepte in gerichteten rekursiven Bewertungsknoten-Hypergraphen

RK für vages Wissen in Form von einfachen SN werden wegen der beschränkten Syntax dieser RS in einigen Fällen sehr aufwendig, da viele Hilfskonstruktionen zur Darstellung der zahlreichen Modifikationen von Kanten und Knotenbewertungen eingeführt werden müssen.

Dabei geht ein Vorteil der SN gegenüber anderen RS, nämlich die einfache, kombinierte Darstellung von Faktenwissen und assoziativen Verbindungen zwischen einzelnen Wissensseinheiten (vgl. WOODS75, S. 44) verloren, da die große Zahl von Hilfskanten und -knoten eine Modellierung assoziativer Verbindungen sehr aufwendig werden läßt.

Zur Zeit versucht man, durch verschiedene Erweiterungen der Syntax von einfachen SN Repräsentationssprachen zu entwickeln, in denen auch komplexe Wissensstrukturen auf natürliche und einfache Weise dargestellt werden können (vgl. z.B. BOLEY76b, HENDRIX76).

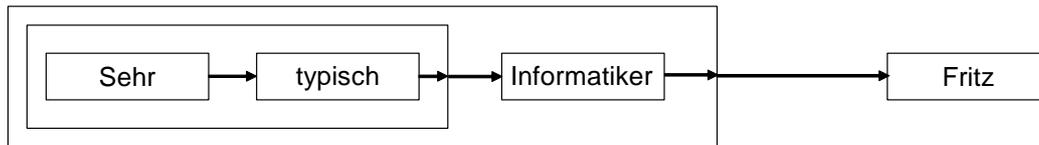
Am Beispiel einer dieser RS, den sog. gerichteten rekursiven Bewertungsknoten-Hypergraphen (engl. directed recursive labelnode hypergraphs, Abk.: DRLHs, vgl. BOLEY76b), die sich gegenüber ähnlichen Ansätzen durch ihre Allgemeinheit und die vollständige Formalisierung ihrer Syntax auszeichnen, wird im folgenden gezeigt, wie solche komplexen SN bei der Darstellung von vagem Wissen eingesetzt werden können.

Neben dem Konzept der gerichteten Hyperkanten, das eine direkte Darstellung n-stelliger Relationen ermöglicht, eignet sich vor allem das Konzept der sog. komplexen Bewertungsknoten zur Repräsentation von vagem Wissen, da eine einfache Repräsentation der Modifikation von Knoten- und Kantenbezeichnungen durch dieses Konzept ermöglicht wird. Durch komplexe Knoten, wie sie für rekursive Graphen charakteristisch sind, können in DRLHs beliebig tief verschachtelte Modifikationen der Knoten eines SN (z.B. in Form von Hecken) beschrieben werden.

Da in DRLHs die für einfache SN übliche Trennung von Knoten und Kantenbezeichnungen durch die Einführung von Bewertungsknoten, welche die Funktion von Knoten- und/oder Kantenbewertungen übernehmen können, aufgehoben wurde, kann auch die Modifikation semantischer Relationen durch komplexe Bewertungsknoten dargestellt werden.

⁵³ Die spezielle Form der Teilnetze ist in diesem Zusammenhang nicht relevant (vgl. dazu SCHUBERT75).

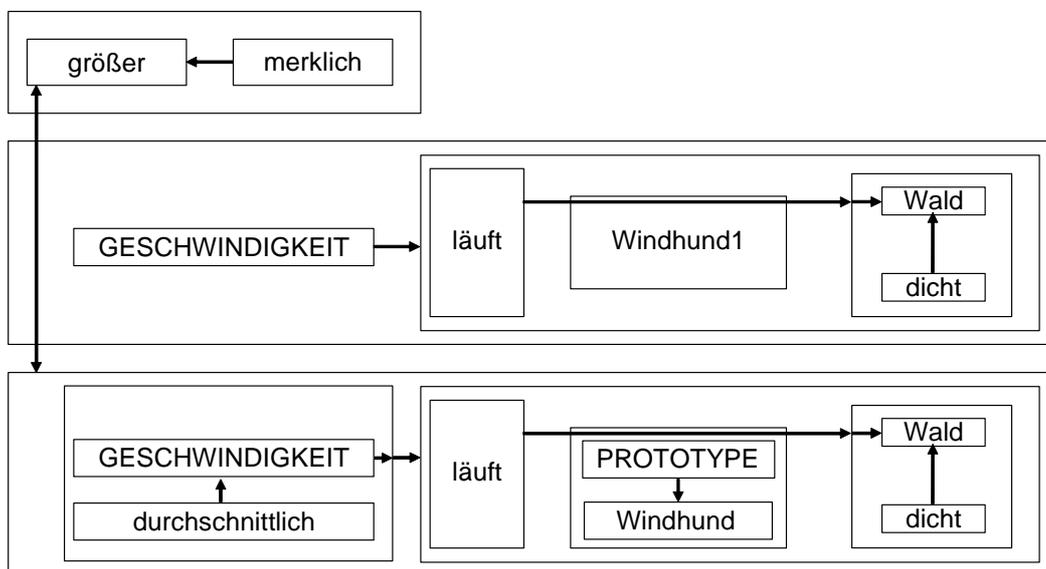
BEISPIEL: Die in Fig. 44 dargestellte RK ist die Diagramm-Form eines DRL_H, durch die der Satz *Fritz ist ein sehr typischer Informatiker* dargestellt werden kann.



FIGUR 44

Das vage Prädikat 'sehr-typischer-Informatiker(x)' wird als komplexer Bewertungsknoten repräsentiert, in dem ein weiterer komplexer Knoten für den Teilausdruck *sehr typisch* eingebettet ist. Die Knoten *typisch* und *Informatiker* fungieren dabei als Kontaktknoten. In der Wissensbasis eines NSS wird diese RK in einer der linearisierten Darstellungsformen für DRL_Hs (z.B. als Relationenlistenmengen, vgl. BOLEY76b) abgespeichert.

Fig. 45 ist eine weitere RK für die intensionale Bedeutung des Satzes *Der Windhund rennt schnell durch den dichten Wald* (vgl. auch 4.4.4.1, 4.4.4.2).



FIGUR 45

Alle Funktionssymbole sind in Fig.45 groß geschrieben. Die Funktion PROTOTYPE liefert als Wert ein Element einer vagen sprachlichen Kategorie, die als F-Menge definiert ist. Die RK in Fig. 45 lehnt sich an die Darstellung von CERONE/SCHUBERT75 an, weist aber an einigen Stellen einen höheren Detaillierungsgrad auf. Das vage, zweistellige Prädikat 'merklich größer' wird als komplexer Bewertungsknoten mit Kontaktknoten 'größer' dargestellt und hat als Argumente Geschwindigkeitsangaben, die ihrerseits durch zwei komplexe Knoten repräsentiert sind.

Doch auch diese RK ist noch unvollständig. Genaugenommen müßte das relative Adjektiv *dicht* noch auf eine Referenzmenge bezogen werden. Da diese Referenzmenge analog zu der für *schnell* angegebenen repräsentiert werden kann, wurde sie in Fig. 45 aus Gründen der Übersichtlichkeit nicht berücksichtigt.

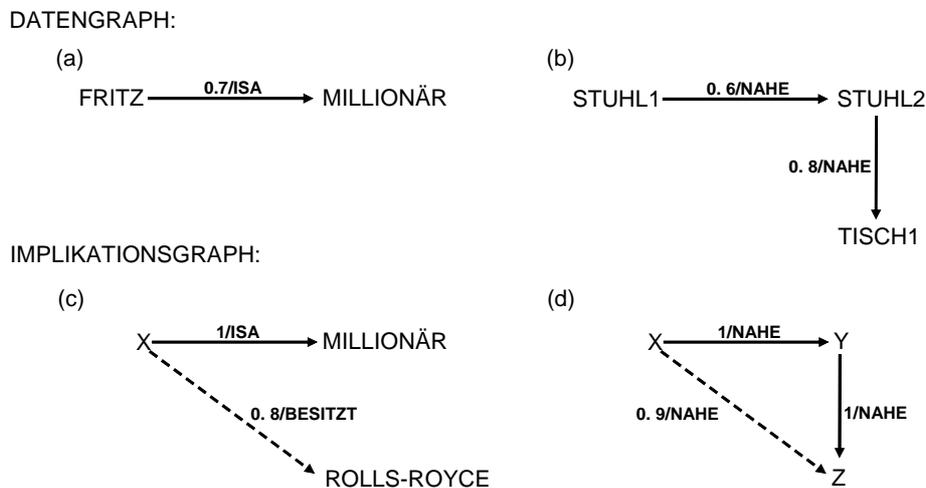
Die große Allgemeinheit von DRL_Hs ermöglicht auch eine einfache Darstellung der extensionalen Semantik vager Ausdrücke, für die in Abschnitt 4.4.3. erweiterte SN verwendet wurden. Die in erweiterten SN auftretenden LISP-Prozeduren können direkt als komplexe Knoten eines DRL_H interpretiert werden (vgl. BOLEY76b).

4.4.5. DIE DARSTELLUNG VAGER INFERENZREGELN IN SN

4.4.5.1. Verallgemeinerte Implikationsgraphen

Im folgenden wird gezeigt, wie einfache vage Inferenzregeln in den von WAHLSTER/v.HAHN76 vorgeschlagenen SN durch eine Verallgemeinerung von Konrads Graphenkalkül (vgl. KONRAD76) für F-Graphen dargestellt werden können.

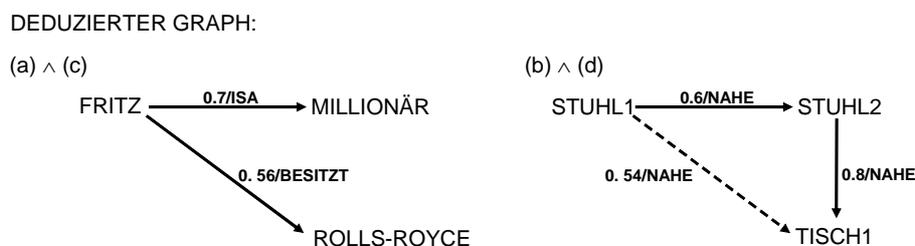
Vage Inferenzregeln der Form $A_1 \wedge A_2 \dots A_n \xrightarrow{\alpha} B$ (α als F-Wert) werden in SN als Implikationsgraphen dargestellt⁵⁴. Implikationsgraphen sind Subgraphen von F-Graphen, die Individuenvariablen als Knoten haben können und aus einem Prämissengraphen und einem Konklusionsgraphen bestehen. Der Prämissengraph kann aus mehreren Prämissenkanten, die stets mit dem F-Wert 1 markiert sind, bestehen (vgl. Fig. 46). Der Konklusionsgraph besteht nur aus genau einer ausgezeichneten Konklusionskante (in Fig. 46 gestrichelt eingezeichnet), die zusätzlich mit beliebigen Werten aus dem reellen Intervall $[0,1]$ bewertet sein kann. Alle Subgraphen des SN, die keine Implikationsgraphen sind, nennen wir Datengraphen (vgl. Fig. 46).



FIGUR 46

Der Inferenzmechanismus besteht nun darin, daß zunächst der Implikationsgraph auf den Datengraphen gelegt wird, wobei der Prämissengraph als Schablone dient. Paßt der Prämissengraph auf den Datengraphen, so wird die Konklusionskante im Datengraphen ergänzt. Dabei wird der F-Wert der deduzierten Kante als Produkt⁵⁵ aus dem F-Wert der Konklusionskante und dem Minimum der F-Werte derjenigen Kanten des Datengraphen berechnet, die mit den Prämissenkanten des Implikationsgraphen 'gematcht' wurden.

In Fig. 47 sind die Ergebnisse des auf die Beispiele aus Fig. 46 angewendeten Inferenzprozesses dargestellt.



FIGUR 47

⁵⁴ Man beachte den Unterschied zu den in Abschnitt 4.2.3. definierten Inferenznetzen.

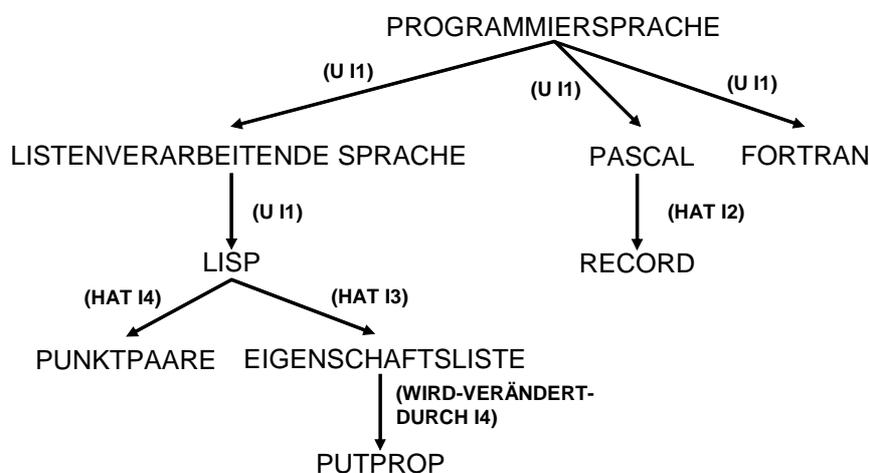
⁵⁵ In diesem Beispiel wird der multiplikative Abtrennungsoperator benutzt (vgl. 4.1.3.)

4.4.5.2. Die 'Davon-wüßte-ich'-Inferenz als Beispiel einer komplexen Schlußregel

Komplizierte Inferenzregeln über vagen, unvollständigen und als SN codierten Wissensbasen werden meist als komplexe Suchprozeduren in einer KI-Programmiersprache (vgl. 4.3.) repräsentiert.

Die 'Davon-wüßte-ich'-Inferenz (engl. lack-of-knowledge inference, vgl. CARBONELL/COLLINS73) wird im folgenden als typisches Beispiel einer vagen Inferenzregel diskutiert.

In einer Wissensbasis sei das in Fig. 48 dargestellte, stark vereinfachte Teilnetz abgespeichert. Da in diesem Teilnetz keine Markierungen für vollständige Mengen (vgl. 4.4.2.4.) vorhanden sind, 'weiß' das System, daß es z.B. noch weitere Programmiersprachen und noch andere Datenstrukturen in PASCAL gibt, die ihm im einzelnen nicht bekannt sind. Um eine Frage wie *Hat LISP Records ?* zu beantworten, wird zunächst von dem Knoten LISP ausgehend nach einer Kante mit der Bezeichnung HAT gesucht, die zu einem Knoten mit der Bezeichnung RECORD führt. Ein solcher Knoten wird nicht gefunden. Da das System 'weiß', daß es nur unvollständiges Wissen hat, bricht es hier nicht ab und antwortet nicht Nein, sondern wendet die 'Davon-wüßte-ich'-Inferenz an.



FIGUR 48

Die Grundidee dieser Inferenzregel ist folgende: Wenn eine erfragte Eigenschaft X bei einem Objekt Y nicht gespeichert ist, dann überprüfe, ob bei einem Y ähnlichen Objekt Z die Eigenschaft X gespeichert ist. Wenn mehr Wissen über das Objekt Y gespeichert ist als über das Objekt Z und die Eigenschaft X für Z gilt, dann trifft X auf Y wahrscheinlich nicht zu. Dieses Ergebnis wird damit begründet, daß über Y mehr gespeichert ist als über Z und daß daher, falls X auch auf Y zutreffen würde, dies bestimmt bekannt wäre.

Der auf dieser Strategie beruhende Suchalgorithmus arbeitet für das Beispiel in Fig. 48 folgendermaßen: über die Oberbegriffs- und Unterbegriffshierarchie wird nach einem LISP ähnlichen Objekt gesucht. Dabei werden PASCAL und FORTRAN gefunden, die wie LISP Unterbegriffe von PROGRAMMIERSPRACHEN sind. Es wird festgestellt, ob von diesen Objekten eine HAT-Kante nach RECORD führt. Dies trifft auf PASCAL zu. An Hand der Werte der I-Markierungen (vgl. 4.4.2.3.) wird der Detaillierungsgrad des Wissens über PASCAL mit dem über LISP verglichen. Da über PASCAL weniger Wissen gespeichert ist als über LISP, geht das System davon aus, daß es 'davon wüßte', wenn es in LISP Records gäbe. Da diese Inferenz in höchstem Maße unsicher und vage ist, wird das negative Ergebnis der Suche durch die Hecke *wahrscheinlich* modifiziert und die Antwort *Wahrscheinlich nicht* generiert.

Dieses Beispiel macht deutlich, daß dem Metawissen eine entscheidende Funktion in vagen Inferenzprozessen zukommt.

4.5 EIN VERGLEICH VON VIER REPRÄSENTATIONSSPRACHEN AM BEISPIEL EINER VAGEN INFERENZ

Am Beispiel einer vagen Inferenz werden in Fig. 49 folgende in den vorangegangenen Abschnitten behandelten Repräsentationssprachen gegenübergestellt:

- Fuzzy Prädikatenlogik (vgl. 4.1.)
- Fuzzy Produktionensysteme (vgl. 4.2.)
- KI-Programmiersprachen: FUZZY (vgl. 4.3.)
- Fuzzy semantische Netze (vgl. 4.4.)

Folgendes in NS formulierte vage Wissen soll in der Wissensbasis eines NSS formal repräsentiert werden:

Regelwissen: *Wenn ein Programm sehr groß ist, ist es meist fehlerhaft.*

Faktenwissen: *HAM-RPM ist ein ziemlich großes Programm.*

Auf die Frage *Ist HAM-RPM fehlerhaft ?* soll das System antworten
Ich bin mir ziemlich sicher, daß HAM-RPM fehlerhaft ist.

Wir vernachlässigen in Fig. 49 Algorithmen zur Analyse und Generierung von NS, welche auch die Überführung von linguistischen Hecken in F-Wahrheitswerte vornehmen (vgl. 3.2.3.).

In der Spalte 'Fuzzy Prädikatenlogik' in Fig. 49 steht SG für *sehr groß* und F für *fehlerhaft*. Die Individuenkonstante 'a' steht für *HAM-RPM*. Es wird F-Resolution mit multiplikativem Abtrennungsoperator benutzt (vgl. 4.1.3.).

In der Spalte 'Fuzzy Produktionensysteme' werden die Produktionen von links nach rechts ausgewertet (Vorwärtsverkettung, vgl. 4.2.1.).

In der Spalte 'Fuzzy semantisches Netz' wird mit verallgemeinerten Implikationsgraphen gearbeitet (vgl. 4.4.5.1.).

Die erste Zeile in Fig. 49 enthält die jeweilige Repräsentationskonstruktion für das Regelwissen, die zweite für das Faktenwissen, die dritte für die Frage und die vierte Zeile enthält die Antwort.

FUZZY PRÄDIKATENLOGIK	FUZZY PRODUKTIONENSYSTEME
1 C ₁ : ¬ SG(x) ∨ F(x) T(C ₁)=0.9	((SEHR-GROSS x) z) → ADD-TO-STM ((FEHLERHAFT x) z • 0.9) (GOAL (FEHLERHAFT x))((FEHLERHAFT x) z) → ADD-TO-STM (GOAL-COMPLETED((FEHLERHAFT x) z))
2 C ₂ : ¬ SG(a) T(C ₂)=0.7	DB: ((SEHR-GROSS HAM-RPM) 0.7)
3 C ₃ : ¬ F(a) T(C ₃)=1	DB: (GOAL (FEHLERHAFT HAM-RPM))
4 T(F(a))=0.63	(GOAL -COMPLETED((FEHLERHAFT HAM-RPM) 0.63))
KI-PROGRAMMIERSPRACHE: FUZZY	FUZZY SEMANTISCHES NETZ
1 (ADD DEDUCE: (PROC DEMON:*ABTRENN THRESH: 0.9 (FEHLERHAFT ?PROGRAMM) (GOAL (SEHR-GROSS IPROGRAMM))))	
2 (ADD (SEHR-GROSS HAM-RPM) 0.7)	HAM-RPM $\xrightarrow{0.7/HAT-EIGENSCHAFT}$ SEHR-GROSS
3 (GOAL (FEHLERHAFT HAM-RPM))	
4 ((FEHLERHAFT HAM-RPM) 0.63)	HAM-RPM $\xrightarrow{0.63/HAT-EIGENSCHAFT}$ FEHLERHAFT

FIGUR 49

4.6. DIE REPRÄSENTATION VON VAGEM WISSEN IN DER REPRÄSENTATIONSSPRACHE 'FRAMES'

4.6.1. EINFÜHRENDE ÜBERSICHT

In Abschnitt 2.2. wurden psychologische Untersuchungen von Rosch diskutiert, die darauf hindeuteten, daß es ausgezeichnete Elemente von unscharfen sprachlichen Kategorien, sog. Prototypen, gibt, die typische Beispiele einer sprachlichen Kategorie sind und u.a. als Repräsentanten der unscharfen Kategorie in modell-gesteuerten, beispielorientierten Inferenzen (engl. reasoning by example) fungieren. In diesem Abschnitt wird nun gezeigt, wie die Repräsentationssprache 'Frames', die MINSKY75 ausgehend von der in der Psychologie erarbeiteten Prototypentheorie (vgl. 2.2.2.) entwickelt hat, zur Darstellung und Verarbeitung von vagem Wissen eingesetzt werden kann.

Wie bei der Untersuchung anderer Repräsentationssprachen in den vorangegangenen Abschnitten so werden auch hier einige Grundkenntnisse über die Repräsentationssprache 'Frames' beim Leser vorausgesetzt.⁵⁶

Ein 'Frame', der zunächst vereinfachend als eine Ansammlung von auf ein bestimmtes Konzept bezogenen Fakten und Prozeduren charakterisiert werden kann, soll im Gegensatz zu den bisher bekannten RS eine objekt-zentrierte Darstellung (vgl. 1.1.3.) und eine Verarbeitung von Wissen in größeren Einheiten ermöglichen. Dabei werden Konzepte aus deklarativen und prozeduralen RS (besonders SN und KI-Programmiersprachen) zu einem einheitlichen Formalismus verschmolzen. Die von MINSKY75 zunächst nur grob skizzierte RS 'Frames' ist inzwischen innerhalb der KI in verschiedenen Arbeiten konkretisiert, weiterentwickelt und in bestehende RS integriert worden.⁵⁷ Die Vielzahl von Adaptionen hat eine Reihe unterschiedlicher Bezeichnungen für Frame-ähnliche Datenstrukturen mit sich gebracht: 'Script' (SCHANKetal.75), 'Schema' (BOBROW/NORMAN75, NORMAN/BOBROW76), 'Unit' (BOBROW/WINOGRAD76). Diese ursprünglich für die Anwendung im Bereich der Szenenanalyse konzipierte RS 'Frames' wird heute in verstärktem Maße auch bei der Konstruktion von NSS berücksichtigt.⁵⁸ Auch in der Linguistik wird das Minskysche Frame-Konzept zur Strukturierung der Lexikonkomponente von Sprachverstehensmodellen eingesetzt (vgl. z.B. FILLMORE75, PETÖFI 76).

4.6.2. DIE DARSTELLUNG VON VAGEM WISSEN IN KRL

4.6.2.1. Das Frame-Konzept von KRL und einige Erweiterungen

Im folgenden orientieren wir uns vorwiegend an einer Variante des Frame-Konzeptes, die der Repräsentationssprache KRL (Knowledge Representation Language, vgl. BOBROW/WINOGRAD76) zugrunde liegt⁵⁹.

Es werden diejenigen Sprachkonstruktionen von KRL, die für eine Repräsentation von vagem Wissen geeignet sind, dargestellt und einige Spracherweiterungen, die sich auf die Darstellung von vagem Wissen beziehen und im folgenden kursiv geschrieben werden, vorgeschlagen und diskutiert⁶⁰.

⁵⁶ vgl. MINSKY75, WINOGRAD74/75b, KUIPERS75, WINSTON77

⁵⁷ vgl. z.B. SCHANKetal.75, CHARNIAK75/77, LAUBSCH75ab, NORMAN/BOBROW76, BOBROWetal.76, BOBROW/WINOGRAD76, GÖRZ/LAGEMANN76

⁵⁸ z.B. in den NSS: GUS (BOBROWetal.76), SAM (SCHANKetal.75), Ms. MALAPROP (CHARNIAK77)

⁵⁹ Die erste Version dieser Sprache wurde unter der Bezeichnung KRL-0 in LISP implementiert.

⁶⁰ Auf eine detaillierte Darstellung der Syntax von KRL wird hier verzichtet (vgl. BOBROW/WINOGRAD76).

DEFINITION: Ein Frame ist eine zusammengesetzte Datenstruktur, die aus einer Menge von innerhalb eines Frames eindeutig bezeichneten Attributen (engl.: slots) besteht.⁶¹ Ein Frame wird mit einem eindeutigen Namen bezeichnet⁶². Jeder Frame enthält ein ausgezeichnetes Attribut, das als SELF bezeichnet wird und an das in Form eines Deskriptors eine Beschreibung der Entität, die durch den Frame repräsentiert wird, gebunden werden kann.

BEISPIEL: Das Konzept 'Person' sei in einem NSS durch folgenden Frame dargestellt:

```
[Person UNIT BASIC
  < SELF > ;keine weitere Beschreibung vorgesehen
  < vorname (A String) >
  < nachname (A String) >
  < wohnsitz { (A Stadt) Hamburg; 0.6/DEFAULT } >
  < alter (A Integer) >
  < äußeres (A Aussehen) >
  < privates (A Privatbereich) > ]
```

Jeder Deskriptor⁶³ ist eine unabhängige Charakterisierung der durch den Frame repräsentierten Entität. Nach dem Prinzip der objekt-bezogenen Prozedurdefinition (engl.: procedural attachment) können neben den Deskriptoren auch Definitionen von Prozeduren an ein Attribut gebunden werden. Außerdem können Deskriptoren durch Angabe eines Meta-Merkmals (engl.: feature) klassifiziert werden (Stichwort: Metawissen). Beispielsweise kennzeichnet das Meta-Merkmal DEFAULT Deskriptoren, die als Standardspezifikationen interpretiert werden (z.B. der Deskriptor 'Hamburg' in dem Frame für 'Person'). Falls für eine Inkarnation eines Frames ein spezieller Wert zu einem durch DEFAULT modifizierten Attribut angegeben wird, so wird die Standardspezifikation⁶⁴ nicht beachtet. In allen anderen Fällen wird nach dem Frame-Konzept die Standardspezifikation als Wert des Attributs aufgefaßt.

Standardspezifikationen werden in sog. Prototype-Frames verwendet, die als typische Beispiele der durch den Frame dargestellten Klasse von Entitäten interpretiert werden. Die Angabe von typischen, einer Erkennungsnorm entsprechenden Merkmalen der Objekte einer sprachlichen Kategorie ist für NSS, die in einer Umgebung arbeiten, in der oft nur vage und unvollständige Informationen zur Verfügung stehen, besonders wichtig.

DEFINITION: Wird durch einen Frame X ein Prototyp einer bestimmten sprachlichen Kategorie und durch einen Frame Y ein bestimmtes Element dieser sprachlichen Kategorie beschrieben, so wird Y als ein Objektframe des Prototyp-Frames X bezeichnet.⁶⁵

In Prototype-Frames geben die an Attribute gebundenen Deskriptoren Bedingungen an, die von den Werten der entsprechenden Attribute eines Objektframes⁶⁶ erfüllt werden müssen. In der Deklaration eines Prototype-Frames werden Attributbezeichner für alle im Zusammenhang mit dem durch den Frame beschriebenen Konzept möglicherweise relevanten Eigenschaften eingeführt. In einem Objektframe brauchen nicht alle für den dazugehörigen Prototype-Frame definierten Attribute mit Werten belegt zu sein (Stichwort: partielles Wissen).

In einem Frame kann der Wert eines Attributs auch aus einem Verweis auf einen anderen Frame

⁶¹ Frame-Definitionen werden durch UNIT gekennzeichnet und durch [...] eingeklammert. Bezeichner für Attribute werden klein geschrieben.

⁶² Bezeichner für Frames beginnen mit Großbuchstaben.

⁶³ Ein Paar aus einem Attributbezeichner (vgl. auch den 'field identifier' in PASCAL-Records) und dem an ihn angeordneten Deskriptor wird durch <.> eingeklammert. Mehrere Deskriptoren eines Attributs werden durch {...} eingeklammert.

⁶⁴ bei Schank als 'blank events' bezeichnet

⁶⁵ Prototype-Frames werden hier nicht als Elemente der Kategorie aufgefaßt (vgl. dagegen 2.2.2.)

⁶⁶ von Minsky als 'terminals' bezeichnet

bestehen (vgl. z.B. 'Aussehen' und 'Privatbereich' in dem oben angegebenen Frame für 'Person'). In der Zeile < wohnsitz { (A Stadt) Hamburg; 0.6/DEFAULT }> des oben dargestellten Frames wird die Standardspezifikation 'Hamburg' durch den F-Wahrheitswert 0.6 modifiziert.⁶⁷ Diese Spracherweiterung kommt der Forderung von Kuipers entgegen, daß Standardspezifikationen in unterschiedlicher Stärke mit den entsprechenden Attributen eines Frames verbunden sein sollten (vgl. KUIPERS75, S.155ff).

Wurde z.B. ein bestimmter Personenwagen in einem Text nur vage beschrieben oder von einem Programm zur automatischen Szenenanalyse nur ungefähr identifiziert, so ist z.B. in dem entsprechenden Objektframe für den Personenwagen kein Wert für das Attribut 'Anzahl-der-Räder' vorhanden. Wird nach der Anzahl der Räder des Personenwagens gefragt oder wird diese in einer Inferenz benötigt, so wird aus dem Prototype-Frame für 'Auto' der Standardwert 4 entnommen. In diesem Fall ist der Standardwert z.B. durch 1/DEFAULT modifiziert, während der Standardwert 4 für das Attribut 'Anzahl-der-Türen' nur mit 0.8/DEFAULT gebunden ist.

Dieser Ansatz erlaubt also, daß die Erwartungs- und Erfahrungswerte für unscharfe sprachliche Kategorien graduell unterschieden werden. Die verschiedenen Bindungsstärken für Standardwerte werden von NSS in der sprachlichen Oberflächenstruktur durch linguistische Hecken ausgedrückt.

Ein wichtiges Konzept der RS 'Frames' besteht darin, daß ein bestimmter Frame Objektframe von mehreren Prototype-Frames sein kann. Ein Objekt kann also hinsichtlich verschiedener Sichtweisen klassifiziert werden.

Durch Deskriptoren der Form (A < Prototype-Frame > WITH < Attribut₁ > = < Wert₁ >... < Attribut_n > = < Wert_n >) kann ein Objektframe in bezug auf verschiedene Prototype-Frames klassifiziert werden. Solche Deskriptoren werden in KRL als 'perspectives' bezeichnet.

BEISPIEL: Die Kontextabhängigkeit der Bedeutung vager, relativer Adjektive kann in KRL durch die Angabe verschiedener, den unterschiedlichen Kontexten entsprechender 'perspectives' in einem Objektframe berücksichtigt werden. Die Prototype-Frames werden dabei als Referenzmenge interpretiert. (vgl.

```
[ Fritz UNIT INDIVIDUAL
  < SELF { (A Student WITH körpergröße = recht/groß ...)
           (A Basketballspieler WITH körpergröße = klein ....) }>..]
```

Jeder Frame gehört zu einer von sieben Klassen, die mit den reservierten Wörtern BASIC, ABSTRACT, SPECIALIZATION, INDIVIDUAL, MANIFESTATION, RELATION und PROPOSITION bezeichnet werden.

Die Klassen BASIC, ABSTRACT und SPECIALIZATION entsprechen genau den drei in Abschnitt 2.2.3. diskutierten Arten von sprachlichen Kategorien (vgl. Fig. 50).

ART DER SPRACHLICHEN KATEGORIE	RESERVIERTE WÖRTER IN KRL	BEISPIEL
Übergeordnete Kategorie	ABSTRACT	[Einrichtungsgegenstand UNIT ABSTRACT...]
Grundkategorie	BASIC	[Stuhl UNIT BASIC....]
Untergeordnete Kategorie	SPECIALIZATION	[Küchenstuhl UNIT SPECIALIZATION...]

FIGUR 50

In Frames vom Typ MANIFESTATION wird u.a. jeweils ein bestimmter Aspekt einer Entität, die durch einen Frame vom Typ INDIVIDUAL beschrieben wird, als eigenständige konzeptuelle Einheit dargestellt (z.B. kann das Individuum Fritz als Geschäftsmann (Manifestation₁) und als Familienvater (Manifestation₂) betrachtet werden).

⁶⁷ Anstelle von Zahlen können auch Hecken verwendet werden.

Durch die verschiedenen Klassenbezeichnungen werden unterschiedliche Arbeitsmodi des KRL-Interpreters gesteuert. So wird z.B. vom Interpreter angenommen, daß ein Objektframe vom Typ INDIVIDUAL nur zu einer Grundkategorie gehört, da sich Grundkategorien aus einer einfachen, disjunkten Einteilung der Elemente eines Weltmodells ergeben (vgl. 2.2.3.).

Spezielle Inferenzprozeduren werden in Frames vom Typ SPECIALIZATION gesucht, während allgemeine Inferenzprozeduren (evtl. bis hin zum automatischen Theorembeweiser, vgl. 4.1.) in Frames vom Typ ABSTRACT gesucht werden.

Beim Entwurf von KRL wurde auch die Repräsentation von unvollständigen Mengen berücksichtigt (vgl. 4.4.2.4.). Alle LISP-Atome, die in einer mit dem Schlüsselwort ITEMS beginnenden Liste enthalten sind, werden als Elemente einer unvollständigen Menge interpretiert.

BEISPIEL: [Fußballnationalmannschaft UNIT INDIVIDUAL
< SELF (A Fußballmannschaft WITH spieler = (ITEMS Vogts Maier)) >..]

Wenn wir zusätzlich vereinbaren, daß die Elemente einer mit dem reservierten Wort FSETOF beginnenden Liste als F-Menge interpretiert und manipuliert werden, so stehen uns neben klassischen und unvollständigen Mengen auch F-Mengen für die Darstellung von vagem Wissen zur Verfügung.

BEISPIEL: Mittelklassewagen =(FSETOF 0/Fiat500 0.7/Golf 1/Opel-Rekord 0.3/Mercedes
0/Rolls-Royce)

4.6.2.2. Meta-Merkmale und Hecken vom Typ 2

Als besonders günstig für die Repräsentation von vagem Wissen erweist sich die Möglichkeit von KRL, Deskriptoren durch Meta-Merkmale zu klassifizieren. Außer dem bereits erwähnten Meta-Merkmal DEFAULT sind in KRL noch CRITERIAL und PRIMARY vorgesehen (vgl. BOBROW/WINOGRAD76, S.17). CRITERIAL kennzeichnet einen Deskriptor, der als grundlegendes Merkmal des durch den Frame repräsentierten Konzepts aufgefaßt wird.⁶⁸ Das Meta-Merkmal PRIMARY markiert Deskriptoren, welche die primäre Sichtweise für die repräsentierten Konzepte angeben, indem sie auf Frames, die als Oberbegriffe interpretiert werden, verweisen.

Alle nicht durch Meta-Merkmale markierten Deskriptoren werden als sekundäre, für das Konzept charakteristische, aber weniger relevante Angaben interpretiert.

Der in KRL integrierte Pattern-Matcher berücksichtigt die Meta-Merkmale, indem er z.B. ein durch einen Frame X repräsentiertes Objekt bereits dann als Element einer durch einen Prototype-Frame Y dargestellten sprachlichen Kategorie erkennt, wenn alle Bedingungen, die durch CRITERIAL-Deskriptoren von Y spezifiziert sind, von X erfüllt werden.

Die Klassifikation von Deskriptoren durch Meta-Merkmale entspricht der Klassifikation von Bedeutungskomponenten eines vagen sprachlichen Terms bei LAKOFF75. In Abschnitt 3.2.2.2. haben wir gezeigt, wie Hecken vom Typ 2 als Operatoren über den Bedeutungskomponenten vager Konzepte dargestellt werden können. In KRL können solche Hecken als NS-Korrelate zu verschiedenen Arbeitsmodi des Pattern-Matchers von KRL aufgefaßt werden (vgl. Fig. 51).

Die in Fig. 51 dargestellte Tabelle ist folgendermaßen zu verstehen: Deskriptoren, die durch ein mit 'W' gekennzeichnetes Meta-Merkmal markiert sind, beschreiben Bedingungen eines Prototype-Frames, die für das zu vergleichende Konzept erfüllt sein müssen. Bei der Markierung 'F' sollen die Bedingungen nicht erfüllt sein und bei 'U' bleiben die Bedingungen unberücksichtigt.

⁶⁸ Dieses Meta-Merkmal entspricht den IMPS in dem Frame-Konzept von WINOGRAD74, S.61, den Relevanz-Markierungen in den SN von QUILLIAN68 (vgl. 4.4.2.2.) und den I-Markierungen von CARBONELL/COLLINS73 (vgl. 4.4.2.3.)

HECKE VOM TYP2	META-MERKMALE VON DESKRIPTOREN			
	'DEFAULT'	'CRITERIAL'	'PRIMARY'	SEKUNDÄR
typisch	W	W	W	W
genaugenommen	U	W	W	F
grob gesprochen	U	F	F	W
letztlich	U	W	F	U

FIGUR 51

Wenn entsprechende Prototype-Frames für die sprachlichen Kategorien *Vogel*, *Säugetier*, *Fisch* und *endlicher Automat* definiert sind und auch die Konzepte *Rotkehlchen*, *Wal*, *Computer* als Frames in KRL repräsentiert werden, dann können folgende Fragen nach einem einfachen, semantisch-orientierten Parsing unmittelbar durch den Pattern-Matcher bearbeitet und positiv beantwortet werden:

Ist ein Rotkehlchen ein typischer Vogel ? ; Alle Bedingungen, die aus Deskriptoren für den *Vogel*-Frame hervorgehen, werden von dem Frame für *Rotkehlchen* erfüllt.

Ist ein Wal genaugenommen ein Säugetier ? ; Die primäre Sichtweise für einen Wal und die grundlegenden Merkmale klassifizieren ihn als Säugetier. Die sekundären Merkmale für den *Säugetier*-Frame wie 'hat-Beine' \neg 'lebt-ständig-im-Wasser' sind allerdings nicht erfüllt.

Ist ein Wal grob gesprochen ein Fisch ? ; Lediglich die sekundären Merkmale von *Fisch* wie 'hat-Flosse' und 'lebt-ständig-im-Wasser' sind für *Wal* erfüllt.

Ist ein Computer letztlich ein endlicher Automat ? ; Die primäre Sichtweise des Frames für *Computer* ist nicht erfüllt, aber alle grundlegenden Merkmale eines endlichen Automaten stimmen mit denen im *Computer*-Frame überein.

Jede Hecke vom Typ 2 spezifiziert also eine Menge von Bedingungen, die von dem Pattern-Matcher für die beiden zu vergleichenden Frames überprüft werden.

4.6.2.3. Die Darstellung der extensionalen Bedeutung vager Konzepte durch objekt-bezogene Prozedurdefinitionen

Eine Darstellung der extensionalen Bedeutung vager sprachlicher Ausdrücke wird in KRL durch objekt-bezogene Prozedurdefinitionen ermöglicht, indem z.B. die extensionale Bedeutung eines vagen Deskriptors wie *groß* als Wert eines Attributs *größe* durch eine an den Deskriptor gebundene Prozedur definiert wird. Auf diese Weise gelingt es, auch mit präzisen Deskriptoren belegte Attribute eines Frames (z.B. *größe* = 190) mit vagen Deskriptoren der entsprechenden Attribute eines anderen Frames (z.8. *größe* = *groß*) zu vergleichen.

In den objekt-bezogenen Prozedurdefinitionen (engl.: procedural attachment) von KRL wurden zwei weit verbreitete Konzepte neuerer Programmiersprachen verschmolzen:

- das Klassenkonzept: Ein Prototype-Frame von KRL ist u.a. vergleichbar mit einer 'class' in SIMULA oder CONCURRENT PASCAL (vgl. BRINCH HANSEN75) oder einem 'cluster' in CLU (vgl. LISKOV76). Auch in der KI bekanntere Sprachen wie SMALLTALK (vgl. LEARNING RESEARCH GROUP76) und PLASMA (vgl. HEWITT75) sind von dem Klassenkonzept stark beeinflusst.
- der Pattern-gesteuerte Prozeduraufruf, der in allen neueren KI-Programmiersprachen möglich ist (vgl. 4.3.2.).

Im folgenden werden bei der objekt-bezogenen Prozedurdefinition sog. 'servant procedures' verwendet, die z.B. den DEDUCE-Prozeduren in FUZZY entsprechen (vgl. 4.3.2.).

BEISPIEL: Die Folge von Schlüsselwörtern 'TRIGGERS (TOESTABLISH...)' veranlaßt den KRL-Interpreter dazu, den darauf folgenden LISP-Ausdruck auszuwerten (vgl. BOBROW/WINOGRAD76, S.29ff). Wir definieren jetzt die in dem Frame für 'Person' vorkommenden Frames 'Aussehen' und 'Privatbereich' (vgl. 4.6.2.1.).

```
(1) [ Privatbereich UNIT SPECIALIZATION
    < SELF { (A Unterscheidungsmerkmal) ;PRIMARY (A Personalakteil)} >
    < hobby (OR (A Tätigkeit) (A Spiel)) >
    < ehfrau (A Person) > ]
```

Der Frame für 'Aussehen' enthält zwei objekt-bezogene Prozedurdefinitionen.

```
(2) [ Aussehen UNIT SPECIALIZATION
    < SELF (A Unterscheidungsmerkmal) >
    < haarfarbe { (A Farbe) (ITEMS blond schwarz) blond;DEFAULT };
                                     CRITERIAL >
    < frisur (ITEMS lang kurz) >
    < gröÙe { groß; 0.4/DEFAULT TRIGGERS(TOESTABLISH (groÙ ?Person)
        (FETCH (?GröÙe !PERSON)) (SUCCEED (groÙ !Person) (SFUNK !GröÙe 170 190)))
        klein TRIGGERS (TOESTABLISH (klein ?Person) (FETCH (?GröÙe !Person))
        (SUCCEED (klein !Person) (*DIF 1 (SFUNK !GröÙe 150 170)))) (A Integer)
        };CRITERIAL >
    < kleidung { (ITEMS elegant sportlich) sportlich; 0.7/DEFAULT} > ]
```

Um die RK möglichst einfach zu gestalten, wurden in (2) anstatt von LISP-Ausdrücken die in Abschnitt 4.3.2. dargestellten Möglichkeiten von FUZZY (kursiv geschrieben) ausgenutzt. Die Attribute 'kleidung' und 'frisur' in (2) sind Beispiele für sekundäre Angaben (vgl. 4.6.2.2.).

```
(3) [ Person1 UNIT INDIVIDUAL
    < SELF (A Person WITH vorname = Peter) > ]
```

```
(4) [ Person2 UNIT INDIVIDUAL
    < SELF (A Person WITH vorname = Dieter) > ]
```

Frames (3) und (4) und die folgenden Frames (5) - (8) sind im Gegensatz zu (1) und (2) Objektframes. Durch einen Deskriptor der Form (THE <Attributbezeichner> FROM <perspective>) kann auf ein bestimmtes Attribut innerhalb eines Frames verwiesen werden.

```
(5) [Personenbeschreibung1 UNIT MANIFESTATION
    < SELF { (THE Äußeres FROM (A Person WITH vorname =Peter))
        (A Aussehen WITH haarfarbe = blond
            gröÙe = groß ;
            kleidung = sportlich
            frisur = kurz ) } > ]
```

Frame (5) ist eine Inkarnation des Prototype-Frames (2).

```
(6) [ Personenbeschreibung2 UNIT MANIFESTATION
    < SELF { (THE Äußeres FROM (A Person WITH vorname=Dieter))
        (A Aussehen WITH frisur = kurz
            haarfarbe = blond
            gröÙe = 180 ) } > ]
```

(7) [Personenbeschreibung3 UNIT MANIFESTATION

< SELF { (THE Privates FROM (A Person WITH vorname = Peter))
(A Privatbereich WITH hobby = Ping-Pong
ehfrau = Emma) } >]

(8) [Personenbeschreibung4 UNIT MANIFESTATION

< SELF { (THE Privates FROM (A Person WITH vorname = Dieter))
(A Privatbereich WITH hobby = Schach) } >]

Enthält die Wissensbasis eines NSS u.a. die Frames (1)-(8), so können z.B. folgende Fragen nach einer syntaktischen und semantischen Analyse mit Hilfe des Pattern-Matchers von KRL beantwortet werden:

- (a) *Sind sich Peter und Dieter in privater Hinsicht sehr ähnlich ?*
(b) *Sind sich Peter und Dieter äußerlich sehr ähnlich ?*

Durch die Eigennamen *Peter* und *Dieter* werden die Frames (3) und (4) angesprochen. Die limitierenden Adverbiale (vgl. 2.3.1.) *in privater Hinsicht* und *äußerlich* werden als Spezifikationen von entsprechenden Attributen einer 'perspective' (THE Privates, THE Äußeres) interpretiert und der vage Term *ähnlich* wird als Vergleichsoperator über den durch die limitierenden Adverbiale angesprochenen Frames (5) und (6) bzw. (7) und (8) aufgefaßt.

Die linguistische Hecke *sehr* steuert die Anzahl der zu vergleichenden Attribute und legt den für eine positive Antwort erforderlichen Gütegrad des 'Matchings' fest (vgl. 4.6.3.).

Da die Frames (7) und (8) kaum Information enthalten und die wenigen Angaben völlig voneinander abweichen, führt der Vergleich zu einem negativen Ergebnis, und Frage (a) wird mit *Nein* beantwortet.

Der zur Beantwortung von Frage (b) erforderliche Vergleich von Frame (5) und (6) ergibt für die Attribute 'haarfarbe' und 'frisur' eine völlige Übereinstimmung. Um das Attribut 'größe' vergleichen zu können, wird die in Frame (2) spezifizierte Prozedur von dem Interpreter auf Frame (6) angewendet. Es ergibt sich der Wert (groß . 0.5). Dieser Wert wird von dem Matcher mit der entsprechenden Angabe in Frame (5), die als (groß . 1) interpretiert wird, verglichen.

Da für das Attribut 'kleidung' in Frame (6) kein Wert spezifiziert ist, wird in dem Prototype-Frame (2) nach einer Standardspezifikation gesucht. Der Matcher findet den Wert (sportlich . 0.7) und vergleicht ihn mit dem Wert (sportlich . 1) aus Frame (5).

Nach einer Auswertung eines Gütekriteriums für Vergleiche, in das z.B. die unterschiedliche Gewichtung einzelner Attribute eines Frames (z.B. durch Meta-Merkmale bestimmt) eingeht, kann die Antwort *Peter und Dieter sind sich äußerlich ziemlich ähnlich* generiert werden.

Die Repräsentation von vagem Wissen in KRL-Frames ermöglicht es also, nicht nur vage Prädikationen (z.B. *ähnlich(Peter,Dieter)*) durch Gradangaben (z.B. *sehr ähnlich*, *ziemlich ähnlich*) sondern auch durch die jeweilige Sichtweise (z.B. *in privater Hinsicht*, *äußerlich*) zu modifizieren (vgl. 1.3.1.).

Die Konzepte der Repräsentationssprache KRL, die sich in den vorangegangenen Abschnitten für eine Darstellung von vagem Wissen als besonders geeignet erwiesen haben, können in folgender Liste zusammengefaßt werden:

- Prototype-Frames
- Standardspezifikationen als vage Erwartungsnormen
- Multiple Deskriptoren
- 'perspectives' zur Darstellung von Referenzmengen vager Ausdrücke
- unvollständige Mengen als Datenstruktur
- explizite Angabe des Wahrheitswertes für Propositionen

Die oben vorgeschlagenen Erweiterungen zur

- expliziten Darstellung von F-Mengen
- Modifikation der Bindungsstärke von Standardspezifikationen
- prozeduralen Einbettung von Wissen über die extensionalen Bedeutung vager sprachlicher Ausdrücke

sollten in Hinblick auf die Repräsentation von vagem Wissen in KRL integriert werden.

Obwohl durch KRL konkretisiert, befindet sich das Frame-Konzept im Vergleich zu den anderen in dieser Arbeit behandelten RS noch immer im Experimentierstadium und ist in der KI nicht unumstritten. Auf die Arbeiten, die sich kritisch mit dem allgemeinen Frame-Konzept (vgl. z.B. FELDMAN75, WILKS76) oder mit KRL speziell (vgl. z.B. BORNAT/WIELINGA76, HAYES77) auseinandersetzen, gehen wir hier nicht ein, da sich die Kritik nicht direkt auf die Darstellung von vagem Wissen in der Repräsentationssprache 'Frames' bezieht.

4.6.3. MÖGLICHKEITEN UND PROBLEME DES 'FUZZY MATCHING'

Für Repräsentationssprachen wie KRL, die auf dem Frame-Konzept aufbauen, sind als Inferenzprozesse vor allem Vergleiche zwischen verschiedenen Frames vorgesehen. Bobrow und Winograd stellen fest:

...reasoning is dominated by a process of recognition in which new objects and events are compared to a stored set of expected prototypes.

(BOBROW/WINOGRAD76, S.18)

Alle mit Inferenzen verbundenen Vergleichsprozesse werden dabei dem sehr weit gefaßten Begriff 'Matching' untergeordnet (vgl. auch MOORE/NEWELL74).

Wir gehen im folgenden kurz auf eine spezielle Klasse solcher Match-Prozesse ein, die im Zusammenhang mit der Darstellung und Verarbeitung von vagem Wissen von Bedeutung sind und meist als 'fuzzy matching' bezeichnet werden. Dabei versteht man unter 'fuzzy matching' alle Match-Operationen, die nicht wie üblich eine binäre Entscheidung sondern einen bestimmten Grad der Übereinstimmung zwischen einem Pattern und einem damit verglichenen Datenobjekt ergeben (vgl. auch BOBROW75, S.27).

Oft wird auch von einem 'close-mismatch' (vgl. BOBROW/NORMAN75, S. 134) oder einem 'best match' (vgl. BOBROW/WINOGRAD76, S.24) gesprochen, wobei man davon ausgeht, daß dem Pattern-Matcher ein Gütekriterium zur Einordnung der Qualität eines Vergleichs zwischen einem Datenobjekt und mehreren Patterns⁶⁹ zur Verfügung steht.

BEISPIEL: Ein Vergleich zwischen der Liste (X Y Z Y) und dem Pattern (X Y Y Z) ergibt einen höheren Grad der Übereinstimmung als ein Vergleich mit dem Pattern (Z X).

Neben einem solchen 'syntaktischen Matching' wird in der KI oft auch ein 'semantisches Matching' angestrebt, bei dem sich z.B. für einen Vergleich von (körpergröße = 180) mit dem Pattern (körpergröße = groß) ein höherer Grad der Übereinstimmung ergibt als beim Vergleich mit dem Pattern (körpergröße = klein).

Obwohl bereits McCARTHY58 forderte, daß Pattern-Matcher entwickelt werden müssen, die auch partielle Erfolge beim Vergleich eines Datenobjekts mit einem Pattern feststellen können, ist meines Wissens bis heute kein NSS entwickelt worden, daß mit einem allgemeinen 'fuzzy matching' arbeitet (vgl. auch WINSTON77, S. 333). Auch die Forderung, ein annäherndes Matchen in KI-Programmiersprachen wie FUZZY (vgl. LeFAIVRE74, S. 152) oder QLISP (vgl. SACERDOTIetal.76, S.21) zu integrieren, bleibt weiterhin unerfüllt.

Vergleichsprozesse zwischen konkret vorgefundenen Sachverhalten und bereits bekannten

⁶⁹ Man spricht dann von 'multiplem Matching'.

Konzepten gehören zu den elementaren kognitiven Prozessen. Beispiele für besonders komplexe Vergleichsprozesse, die nur selten unmittelbar zu einer binären Entscheidung führen sind

- die medizinische Diagnose : eine Menge von Symptomen wird mit einer Reihe von Krankheitsbildern ('Krankheits-Frames') verglichen.
- die juristische Subsumtion : die Fallwirklichkeit wird mit unterschiedlichen Tatbeständen verglichen.

Bei der Formalisierung und Algorithmisierung solcher Vergleichsprozesse im Rahmen der KI stößt man sofort auf semantische Probleme:

In general, the problem of finding the closest match to a signal is extremely complex, since the notion of 'closeness' depends on interactions between both the uses and the meaning of the symbols.

(WINOGRAD76, S.148)

Jede Art von 'fuzzy matching' erfordert also selbst wieder Wissen über die zu vergleichenden symbolischen Strukturen. Dies scheint der eigentliche Grund dafür zu sein, daß bis heute noch kein allgemeines Verfahren für ein 'fuzzy matching' entwickelt wurde.

In Abschnitt 4.6.2.2. haben wir gezeigt, wie einfache Vergleiche zwischen Datenobjekten und Patterns (dort: Objektframes und Prototype-Frames) bei einer Berücksichtigung von Metawissen (dort in Form von Meta-Merkmalen) verschiedene Abstufungen der Übereinstimmung ergeben können. Auch ein 'fuzzy matching' zwischen vagen und präzisen Werten einer Variablen (vgl. 4.3.4., 4.6.2.3.) war nur durch die zusätzliche Einbettung von Wissen über die extensionale Bedeutung vager sprachlicher Ausdrücke möglich.

Abschließend stellen wir daher fest, daß die Probleme des 'fuzzy matching' nicht durch das Finden des 'richtigen' Algorithmus und der 'richtigen' Gütekriterien gelöst werden können, sondern unmittelbar mit der adäquaten Repräsentation von vagem Wissen und vor allem von Metawissen zusammenhängen.

5 ABSCHLIESSENDE ÜBERSICHTEN

5.1. ZUR KRITIK AN DER VERWENDUNG DER 'FUZZY LOGIC' IN KI-SYSTEMEN

Im folgenden Abschnitt setzen wir uns mit einigen in der KI vorgebrachten Kritiken auseinander, die sich gegen eine Integration der Theorie der F-Mengen und der F-Logik in Repräsentationssprachen richten.

Zunächst muß festgestellt werden, daß sich diese Kritik nicht gegen eine Integration von vagem Wissen und vagen Inferenzen in NSS richtet. Es ist in der KI allgemein anerkannt, daß in NSS auch auf vagem Wissen basierende Inferenzen berücksichtigt werden müssen.

Es wird also nicht das Ziel, das man bei der Entwicklung einer Theorie der F-Mengen und der F-Logik verfolgt, sondern die spezielle Methode, die mit diesem Ansatz gegeben ist, kritisiert.

Die in der KI am häufigsten zitierte kritische Anmerkung zur Theorie der F-Mengen ist McCarthys Frage

Where do all those numbers come from ? .

Diese Kritik bezieht sich auf die meist in Form von Zahlen aus dem reellen Intervall $[0,1]$ angegebenen Zugehörigkeitswerte von Elementen zu einer F-Menge.

Bereits in Abschnitt 4.1.2. wurde darauf hingewiesen, daß eine F-Logik axiomatisch ohne Rückgriff auf Zahlenwerte eingeführt werden kann:

...we are able to both do without the numbers in fuzzy reasoning and also to introduce them if required, not on an arbitrary basis, but directly from the axiom schema....

(GAINES76, S.558)

Da aber aus praktischen Erwägungen heraus (vgl. 4.1.2.) in den vorgeschlagenen RK häufig Zahlenwerte benutzt werden, kann McCarthys Frage auch als Hinweis auf das methodische Problem verstanden werden, wie die Zahlenwerte ggf. zu ermitteln sind.

Neben dem psychologischen Experiment mit Testpersonen (vgl. 3.3.) dient beim Aufbau von KI-Systemen in einigen Fällen die Befragung einer einzelnen Person (des 'Experten', vgl. 4.2.1.) der Ermittlung solcher Zahlenwerte. Die in der KI am häufigsten verwendete Methode ist allerdings die Introspektion, die auch in dieser Arbeit bei der Konstruktion von Beispielen angewandt wurde. Es ist offenkundig, daß dieses in der KI praktizierte Verfahren von einem wissenschaftstheoretischen Standpunkt aus betrachtet problematisch ist. Allerdings trifft diese Kritik dann allgemein auf die Konstruktion von Modellen in der KI zu. Wir könnten also an McCarthy die Gegenfrage stellen: Wo kommen die Produktionen der formalen Grammatiken her, die von Algorithmen zur syntaktischen Analyse in NSS verwendet werden? Wo kommen die z.B. im Prädikatenkalkül formulierten Inferenzregeln her? usw.

Hayes greift die Kritik von McCarthy auf (vgl. HAYES74, S. 77ff). Bezogen auf Adjektive wie *groß*, *alt* und *teuer* stellt Hayes fest *These are precise words but they refer to vague measuring scales* (ebd., S.77). Obwohl Hayes nicht definiert, was er unter Vagheit versteht, kann man aus seinem Beispiel *when I say a heap is small, I mean just that* (ebd.) schließen, daß er sich gegen die Auffassung wendet, für ein Wort wie *klein* sei in der NS keine 'klare' Semantik festgelegt. Diese Auffassung wird aber in den Arbeiten, die eine Integration von Konzepten der F-Logik in bestehende RS vorschlagen, nicht vertreten, sondern es wird umgekehrt in diesen Arbeiten gerade der Versuch gemacht, die Semantik solcher Ausdrücke formal zu explizieren.

Auch die Feststellung von Hayes, daß sich diese Adjektive auf 'vage Skalen' beziehen kann nicht als Argument gegen eine F-Logik akzeptiert werden, da Zadehs Anwendung der F-Mengen vom Typ 2 (vgl. 3.1.2., 3.2.1.) zur Explikation der extensionalen Semantik von nach unserer Definition (vgl. 1.3.1.) vagen Adjektiven auf solchen 'vagen Skalen' beruht:

...a role comparable to that of a unit of measurement is played by one or more primary fuzzy sets from which other sets can be generated through the use of linguistic modifiers such as very, quite, more or less, extremely, essentially, completely etc.

(ZADEH76, S. 250)

Die Beobachtung von Hayes, daß *an intermediate heap will be considered small if it began as small and grew, and considered large if it began as large and shrunk* kann nicht als Argument gegen eine F-Logik akzeptiert werden, sondern lediglich als Hinweis auf die Kontextabhängigkeit der Bedeutungen von Wörtern wie *groß* und *klein*. Diese Kontextabhängigkeit kann aber, wie wir in dieser Arbeit gezeigt haben (vgl. z.B. 3.2.1, 4.3.2.2., 4.4.3., 4.6.2.3.), in einer Explikation der Semantik relativer Adjektive durch F-Mengen durchaus berücksichtigt werden. Außerdem ist die Beobachtung von Hayes m.E. kontraintuitiv. Ein Haufen mittlerer Größe, der zunächst 'groß' ist und dann verkleinert wird, wird meist als *kleiner Haufen* bezeichnet, wobei die Referenzmenge zu dem relativen Adjektiv *klein* der ursprüngliche, 'große' Haufen ist.

Auch einige Bemerkungen von Minsky in seiner grundlegenden Arbeit über Frames werden häufig als Argumente gegen eine Anwendung der F-Logik in der KI interpretiert (vgl. MINSKY75, S.272ff). Minsky stellt zunächst fest

People are very poor at handling magnitudes or intensities on any absolute scale (MINSKY75, S. 273).

One hardly ever needs quantitative precision (ibd. S. 274).

Diese Beobachtungen widersprechen aber nicht sondern decken sich mit entsprechenden Feststellungen Zadehs (vgl. 3.2.1.). Minsky wendet sich gegen 'quantitative Modelle' in der KI:

...the output of a quantitative mechanism, be it numerical, statistical, analog, or physical (nonsymbolic) is too structureless and uninformative to permit further analysis. A number cannot reflect the considerations that formed it (ibd. S. 275).

Gegen diese sehr elementare Feststellung kann nichts eingewendet werden. Allerdings trifft diese Aussage, interpretiert man sie als Kritik an den in der F-Logik oft als Wahrheitswert verwendeten Zahlen, auch auf die zweiwertige Logik zu. Die Auswertung eines einfachen booleschen Ausdrucks im Sinne der zweiwertigen Logik ergibt auch einen Wahrheitswert, aus dem sich nicht rekonstruieren läßt, wie er zustande gekommen ist.

Außerdem zeigt ein Beispiel wie das MYCIN-System (vgl. auch SCOTTetal. 77) , daß Inferenzen, die auf einem 'quantitativen Modell' beruhen, durchaus nachträglich begründet werden können, indem alle Zahlenwerte in entsprechende linguistische Hecken transformiert und die angewandten Inferenzregeln 'vorgelesen' werden.

Die Skepsis Minskys und der bisher genannten Kritiker gegenüber RK, in denen Zahlen auftreten, ist allerdings insofern berechtigt, als ein wesentlicher Fortschritt der KI gegenüber den z.B. im Rahmen der Kybernetik und statistischen Entscheidungstheorie entwickelten Modellen darin besteht, die Mächtigkeit symbolischer, diskreter Darstellungsformen erkannt zu haben. In dieser Arbeit wurde daher auch soweit wie möglich von den in der KI entwickelten RS bei der Repräsentation von vagem Wissen Gebrauch gemacht und nur dort, wo es sich als notwendig erwies (z.B. bei der Repräsentation der Hecken vom Typ 1) wurden Konzepte aus der Theorie der F-Mengen verwendet.

Wie wir gezeigt haben, gibt es auch Bereiche der Repräsentation von vagem Wissen, die nicht befriedigend durch eine Integration von Konzepten der F-Logik in bestehende RS bearbeitet werden können.

Besonders für die Repräsentation von komplexen Inferenzstrategien über vagem Wissen (vgl. z.B. 4.4.5.2.) ist die F-Logik in ihrer derzeitigen Form ungeeignet. Auch die Teilnehmer an einer in SRIDHARAN75 (vgl. Kap. 'Methods of inference') nachgedruckten Diskussion, in der es um die Repräsentation von vagen Schlußregeln in medizinischen Expertensystemen ging, kamen zu dieser Auffassung. In der genannten Diskussion wurde abschließend festgestellt, daß es zu den mit einer numerischen Bewertung der Inferenzregeln arbeitenden Modellen zur Zeit keine praktikable Alternative gibt.

Allgemein kann man feststellen, daß alle bisher implementierten NSS, in denen Aspekte der Repräsentation von vagem Wissen berücksichtigt werden, mehr oder weniger stark von Konzepten der F-Logik beeinflußt sind.

Beurteilt man die an der F-Logik orientierten RK stark vereinfachend nur nach dem folgenden von WILKS77 verwendeten Kriterium, so beweisen die im folgenden Abschnitt 5.2. dargestellten, bereits implementierten Systeme, daß der hier verfolgte Ansatz im Rahmen der KI sinnvoll ist:

It seems to me worth emphasising again that there can be no other ultimate test of a system for understanding natural language than its success in doing some specific task, and that to pretend otherwise is to introduce enormous confusion. Considerations of logic or psychological plausibility may indeed be suggestive in the construction of AI language systems but that is quite another matter from their ultimate accountability, which can only be whether or not they work.

(WILKS77, .S. 72)

5.2. DIE REPRÄSENTATION VON VAGEM WISSEN IN BESTEHENDEN NSS: PERFORMANZ UND ANWENDUNG DER SYSTEME

In diesem letzten Abschnitt greifen wir eine Fragestellung, mit der wir uns bereits am Anfang dieser Arbeit auseinandergesetzt haben (vgl. 1.2.2.), nochmals auf, indem wir am Beispiel einiger implementierter NSS zeigen, welcher Stellenwert der Repräsentation und Verarbeitung von vagem Wissen bei der Entwicklung und besonders beim praktischen Einsatz von NSS zukommt.

Vergleicht man die wenigen, vollständig implementierten NSS, bei deren Entwurf Probleme der Repräsentation und Verarbeitung von vagem Wissen berücksichtigt wurden, so fällt zunächst auf, daß diese Systeme vorwiegend mit Mikrowelten arbeiten, in denen räumliche Beziehungen zwischen einfachen Objekten im Vordergrund stehen. Dies ist darauf zurückzuführen, daß trotz der einfachen Grundstruktur solcher Mikrowelten gerade bei der Verbalisierung räumlicher Beziehungen in Ausdrücken wie *x ist vor y*, *x ist neben y*, *x ist in der Nähe von y* die Vagheit vieler NS-Formulierungen besonders deutlich erkennbar ist. In mehreren Arbeiten zur KI wird daher gefordert, daß in Dialogen über solche räumlichen Beziehungen auch Formulierungen wie *ein kleines bißchen hinter x*, *weit rechts von y* usw. vom System akzeptiert bzw. generiert werden (vgl. WINSTON75, S. 201, BADLER 75, S.198ff, SONDHEIMER76).

Besonders die sog. FUZZY ROBOT USERS GROUP, die von J. Goguen geleitet wird, hat sich in den letzten beiden Jahren intensiv mit der Implementation von NSS beschäftigt, die vage Formulierungen über räumliche Beziehungen analysieren. So wurde z.B ein System implementiert, das über einer Winograds Blockwelt ähnlichen Mikrowelt arbeitet, aber im Gegensatz zu dem SHRDLU-System systematisch die Unschärfe sprachlicher Ausdrücke berücksichtigt (vgl. SHAKET76).

Shaket beschäftigt sich ausschließlich mit den referenzsemantischen Aspekten eines Dialogs. Das System reagiert auf NS-Eingaben eines Benutzers, in denen ein Objekt oder eine Menge von Objekten der Mikrowelt durch vage Formulierungen angesprochen werden, mit der Ausgabe von eindeutigen Bezeichnungen für diese Objekte.

BEISPIEL: BENUTZER:⁷⁰ *Finde einen sehr langen Stab !*
SYSTEM: *Objekt1*
BENUTZER: *Finde einen winzigen Quader !*
SYSTEM: *Objekt3*
BENUTZER: *Finde einen großen, länglichen Block in der Mitte !*
SYSTEM: *Objekt4*

In Shakets System sind nicht nur Adjektive wie *länglich*, *kurz*, *dick* usw. sondern auch Nomen wie *Latte* oder *Stab* als F-Mengen definiert. Nach dem sog. 'Prinzip der maximalen Bedeutsamkeit' wird bei Shaket die Bezeichnung für den Gegenstand ausgegeben, aus dessen Eigenschaften sich der höchste Zugehörigkeitswert für die durch die vage Formulierung des Benutzers spezifizierte F-Menge ergibt. Insgesamt sind sechs Hecken vom Typ 1 in Shakets System als Operatoren über F-Mengen implementiert.

Zu Shakets System muß allerdings kritisch angemerkt werden, daß die direkte prozedurale Repräsentation von Wissen in der Programmiersprache APL den innerhalb der KI heute gültigen Standards nicht mehr entspricht. Auch die Tatsache, daß Shaket für jedes Objekt exakte, vollständige Angaben über dessen Höhe, Breite, Dichte und die Wellenlänge des von ihm reflektierten Lichtes (sic !) voraussetzt, muß sich negativ auf die Beurteilung eines Systems auswirken, das beansprucht, vage Informationen verarbeiten zu können. Die sprachlichen und deduktiven Fähigkeiten des Systems sind außerdem verglichen mit NSS wie SHRDLU, MARGIE, SAM sehr begrenzt.

Ein anderes in der FUZZY ROBOT USERS GROUP entwickeltes NSS wird als Schnittstelle zu einem simulierten Roboter eingesetzt, der einen Weg durch einen vorgegebenen Irrgarten finden soll. Das NSS analysiert vage Hinweise des Benutzers, die dem Roboter bei seiner Aufgabe helfen sollen (vgl. GOGUEN75, GERSHMAN76).

⁷⁰ Alle folgenden Beispiele wurden soweit erforderlich vom Englischen ins Deutsche übertragen.

BEISPIEL: BENUTZER: *Gehe zunächst ein kurzes Stück südwärts und dann ziemlich lange in östlicher Richtung.*

Das System bezieht diese Hinweise nach einer Analyse in seinen Lösungsplan ein.

Ein von DENOFSKY76 entwickeltes System 'versteht' vage räumliche Beziehungen, die durch die Verbindung von Wörtern wie nahe und weit mit linguistischen Hecken vom Typ 1 ausgedrückt werden können. Bei der Entwicklung des Systems wurde vor allem die Kontextabhängigkeit der Interpretation solcher Formulierungen berücksichtigt.

Auch in dem System HAM-RPM wird die Unscharfe sprachlicher Ausdrücke im Bereich der räumlichen Beziehungen berücksichtigt (vgl. WAHLSTER/v.HAHN76, S. 220ff).

BEISPIEL: Folgender Dialogausschnitt, der sich auf das in HAM-RPM gegenwärtig als Mikrowelt benutzte Interieur bezieht, veranschaulicht diese Feststellung.

BENUTZER: *Wo steht der Tisch ?*
SYSTEM: *In etwa rechts neben dem weißen Schrank.*

Zwei experimentelle Systeme, die nicht an geometrische Mikrowelten gebunden sind, wurden von LeFAIVRE74 und BALZER75 implementiert. LeFaivre hat in der von ihm entwickelten KI-Programmiersprache FUZZY (vgl. 4.3.2.) ein sehr einfaches, kleines FAS implementiert, das einige vage Adjektive und linguistische Hecken mit Hilfe von in der assoziativen Datenbasis gespeichertem Wissen analysiert und bei der Generierung von Antworten verwendet (vgl. LeFAIVRE74, S.134ff).

BEISPIEL: BENUTZER: *Ist eine Schlange sehr hübsch ?*
SYSTEM: *Nein, eine Schlange ist ziemlich häßlich .*

Das System von BALZER75 wurde innerhalb des KI-Forschungsschwerpunktes 'Automatische Programmierung' entwickelt und partiell implementiert. Es generiert einfache Programme aus z.T. unvollständigen und vagen Programmspezifikationen.

Die bisher erwähnten Systeme wurden nicht für eine spezielle, praktische Anwendung konzipiert sondern als experimentelle Systeme, in denen u.a. verschiedene Methoden zur Darstellung und Verarbeitung von vagem Wissen erprobt werden sollen.

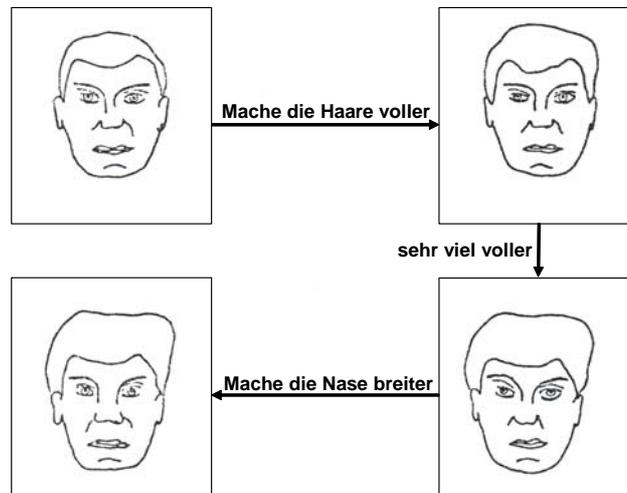
Als Beispiele für Programme, die einen klaren Anwendungsbezug aufweisen, sei hier nochmals auf das MYCIN-System (vgl. SHORTLIFFE74 und die unter dem Namen SCHOLAR implementierten NSS (vgl. CARBONELL/COLLINS73) hingewiesen. Beide Systeme können zumindest als Pilotprojekte für eine praktische Anwendung von KI-Systemen betrachtet werden, in denen die Darstellung von vagem Wissen berücksichtigt wurde (vgl. 4.2.2., 4.4.2.3.).

Sowohl in MYCIN als auch in den SCHOLAR-Systemen enthält die Wissensbasis vage Inferenzregeln, die auf die eingegebenen oder in der Wissensbasis abgespeicherten Aussagen angewandt werden. In beiden Systemen spielt die Verarbeitung von vagem Wissen und die Analyse bzw. Generierung von vagen sprachlichen Formulierungen eine entscheidende Rolle bei der Gestaltung des Dialogs zwischen dem System und seinen 'naiven' Benutzern (vgl. 1.2.2.), nicht zuletzt auch unter dem Gesichtspunkt des sog. 'human engineering'.

Am Schluß dieser Arbeit soll am Beispiel eines weiteren NSS gezeigt werden, daß es bereits heute Situationen gibt, in denen ein beliebiger Bürger in direkten Kontakt mit einem NSS kommt, das auf in der KI entwickelten Techniken zur Repräsentation von vagem Wissen basiert.

Für den Erkennungsdienst der amerikanischen Polizei wurde eine NS-Schnittstelle zu einem System entwickelt, mit dem ein Zeuge im Dialog eine Phantomzeichnung entwerfen kann (vgl. RHODES/KLINGER77). Ausgehend von einer auf einem graphischen Display dargestellten Skizze eines Gesichtes, gibt der Benutzer NS-Beschreibungen des Aussehens der von ihm erkannten Person in das als SKETCH bezeichnete System ein, auf die das System mit einer entsprechenden Veränderung der Phantomzeichnung reagiert. Der zeichnerisch meist unbegabte Benutzer kann mit diesem System sukzessive eine Skizze entwerfen, die schließlich seiner Erinnerung an das Aussehen der gesuchten Person in etwa entspricht. Da bei der sprachlichen Beschreibung von Gesichtern vage Formulierungen wie *ovales Gesicht, länglicher Mund, breite Nase, große Ohren, sehr volles Haar*

usw. gehäuft auftreten, ist für die genannte Anwendung eine referenzsemantische Analyse vager, relativer Adjektive und linguistischer Hecken unbedingt erforderlich (vgl. Fig. 52).



FIGUR 52

Obwohl SKETCH auf einem Kleinrechner vom Typ IMLAC mit einem geringen Hauptspeicherausbau von 8KWö bei einer Wortlänge von 16 Bits implementiert wurde, sind in diesem System neben einem sehr einfachen Parser und Möglichkeiten zur Berücksichtigung des Dialogkontextes sowie zur Auflösung einfacher Mehrdeutigkeiten vor allem auch Algorithmen zur Analyse vager Prädikationen und linguistischer Hecken realisiert worden. Konzepte wie *breite Nase* oder *breiter Mund* werden in SKETCH als F-Mengen dargestellt, die durch linguistische Hecken vom Typ 1 modifiziert werden können.

Obwohl in SKETCH der Dialogkontext stark eingeschränkt ist und komplexe Repräsentationsprobleme wie die Darstellung vager Inferenzregeln oder Hecken vom Typ 2 völlig ausgeklammert wurden, zeigt dieses im Vergleich mit den anderen in dieser Arbeit betrachteten KI-Systemen beinahe triviale Programm, daß in der KI entwickelte Techniken und Konzepte bereits heute unmittelbaren Einfluß auf Computeranwendungen in der alltäglichen Praxis haben.

LITERATURVERZEICHNIS

- ABELSON73 *Abelson, R. P.: The structure of belief systems. SCHANK/COLBY73, S. 287-339*
- ADJUKIEWICZ75 *Adjukiewicz, K.: Pragmatic logic. Dordrecht: Reidel 1975*
- ALSTON64 *Alston, W.: Philosophy of language. Englewood Cliffs: Prentice-Hall 1964*
- ALTMAN76 *Altman, H.: Die Gradpartikeln im Deutschen. Tübingen: Niemeyer 1976*
- BADLER75 *Badler, N. I.: Temporal scene analysis: Conceptual descriptions of object movements. Toronto Univ., Dep. of Computer Sc., Techn. Report No. 80, Feb. 1975*
- BALZER75 *Balzer, R. M.: Imprecise program specification. Univ. of Southern California, Inf. Sc. Inst., Report ISI/RR-75, Dez. 1975*
- BARTSCH72 *Bartsch, R.: Adverbialsemantik. Frankfurt: Athenäum 1972*
- BARTSCH/VENNEMANN73 *Bartsch, R./Vennemann, T.: Semantic structures. Frankfurt: Athenäum 1973²*
- BAUMGART72 *Baumgart, B. G.: Micro Planner alternate reference manual. Stanford Univ., AI-Lab., Operating note No. 67, April 1972*
- BECKER73 *Becker, J. D.: A model for the encoding of experimental information. SCHANK/COLBY73 S. 396-434*
- BERKOWSKI70 *Berkowski, L. (ed.): Jan Lukasiewicz, selected works. Amsterdam: North-Holland 1970*
- BIERWISCH67 *Bierwisch, M.: Some semantic universals of German adjectivals. Foundations of Language 3, 1967, S. 1-36*
- BLACK37 *Black, M.: Vagueness. Philosophy of Science 4, 1937, S. 427-455*
- BLACK49 *Black, M. : Vagueness. An Exercise in logical analysis. Black, M.: Language and philosophy. Studies in method. N.Y.: Cornell Univ. Press 1949*
- BLACK70 *Black, M.: Reasoning with loose concepts. Black, M.: Margins of precision. Essays in logic and language. N.Y.: Cornell Univ. Press 1970, S. 1-13*
- BLACK73 *Black, M.: Sprache. München: Fink 1973*
- BLAU73 *Blau, U.: Zur 3-wertigen Logik der natürlichen Sprache. Papiere zur Linguistik 4, 1973, S. 20-96*
- BLEDISOE75 *Bledsoe, W. W.: Non-resolution theorem proving. Univ. of Texas at Austin, Report ATP29, Sept. 1975*
- BOBROW75 *Bobrow, D. G.: Dimensions of representation. BOBROW/COLLINS75, S. 1-34*
- BOBROW/COLLINS75 *Bobrow, D. G./Collins, A. M. (eds.): Representation and understanding: Studies in cognitive science. N.Y.: Academic 1975*
- BOBROW/NORMAN75 *Bobrow, D. G./Norman, D. A.: Some principles of memory schemata. BOBROW/COLLINS75, S. 131-150*
- BOBROW/RAPHAEL74 *Bobrow, D. G. /Raphael, B.: New programming languages for artificial intelligence. Computer Surveys, 6, 3, Sept. 1974, S. 153-174*
- BOBROW/WINOGRAD76 *Bobrow, D. G./Winograd, T.: An overview of KRL, A knowledge representation language. Xerox Palo Alto Research Center, Juni 1976*
- BOBROWetal.76 *Bobrow, D. G. / Kaplan, R. M. / Kay, M. / Norman, D. A. / Thompson, H. / Winograd, T.: GUS- A frame-driven dialog system. Xerox Palo Alto Research Center 1976*

- BÖCKERetal.75 Böcker, H. -D. / Schümann, M. / Ueckert, H.: Einführung in LISP 1.6. Univ. Hamburg, Vervielf., 1975
- BOLEY74 Boley, H.: Einfache natürlichsprachliche Dialoge mit einem semantischen Netzwerkprogramm. Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-10/74, 1974
- BOLEY75 Boley, H.: Zur Repräsentation von Wissen. Brunnstein, K. (ed.): Zur Konzeption eines Frage-Antwort-Systems für den computergestützten Unterricht. Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-24/75, 1975
- BOLEY76a Boley, H.: A theory of representation (-languages, -constructions, and -relations). Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-38/76, 1976
- BOLEY76b Boley, H.: Directed recursive labelnode hypergraphs: a new representation-language. Advances in natural language processing, Preprints of a workshop held at the Univ. of Antwerp, Okt. 1976 (erscheint auch in: Artificial Intelligence, 8,3, 1977)
- BOLINGER72 Bolinger, D.: Degree words. Paris: Mouton 1972
- BORGIDA75 Borgida, A. T.: Topics in the understanding of English sentences by computer. Univ. of Toronto, Dep. of Computer Sc., Techn. Report No. 78, Feb. 1975
- BORNAT/WIELINGA76 Bornat, R. / Wielinga, B.: An underview of KRL - is it a programming language? AISB Newsletter, 24, Nov. 1976, S. 24-26
- BRACHMAN76 Brachman, R. J.: What's in a concept: Structural foundations for semantic networks. COLING 76, Preprints of 6th intern. Conf. on Comp. Linguistics, Ottawa 1976
- BRINCH HANSEN75 Brinch Hansen, P.: Concurrent Pascal report. California Inst. of Technology, Inform. Sc. , Pasadena 1975
- BROWN/BURTON75 Brown, J.S. / Burton, R. R.: Multiple representations of knowledge. BOBROW/COLLINS75, S. 311-349
- BROWNetal.77 Brown, D. C. / Buttelmann, H. W. / Chandrasekaran, B. / Kwasny, S. C. / Sondheimer, N. K.: Natural language graphics. SIGART Newsletter, 61, Feb. 1977, S. 57-58
- BRUCE71 Bruce, B.: A model for temporal references and its application in a question answering program. Artificial Intelligence, 2, 1971
- BRUCE75 Bruce, B.: Belief systems and language understanding. BBN-Report No. 2973, AI-Report No. 21, Jan. 1975, Cambridge, Mass.
- BRUNNSTEIN77 Brunnstein, K. (ed.): Gesellschaftliche Auswirkungen großer Informationssysteme aus der Sicht verschiedener Disziplinen. Arbeitspapiere eines Werkstattgesprächs. Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-46/77, März 1977
- CARBONELL/COLLINS73 Carbonell, J. R. / Collins, A. M.: Natural semantics in artificial intelligence. IJCAI3, 1973, S. 344-351
- CERCONE/SCHUBERT75 Cercone, C. / Schubert, L.: Toward a state based conceptual representation. IJCAI4, 1975, S. 83-90
- CHANDRASEKARAN75 Chandrasekaran, B.: Artificial intelligence: the past decade. Rubinoff, M. / Yovits, M. C. (eds.) Advances in computers, 13, N.Y.: Academic 1975, S. 169-232
- CHANG76 Chang, C. L.: DEDUCE - A deductive query language for relational data bases. Chen, C. H. (ed.): Pattern recognition and artificial intelligence. N.Y. : Academic 1976, S. 108-134

- CHANG/LEE73 *Chang, C. L. / Lee, R. C. T.: Symbolic logic and mechanical theorem proving. N.Y.: Academic 1973*
- CHARNIAK75 *Charniak, E.: Organization and inference in a frame-like system of common-sense knowledge. NASH-WEBBER/SCHANK75, S. 46-55*
- CHARNIAK76 *Charniak, E.: Inference and knowledge I, II. CHARNIAK/WILKS76, S. 1-21, 129-154*
- CHARNIAK77 *Charniak, E.: A framed painting. Univ. Genf, Inst. for Semantic and Cognitive Studies, Report, 1977*
- CHARNIAK/WILKS76 *Charniak, E. / Wilks, Y. (eds.): Computational semantics. An introduction to artificial intelligence and natural language comprehension. Amsterdam: North-Holland 1976*
- CHOMSKY73 *Chomsky, N.: Aspekte der Syntax-Theorie. Frankfurt: Suhrkamp 1973 (Original: 1965)*
- CODD74 *Codd, E. F.: Seven steps to rendezvous with the casual user. Klimbie, J. W. / Koffman, K. L. (eds.): Data base management. Amsterdam: North-Holland 1974, S. 179-199 (nachge-druckt in: Tsichritzis, D. / Chamberlin, D. D. / Taylor, R. W. (eds.): Tutorial on data base languages and systems. San Francisco, 1976, S. 38-55)*
- COFER76 *Cofer, C. N. (ed.): The structure of human memory. San Francisco: Freeman 1976*
- COLBY75 *Colby, K. M.: Artificial paranoia: A computer simulation of paranoid processes. N.Y.: Pergamon 1975*
- COLLINS/WARNOCK74 *Collins, A. M. / Warnock, E. H. : Semantic networks. BBN-Report No. 2833, AI-Report No. 15, Mai 1974, Cambridge, Mass.*
- COLLINSetal.75 *Collins, A. M. / Warnock, E. H. / Aiello, N. E. / Miller, M. L.: Reasoning from incomplete knowledge. BOBROW/COLLINS75, S. 383-415*
- DAMERAU75 *Damerau, F. J.: On fuzzy adjectives. IBM-Research Report RC5340, N.Y., März 1975*
- DAMERAU76 *Damerau, F. J.: Automated language processing. Williams, M. E. (ed.): Annual Review of Inf. Sc. and Technology, 11, 1976, S.107-161*
- DATE75 *Date, C. J.: An introduction to database systems. Reading: Addison-Wesley 1975*
- DAVIES73 *Davies, D. J. M.: POPLER 1.5 reference manual. Univ. of Edinburgh, TPU Report No. 1, Mai 1973*
- DAVIS/KING75 *Davis, R. / King, J.: An overview of production systems. Stanford Univ., AI-Lab., Memo-AM-271, Okt. 1975*
- DEHNING/MAASS77 *Dehning, W. /Maaß, S. : Kommunikative Aspekte der Mensch-Computer Interaktion. Univ. Hamburg, Inst. f. Informatik, IfI -HH-M-43/77 1977*
- DEMANT71 *Demant, B.: Fuzzy retrieval structures. Angewandte Informatik, 11, 1971, S. 500-502*
- DENOFOSKY76 *Denofsky, M. E.: How near is near?. MIT-AI-Lab., Memo 344, Feb. 1976*
- DIENES49 *Dienes, Z. P.: On an implication function in many-valued systems of logic. Journal of Symbolic Logic, 14, 1949, S. 95-97*
- DIETRICH/KLEIN74 *Dietrich, R. / Klein, W.: Computerlinguistik – eine Einführung. Stuttgart: Kohlhammer 1974*
- DILGER76 *Dilger, W.: Verbindungsgraph und Auswahlfunktion. Inst. f. Deutsche Sprache, Mannheim, Int. Papier, Aug. 1976*

- DREYFUSSetal,75 *Dreyfuss, G. R. / Kochen, M. / Robinson, J. / Badre, A. N. : On the psycholinguistic reality of fuzzy sets. Grossman, R. E. / San, L. J. / Vance, T. J. (eds.): Functionalism. Univ. of Chicago Press 1975*
- DUDAetal.76 *Duda, R. O. / Hart, P. E. / Nilsson, N. J.: Subjective bayesian method for rule-based inference systems. Stanford Res. Inst., Menlo Park, Techn. Note 124, Jan. 1976*
- EISENBERG76 *Eisenberg, P. : Oberflächenstruktur und logische Struktur. Untersuchungen zur Syntax und Semantik des Deutschen Prädikatadjektivs, Tübingen: Niemeyer 1976*
- EISENBERG76b *Eisenberg, P. (ed.): Maschinelle Sprachanalyse. Beiträge zur automatischen Sprachverarbeitung 1, Berlin: De Gruyter 1976*
- EISENBERG76c *Eisenberg, P. (ed.): Semantik und Künstliche Intelligenz. Beiträge zur automatischen Sprachverarbeitung 2, Berlin: De Gruyter 1976*
- ELCOCK/MICHIE77 *Elcock, E. W./ Michie, D. (eds.): Machine representations of knowledge. (= Machine Intelligence 8) N.Y.: Wiley 1977*
- EPSTEINetal.74 *Epstein, G. / Frieder, G. / Rine, D. C.: The development of multiple-valued logic as related to computer science. Computer, 7, 4, 1974, S. 20-32*
- FELDMAN75 *Feldman, J.: Bad-mouthing frames. NASH-WEBBER/SCHANK75, S. 102-103*
- FIKES75 *Fikes, R. E.: Deductive retrieval mechanisms for state description models. IJCAI4, S. 98-106*
- FILLMORE75 *Fillmore, C. J.: An alternative to checklist theories of meaning. Proc. of 1. annual meeting of the Berkeley Linguistic society, Berkeley 1975, S. 123-131*
- FINE75 *Fine, K.: Vagueness, truth and logic. Synthese, 30, 1975, S. 265-300*
- FISCHER/BÖCKER76 *Fischer, G. / Böcker, H.-D.: Modelle und Heuristiken- Wissen zur Einschränkung und Lenkung von Suchprozessen. LAUBSCH/SCHNEIDER76, S. 51-72*
- FOSTER77 *Foster, J. M.: Programming language design for the representation of knowledge. ELCOCK/MICHIE 77, S. 209-222*
- FRASER75 *Fraser, B.: Hedged performatives. Cole, P. / Morgan, J. L. (eds.): Syntax and semantics, 3, 1975, N.Y.: Academic, S. 187-210*
- FREEDLE/CAROLL72 *Freedle, O. / Carroll, J. B. (eds.): Language comprehension and the aquisition of knowledge. N.Y.: Wiley 1972*
- FREGE62 *Frege, G.: Funktion, Begriff, Bedeutung. (Patzig, G. (ed.)), Göttingen 1962*
- GAINES76 *Gaines, B. R.: Foundations of fuzzy reasoning. Int. Journal of Man-Machine Studies, 8, 1976, S. 623-668*
- GAINES/KOHOUT77 *Gaines, B. R. / Kohout, L.: The fuzzy decade: a bibliography of fuzzy systems and closely related topics. Int. Journal of Man-Machine Studies, 9, 1977, S. 1-68*
- GENRICH75 *Genrich, H. J.: Belästigung der Menschen durch Computer. Mühlbacher, J. (ed.): 5. GI-Jahrestagung, Berlin 1975, S. 94-106*
- GERSHMAN76 *Gershman, A.: Fuzzy set methods for understanding vague hints for maze running. Univ. of California, Los Angeles, Computer Sc. Dep., Report, 1976*

- GÖRZ/LAGEMANN76 *Görz, G. / Lagemann, H. H.: Bemerkungen zum Frame-Problem. LAUBSCH/SCHNEIDER76, S. 73-90*
- GOGUEN69 *Goguen, J. A.: The logic of inexact concepts. Synthese, 19, 1969, S. 325-373*
- GOGUEN73 *Goguen, J. A.: Axioms, extensions and applications for fuzzy sets: languages and the representation of concepts. IBM-Research, N.Y., Report RC4547, Sept. 1973*
- GOGUEN74 *Goguen, J. A.: Concept representation in natural and artificial languages: Axioms, extensions and applications for fuzzy sets. Int. Journal of Man-Machine Studies, 6, 1974, S. 513-561*
- GOGUEN75 *Goguen, J. A.: On fuzzy robot planning. ZADEHet al.75, S. 429-447*
- GOGUEN76 *Goguen, J. A.: Robust programming languages and the principle of maximal meaningfulness. Milwaukee Symposium on automatic computation and control, 1976, S. 87-90*
- GREEN63 *Green, C. C.: Application of theorem proving to problem solving. IJCAI1, 1969, S. 219-239*
- GREEN/RAPHAEL68 *Green, C. / Raphael, B.: The use of theorem-proving techniques in question-answering systems. Proc. ACM Nat. Conf., 1968, S. 169-181*
- GREGG74 *Gregg, L. W. : Knowledge and cognition. N.Y.:Wiley 1974*
- GRISHMAN76 *Grishman, R.: A survey of syntactic analysis procedures for natural language, Americ. Journal of Comp. Linguistics, Microfiche 47, 1976*
- v.HAHNetal.76 *v.Hahn, W. / Henskes, D. T. / Hoepfner, W./Wahlster, W.: HAM-RPM: ein Redepartnermodell als Simulationsprogramm. Braunmüller, K. /Kürschner, W. (eds.): Grammatik. Akten des 10. Linguistischen Kolloquiums. Tübingen 1975, Bd. 2, Tübingen: Niemeyer 1976, S. 337-357*
- HART76 *Hart, P. E.: Artificial intelligence. Stanford Research Inst., Menlo Park, Techn. Note 126, Feb. 1976*
- HARWOOD/HANNA76 *Harwood, W. T. / Hanna, F. K. : A distributed activity processing system for AI. AISB Summer Conf., Proc., Edinburgh 1976, S. 130-136*
- HAYES69 *Hayes, P. J.: Robotologic. Machine Intelligence, 5, 1969, S. 533-554*
- HAYES71 *Hayes, P. J.: A logic of actions. Machine Intelligence, 6, 1971, S. 495-520*
- HAYES74 *Hayes, P. J.: Some problems and non-problems in representation theory. AISB Summer Conf., Proc, Univ. of Sussex, 1974, S. 63-79*
- HAYES75 *Hayes, P. J.: Computer programming as a cognitive paradigm. Nature,254, April 1975, S. 563-566*
- HAYES77 *Hayes, P. J.: In defense of logic. erscheint in: Advance Papers, IJCAI5, MIT, 1977*
- V.HELD76 *v.Held, W.: Zur Beschreibung und Darstellung begrifflicher Komponenten von Ausdrucksbedeutungen. Braunmüller, K. / Kürschner, W. (eds.): Grammatik. Akten des 10. Linguistischen Kolloquiums. Tübingen 1975, Bd. 2, Tübingen: Niemeyer 1976, S. 167-176*
- HEMPEL39 *Hempel C.G.: Vagueness and logic. Philosophy of Science,6, 1939, S. 163-180*

- HEMPHILL75 *Hemphill, L. G.: A conceptual approach to automated language understanding and belief structures: with disambiguation of the word 'for'. Stanford Univ., Report Stan-CS-75-534, 1975*
- HENDRIX73 *Hendrix, G. G.: Modelling simultaneous actions and continous processes. Artificial Intelligence, 4, 1973, S. 145-180*
- HENDRIX76 *Hendrix, G. G.: The representation of semantic knowledge. Walker, D. E. (ed.): Speech understanding research. Stanford Research Inst., Final techn. Report, Okt. 1976*
- HERSH/CARAMAZZA76 *Hersch, H. M. / Caramazza, A.: A fuzzy set approach to modifiers and vagueness in natural language. Journal of Exp. Psychology: General, 105,3, 1976, S. 254-276*
- HEWITT71 *Hewitt, C.: Procedural embedding of knowledge in PLANNER. IJCAI2, 1971, S. 169-182*
- HEWITT75 *Hewitt, C.: How to use what you know. IJCAI4, 1975, S. 189-198*
- HOEPPNER77 *Hoepfner, W.: Derivative Wortbildung in automatischen Analysesystemen. Univ. Hamburg (erscheint demnächst)*
- HOWEetal.75 *Howe, J. A. M. / Knapman, J. / Noble, H.M. / Weir, S. / Young, R. M.: Artificial intelligence and the representation of knowledge. Univ. of Edinburgh, Dep. of AI, Research Report No. 5, 1975*
- HUGHES/CRESSWELL68 *Hughes, G. E. / Cresswell, M. J.: An introduction to modal logic. London: Methuen 1968*
- HUNT75 *Hunt, E. B.: Artificial intelligence. N.Y.: Academic 1975*
- IJCAI1 *Proc. of the first International Joint Conf. on Artificial Intelligence, Washington, D.C. 1969*
- IJCAI2 *Advance Papers of the second International Joint Conf. on Artificial Intelligence, London, 1971*
- IJCAI3 *Proc. of the third International Joint Conf. on Artificial Intelligence, Stanford Univ., 1973*
- IJCAI4 *Advance Papers of the fourth International Joint Conf. on Artificial Intelligence, Tiflis, 1975*
- ITZINGER76 *Itzinger, O.: Methoden der maschinellen Intelligenz. Eine Einführung. München: Hanser 1976*
- JACKSON74 *Jackson, P. C.: Introduction to artificial intelligence. N.Y.: Mason & Lipscomb 1974*
- KATZ67 *Katz, J. J.: Recent issues in semantic theory. Foundations of Language, 3, 1967, S. 124-194*
- KAUFMANN75 *Kaufmann, A.: Introduction to the theory of fuzzy subsets. Bd. 1, N.Y.: Academic 1975*
- KEENAN72 *Keenan, E. L.: On semantically based grammar. Linguistic Inquiry, 3, 4, 1972, S. 413-461*
- KING/HAYES76 *King, M. / Hayes, P.: An introduction to programming in LISP. CHARNIAK/WILKS76, S. 235-281*
- KLING73 *Kling, R.: Fuzzy Planner, Computing inexactness in a procedural problem-solving language. Wisconsin Univ., Madison, Report TR-168, Feb. 1973*
- KLOCKOW76 *Klockow, R.: Gänsefüßchen-Semantik. Eine Ergänzung zu Lakoffs 'Hedges'. Weber, H. / Weydt, H. (eds.): Sprachtheorie und Pragmatik. Akten des 10. Linguistischen Kolloquiums. Tübingen 1975, Bd. 1, Tübingen: Niemeyer 1976, S. 235-245*
- KNUTH68 *Knuth, B.: Semantics of context-free languages. Math. System Theory, 2, 1968, S. 127-145*

- KOCHEN75 *Kochen, M.: Applications of fuzzy sets in psychology. ZADEHetal.75, S. 395-408*
- KOCHEN/BADRE74 *Kochen, M. / Badre, A. N.: On the precision of adjectives which denote fuzzy sets. Journal of Cybernetics, 4, 1974, S. 49-59*
- KONRAD76 *Konrad, E.: Ein Graphenkalkül zur Darstellung und Inferenz von Wissen. LAUBSCH/SCHNEIDER76, S. 105-118*
- KOOIJ71 *Kooij, J. G.: Ambiguity in natural language. An investigation of certain problems in the linguistic description. Amsterdam: North-Holland 1971*
- KOWALSKI74 *Kowalski, R.: Predicate logic as a programming language. Proc. of IFIP, Aug 1974, S. 569-574*
- KOWALSKI75 *Kowalski, R.: A proof procedure using connection graphs. JACM, 22, 4, 1975, S. 572-595*
- KUIPERS75 *Kuipers, B.: A frame for frames: Representing knowledge for recognition. BOBROW/COLLINS75 S. 151-184*
- v.KUTSCHERA76 *Kutschera, F.: Einführung in die intensionale Semantik. Berlin: De Gruyter 1976*
- LAKOFF70 *Lakoff, G.: A note on vagueness and ambiguity. Linguistic Inquiry, 1, 1970, S. 357-359*
- LAKOFF73 *Lakoff, G.: Fuzzy grammar and the competence/performance terminology game. Papers from the 9th regional meeting of the Chicago Linguistic Society, April 1973, S. 271-291*
- LAKOFF75 *Lakoff, G.: Hedges: A study in meaning criteria and the logic of fuzzy concepts. Hockney, D. / Harper, W. / Freed, B. (eds.): Contemporary research in philosophical logic and linguistic semantics. Dordrecht: Reidel 1975, S. 221-272 (vorher in: Papers from the 8th regional meeting of the Chicago Linguistic Society 1972)*
- LAUBSCH75a *Laubsch, J. H.: Some thoughts about representing knowledge in instructional systems. IJCAI4, 1975, S. 122-125*
- LAUBSCH75b *Laubsch, J. H.: Organisation eines semantischen Gedächtnismodells mit Mikro-Planner. VEENKER 75b, S. 110-122*
- LAUBSCH/SCHNEIDER76 *Laubsch, J. H. / Schneider, H.-J. (eds.): Dialoge in natürlicher Sprache und Darstellung von Wissen. Workshop der GI-Fachgruppe Künstliche Intelligenz, Freudenstadt 1976*
- LEARNING RESEARCH GROUP76 *Learning Research Group: Personal dynamic media. Xerox Palo Alto Research Center, Report SSL 76-1, 1976*
- LEAVENWORTH/SAMMET74 *Leavenworth, B. / Sammet, J. E.: An overview of nonprocedural languages. IBM-Research, Report RC 4685, N.Y. 1974*
- LEE72 *Lee, R. C. T.: Fuzzy logic and the resolution principle. JACM, 19, 1972, S. 109-119*
- LeFAIVRE74 *LeFaivre, R. A.: Fuzzy problem-solving. Univ. of Wisconsin, Madison, Techn. Report No. 37, Sept. 1974*
- LeFAIVRE75 *LeFaivre, R. A.: The representation of fuzzy knowledge. Journal of Cybernetics, 4, 2, 1975, S. 57-66*
- LeFAIVRE76 *LeFaivre, R. A.: Procedural representation in a fuzzy problem-solving system. Proc. Nat. Computer Conf., 1976, S. 1069-1074*
- LeFAIVRE77a *LeFaivre, R. A.: FUZZY reference manual. Rutgers Univ., Computer Sc. Dep., März 1977*

- LeFAIVRE77b *LeFavre, R. A.: Fuzzy representation and approximate reasoning. Rutgers Univ., Computer Sc. Dep., Report RUCBM-TR-78, 1977*
- LEISI53 *Leisi, E.: Der Wortinhalt. Seine Struktur im Deutschen und Englischen. Heidelberg 1953*
- LENDERS75 *Lenders, W.: Semantische und argumentative Text-deskription. Ein Beitrag zur Simulation sprachlicher Kommunikation. Hamburg: Buske 1975*
- LEWIS74 *Lewis, D.: Prinzipien der Semantik. Kanngiesser, S. / Lingrün, G. (eds.): Studien zur Semantik. Kronberg/Ts.: Scriptor 1974, S. 136-197*
- LISKOV76 *Liskov, B. : An introduction to CLU. MIT, Lab. for Computer Sc., CGS Memo 136, 1976*
- MACVICAR WHELAN74 *MacVicar Whelan, P. J.: Fuzzy sets, the concept of height, and the hedge 'very'. Grand Valley State Colleges, Allendale, Michigan, Physics Dep., Techn. Note 1, 1974*
- MACVICAR WHELAN76 *MacVicar Whelan, P. J.: Fuzzy sets for man-machine interaction. Intern. Journal of Man-Machine Studies,8, 1976, S. 687-697*
- MALHOTRA75 *Malhotra, A.: Design criteria for a knowledge-based English language system for management: An experimental analysis. MIT, Proj. MAC, Techn. Report TR-146, Feb. 1975*
- MALHOTRA/WLADAWSKY75 *Malhotra, A. / Wladawsky, I.: The utility of natural language systems. IBM-Research Report RC5739, N.Y.,1975*
- MANN75 *Mann, W. C: Why things are so bad for the computer-naive user. Southern California Univ., Marina del Rey, Report ISI/RR-75-32, 1975*
- MARCOTTYetal.76 *Marcotty, M. / Ledgard, H. F. / Bochmann, G. V.: A sampler of formal definitions. Computing Surveys,8,2, 1976, S. 191-276*
- McCARTHY58 *McCarthy, J.: Programs with common sense. Proc. Teddington Conf. on Mechanisation of Thought Processes, Bd. 1, London 1958, S. 77-84*
- McCARTHY/HAYES63 *McCarthy, J. / Hayes, P. J.: Some philosophical problems from the standpoint of artificial intelligence. Machine Intelligence, 4, 1969, S. 463-502*
- McDERMOTT75 *McDermott, D. V.: Very large Planner-type data bases. MIT, AI-Lab., Report No.339, 1975*
- McDERMOTT76 *McDermott, D. V.: Artificial intelligence meets natural stupidity. SIGART Newsletter, No. 57, April 1976, S. 4-9*
- McDERMOTT/SUSSMAN72 *McDermott, D. V. / Sussman, G. J.: The CONNIVER reference manual. MIT, AI-Lab., Memo No. 259a, 1972*
- McDERMOTTetal.76 *McDermott, J. / Newell, A. / Moore, J.: The efficiency of certain production system implementations. Carnegie-Mellon Univ., Report, Sept. 1976*
- MELLI74 *Melli, L. F.: The 2. PAK language. Primitives for AI applications. Univ. of Toronto, Dep. of Computer Sc., Techn. Report No. 73, Dez. 1974*
- MICHALSKI77 *Michalski, R. S.: Variable-valued logic and ist application to pattern recognition and machine learning. Rine, D. G. (ed.): Computer science and multiple-valued logic. Amsterdam: North-Holland 1977, S. 506-534*
- MICHIE74 *Michie, D.: On machine intelligence. Edinburgh Univ. Press 1974*

- MILLER/BECKER74 *Miller, L. A. /Becker, C. A.: Programming in natural language. IBM-Research, Report RC5137, N.Y., Nov. 1974*
- MINSKY68 *Minsky, M. (ed.): Semantic information processing. Cambridge: MIT Press 1968*
- MINSKY75 *Minsky, M.: A framework for representing knowledge. Winston, P.: The psychology of computer vision. N.Y.: McGraw-Hill 1975, S. 211-277*
- MITTELSTEINetal.76 *Mittelstein, M. / Nebel, B. / Pretschner, B. / Scheffe, P.: HASY- Ein Programm zur syntaktischen Analyse natürlicher Sprachen. Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-10/76, 1976*
- MOORE66 *Moore, R. E.: Interval analysis. Englewood Cliffs: Prentice-Hall 1966*
- MOORE/NEWELL74 *Moore, J. / Newell, A.: How can Merlin understand? GREGG74, 19 74*
- MORGAN76 *Morgan, C. G.: Methods for automated theorem-proving in non-classical logics. IEEE Trans. Comp., C-25,8, 1976, S. 852-862*
- MOTSCH64 *Motsch, W. : Syntax des Deutschen Adjektivs. Berlin: Akademie 1964*
- MÜLLER77 *Müller, B. : Zum Problem von Frage-Antwort-Systemen im Rechtswesen. Gesellschaft f. Mathematik und Datenverarbeitung, Birlinghoven, Entwicklungsunterlage FRAGANT 11, 1977*
- MUNSON71 *Munson, J.: Robot planning, execution and monitoring in an uncertain environment. IJCAI2, 1971 , S. 338-349*
- MYLOPOULOSetal.75 *Mylopoulos, J. / Borgida, A. / Cohen, P. / Roussopolous, N. / Tsotsos, J. / Wong, H.: TORUS- A natural language understanding system for data management, IJCAI4, 1975, S. 414-421*
- NASH-WEBBER/BRUCE76 *Nash-Webber, B. / Bruce, B.: Evolving uses of knowledge in speech understanding systems. Preprints of the Comp. Linguistics Conf. Ottawa, 1976*
- NASH-WEBBER/SCHANK75 *Nash-Webber, B. / Schank, R. (eds.): Theoretical issues in natural language processing. Proc. of an interdisciplinary workshop, Cambridge, Mass., 1975*
- NEGOITA/RALESCU75 *Negoita, C. V. / Ralescu, D. A.: Applications of fuzzy sets to systems analysis. Basel: Birkhäuser 1975*
- NEUHAUS71 *Neuhaus, J. H.: Beschränkungen in der Grammatik der Wortableitungen im Englischen. Diss. Saarbrücken 1971*
- NEWELL70 *Newell, A.: Remarks on the relationship between artificial intelligence and cognitive psychology. Banerji, R. / Mesarovic, M. (eds.): Theoretical approaches to non-numerical problem-solving. N.Y. : Springer 1970, S. 363-400*
- NEWELL73 *Newell, A.: Production systems: Models of control structures. Chase, W. (ed.): Visual information processing. N.Y.: Academic 1973, S. 463-526*
- NEWELL/SIMON72 *Newell, A. / Simon, H. A.: Human problem solving. Englewood Cliffs: Prentice Hall 1972*
- NEWELL/SIMON76 *Newell, A. / Simon, H. A.: Computer science as empirical inquiry: Symbols and search. CACM, 19,3, März 1976, S. 113-126*
- NILSSON71 *Nilsson, N.: Problem-solving methods in artificial intelligence. N.Y.: McGraw-Hill 1971*

- NILSSON74 *Nilsson, N.: Artificial intelligence. Inform. Processing 74, Conf. Proc., Amsterdam: North-Holland 1974, S. 778-801*
- NILSSON75 *Nilsson, N.: Problem solving: Diagnosis of equipment. Nilsson, N. (ed.): Artificial intelligence - Research and applications. Stanford Research Inst., Menlo Park, 1975, S. 219-227*
- NOBLE75 *Noble, H.: SNAP - a program description. HOWE et al.75, S. 28-40*
- NORMAN/BOBROW76 *Norman, D. A. / Bobrow, D. G.: On the role of active memory processes in perception and cognition. COFER76, S. 114-131*
- NORMAN/RUMELHART75 *Norman, D. A. / Rumelhart, D. E.: Explorations in cognition. San Francisco: Freeman 1975*
- OSGOODetal.57 *Osgood, C. E. / Suci, G. J. / Tannenbaum, P. H.: The measurement of meaning. Urbana: Univ. of Illinois Press 1957*
- PEIRCE02 *Peirce, C. S.: Vagueness. Baldwin, J. M. (ed.): Dictionary of philosophy and psychology, Bd. 2, N.Y.: Macmillan 1902*
- PETÖFI76 *Petöfi, J. S.: A frame for frames. Second annual meeting of the Berkeley Linguistic Society, 1976, S. 319-329*
- POLYA75 *Pólya, G.: Mathematik und plausible Schliessen. Bd. 2: Typen und Strukturen plausibler Folgerung. Basel: Birkhauser 1975²*
- POPPER76 *Popper, K. R.: Unended quest. London: Fontana 1976*
- POST43 *Post, E.: Formal reduction of the general combinatorial decision problem. Americ. Journal Math., 65, 1943, S. 197-215*
- QUILLIAN68 *Quillian, M. R.: Semantic memory. MINSKY68, S. 227-270*
- RAPHAEL68 *Raphael, B.: SIR: A computer program for semantic information retrieval. MINSKY68, S. 33-145*
- RAPHAEL76 *Raphael, B.: The thinking computer: Mind inside matter. San Francisco: Freeman 1976*
- RAPHAEL/ROBINSON72 *Raphael, B. / Robinson, A. E. : Bibliography on computer semantics. Stanford Research Inst., AI-Center, Techn. Note 72, 1972*
- REDDY73 *Reddy, D. R.: Eyes and ears for computers. Einsele, Th. / Giloi, W. / Nagel, H.-H. (eds.): NTG/GI Fachtagung 'Cognitive Verfahren und Systeme', (Lecture Notes in Economics and mathematical systems, Bd. 83) Berlin, Heidelberg, N.Y.: Springer 1973, S. 1-28*
- REDDY76 *Reddy, D. R.: Speech recognition by machine: A review. Proc. IEEE, 64, 4, April 1976, S. 501-531*
- REITER75 *Reiter, R.: Formal reasoning and language understanding. NASH-WEBBER/SCHANK75, S. 194-198*
- RESCHER69 *Rescher, N.: Many-valued logic. N.Y.: McGraw-Hill 1969*
- RHODES/KLINGER77 *Rhodes, M. L. / Klinger, A.: Conversational text input for modifying facial images. SIGART Newsletter, 61, Febr. 1977, S. 58-61*
- RIEGER74 *Rieger, B.: Eine 'tolerante' Lexikonstruktur. Zur Abbildung natürlchsprachlicher Bedeutung auf 'unscharfe' Mengen in Toleranzräumen. Zeitschrift für Literaturwissenschaft und Linguistik, 16, 1974, S. 31-47*
- RIEGER76a *Rieger, B.: Theorie der unscharfen Mengen und empirische Textanalyse. Inst. f. mathematisch-empirische Systemforschung, Aachen, 1976, Bericht*

- RIEGER76b *Rieger, B.: Fuzzy structural semantics. On a generative model of vague natural meaning. Inst. f. mathematisch-empirische Systemforschung, Aachen, 1976, Bericht*
- ROBINSON65 *Robinson, J. A.: A machine-oriented logic based on the resolution principle. JACM,12,1, Jan. 1965, S. 23-41*
- ROSCH73 *Rosch, E.: Natural categories. Cognitive Psychology, 4, 1973, S. 328-350*
- R0SCH75a *Rosch, E.: Cognitive representations of semantic categories. Journal of Exper. Psychology: General, 104, 1975, S. 192-233*
- ROSCH75b *Rosch, E.: Cognitive reference points. Cognitive psychology, 7, 1975, S. 573-605*
- R0SCH/MERVIS75 *Rosch, E. / Mervis, C.: Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 1975, S. 573-605*
- ROSENBERG75 *Rosenberg, R. S.: Artificial intelligence and linguistics: A brief history of a one-way relationship. Proc. of the first annual meeting of the Berkeley Linguistic Society, 1975, S. 379-392*
- ROSENFELD75 *Rosenfeld, A.: Fuzzy graphs. ZADEHetal.75, S. 77-95*
- ROSS70 *Ross, J. R.: A note on implicit comparatives. Linguistic Inquiry, 1, 1970, S. 336-363*
- RULIFSONetal.72 *Rulifson, J. F. / Derksen, J. A. / Waldinger, R. J.: QA4: A procedural calculus for intuitive reasoning. Stanford Research Inst., Techn. Note 73, Nov. 1972*
- RUMELHART/NORMAN73 *Rumelhart, D. / Norman, D.: Active semantic networks as a model of human memory. IJCAI3, 1973, S. 450-457*
- RUMELHARTetal.72 *Rumelhart, D. / Lindsay, P. H. / Norman, D. A.: A process model for long-term memory. Tulving, E. / Donaldson, W. (eds.): Organization of memory. N.Y.: Academic 1972, S. 197-246*
- RUSSELL23 *Russell, B.: Vagueness. Australian Journal of Psychology and Philosophy, 1, 1923, S. 84-92*
- RUSSELL76 *Russell, S.: Computer understanding of metaphorically used verbs. Americ. Journal of Comp. Linguistics, Microfiche 44, 1976*
- RUSTIN73 *Rustin, R. (ed.): Natural language processing. Englewood Cliffs: Prentice Hall 1973*
- SACERDOTIetal.76 *Sacerdoti, E. D. / Fikes, R. E. / Reboh, R. / Sagalowicz, D. / Waldinger, R. J. / Wilber, M. B.: QLISP: A language for the interactive development of complex systems. Stanford Research Inst., AI-Center, Techn. Note 120, März 1976*
- SAGERetal.77 *Sager, N. / Hirshman, L. / Grishman, R. / Insolio, C.: Transforming medical records into a structured data base. SIGART Newsletter,61, Feb. 1977, S. 38-39*
- SANDEWALL71 *Sandewall, E.: Formal methods in the design of question-answering systems. Artificial Intelligence, 2, 1971, S. 129-145*
- SANDEWALL73 *Sandewall, E.: Conversion of predicate-calculus axioms, viewed as non-deterministic programs, to corresponding deterministic programs. IJCAI3, 1973, S. 230-234*
- SANDEWALL75 *Sandewall, E.: Ideas about management of LISP data bases. IJCAI4, 1975, S. 585-592*
- SANDEWALL77 *Sandewall, E.: Some observations on conceptual programming. ELCOCK/MICHIE77, S. 223-265*

- SCHAFF68 Schaff, A.: *Unschärfe Ausdrücke und die Grenzen ihrer Präzisierung*. Schaff, A.: *Essays über die Philosophie der Sprache*. Frankfurt: Europäische Verlagsanstalt 1968, S. 65-94
- SCHANK75 Schank, R. C. (ed.): *Conceptual information processing*. Amsterdam: North-Holland 1975
- SCHANK/COLBY73 Schank, B. C. / Colby, K. M. (eds.): *Computer models of thought and language*. San Francisco: Freeman 1973
- SCHANKetal.75 Schank, R. C. et al.: *SAM- A story understander*. Yale Univ., New Haven, Report 43, Aug. 1975
- SCHEFE73 Schefe, P.: *Simulation kognitiver Prozesse - Wissenschaftstheoretische Aspekte*. Univ. Hamburg, Inst. f. Informatik, Vervielf. 1973
- SCHEFE76 Schefe, P.: *Ein dependentiell-transformationeller Parser für den Mensch-Maschinen-Dialog in natürlicher Sprache*. LAUBSCH/SCHNEIDER76, S. 175-191
- SCHEFE77 Schefe, P.: *Technische und nicht-technische Aspekte der Mensch-Maschine-Kommunikation*. BRUNNSTEIN77, 10.1.1-10.1.12
- SCHMIDT74 Schmidt, C.: *The relevance to semantic theory of a study of vagueness*. Papers from the 10th regional meeting of the Chicago Linguistic Society, April 1974, S. 617-630
- SCHOTSCH75 Schotsch, P. K.: *Fuzzy modal logic*. Proc. of 1975 Inter. Sympos. on multiple-valued logic, Bloomington, Indiana, S. 176-182
- SCHWEPPE75 Schweppe, H.: *Beziehungen zwischen Künstlicher Intelligenz und Data Base Management*. VEENKER 75b, S. 61-78
- SCHWIND75 Schwind, C.: *Generating hierarchical semantic networks from natural language discourse*. IJCAI4, 1975, S. 429-434
- SCHWIND76 Schwind, C.: *Darstellung von natürlichsprachlichen Texten durch zustandslogische Ausdrücke und Modelle*. LAUBSCH/SCHNEIDER76, S. 192-203
- SCHUBERT75 Schubert, C. K.: *Extending the expressive power of semantic networks*. IJCAI4, 1975, S. 158-164
- SCOTTetal.77 Scott, C. A. / Clancey, A. / Davis, R. / Shortliffe, E. H.: *Explanation capabilities of production-based consultation systems*. Americ. Journal of Comp. Linguistics, Microfiche 62, 1977
- SCRAGG76 Scragg, G. W.: *Semantic nets as memory models*. CHARNIAK/WILKS76, S. 101-127
- SGALL75 Sgall, P. J.: *Linguistics and artificial intelligence*. Prague Bull. Math. Linguistics, 24, 1975, S. 5-33
- SHAKET76 Shaket, E.: *Fuzzy semantics for a natural-like language defined over a world of blocks*. Univ. of California, Los Angeles, AI-Memo No. 4, 1976
- SHORTLIFFE74 Shortliffe, E. H.: *MYCIN: A rule-based computer program for advising physicians regarding antimicrobial therapy selection*. Stanford Univ., AI-Lab., Memo 251, STAN-CS-74-465, 1974
- SHORTLIFFE/BUCHANAN75 Shortliffe, E. H. / Buchanan, B. G.: *A model of inexact reasoning in medicine*. Math. Biosc., 23, 1975

- SHORTLIFFEetal.75 *Shortliffe, E. H. / Davis, R. / Anline, S. G. / Buchanan, B. G./ Green, C. C. / Cohen, S. N.: Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. Computers and Biomedical Research, 8, 1975, S. 303-320*
- SIKLOSSY76 *Siklóssy, L.: Let's talk LISP. Englewood Cliffs: Prentice-Hall 1976*
- SIMMONS72 *Simmons, R. F.: Some semantic structures for representing English meanings. FREEDLE/CAROLL 72, S. 71-97*
- SIMMONS73 *Simmons, R. F.: Semantic networks: their computation and use for understanding English sentences. SCHANK/ COLBY73, S. 63-113*
- SIMON73 *Simon, H. A.: The structure of ill structured problems. Artificial Intelligence, 4, 1973, S. 181-201*
- SIMON/SIKLOSSY72 *Simon, H. A. / Siklóssy, L. (eds.): Representation and meaning. Experiments with information processing systems. Englewood Cliffs: Prentice-Hall 1972*
- SINOWJEW68 *Sinowjew, A. A.: Über mehrwertige Logik. Berlin: Akademie 1968*
- SKUCE75 *Skuce, D.: An english-like language for qualitative scientific knowledge. IJCAI4, 1975, S. 593-599*
- SLOMAN71 *Slovan, A.: Interactions between philosophy and artificial intelligence: the role of intuition and non-logical reasoning in intelligence. Artificial Intelligence, 2, 1971, S. 209-225*
- SONDHEIMER76 *Sondheim, N. K.: Spatial reference and natural-language machine control. Intern. Journal of Man-Machine Studies,8, 1976, S. 329-336*
- SRIDHARAN75 *Sridharan, D. (ed.): Proceedings of the first annual artificial intelligence in medicine workshop. Rutgers Univ., New Brunswick, 1975*
- SRIDHARAN76 *Sridharan, N. S.: The frame and focus problems in AI: Discussion in relation to the Believer system. AISB Summer Conf, Proc., Edinburgh, Juli 1976, S. 322-333*
- STAMMERJOHANN75 *Stammerjohann, H. (ed.): Handbuch der Linguistik. Allgemeine und Angewandte Sprachwissenschaft. München: Nymphenburger 1975*
- STEELS76 *Steels, L.: Introduction to natural language processing. Advances in natural language processing, Preprints of a workshop held at the Univ. of Antwerp, Okt. 1976*
- SUSSMANetal.70 *Sussman, G. J. / Winograd, T. / Charniak, E.: Micro-PLANNER reference manual. MIT, AI-Lab., Memo 203a, 1970*
- SUTHERLAND76 *Sutherland, G.: Implementation of inference nets - II. Stanford Research Inst., AI-Center, Techn. Note 122, Jan. 1976*
- UHR75 *Uhr, L.: Toward integrated cognitive systems, which must make fuzzy decisions about fuzzy problems. ZADEHetal.75, S. 353-393*
- VEENKER75a *Veenker, G. (ed.): Künstliche Intelligenz-Forschung in der BRD. Informatik-Berichte der Univ. Bonn, Nr. 5, Bonn, April 1975*
- VEENKER75b *Veenker, G. (ed.): Zweites Treffen der GI-Fachgruppe Künstliche Intelligenz. Univ. Dortmund, Abt.: Informatik, Bericht Nr. 13*

- VENNEMANN73 Vennemann, T.: *Explanation in syntax*. Kimball, J.P. (ed.): *Syntax and semantics*, Bd. 2, N.Y. : Seminar Press 1973, S. 1-50
- WAHLSTER75 Wahlster, W. : *Pattern Matching und die Transformation von S-Expressions*. Wittig, T. (ed.): *LISP 1.6 II. Eine Erweiterung der Stanford-Version*. Univ. Hamburg, Inst. f. Informatik, IfI-HH-M-19/75, 1975, S. 7-11
- WAHLSTER77 Wahlster, W.: *HAM-RPM - A knowledge-based conversationalist*. *SIGART Newsletter*, 61, Feb. 1977, S. 36-37
- WAHLSTER/v.HAHN76 Wahlster, W. / v.HAHN, W.: *Einige Erweiterungen des natürlichsprachlichen AI-Systems HAM-RPM*. LAUBSCH/SCHNEIDER76, S. 204-225
- WALKER73 Walker, D. E.: *Automated language processing*. Cuadra, C. A. (ed.): *Annual review of information science and technology*, 8, 1973, S. 69-119
- WALTZ77 Waltz, D. L.: *Natural language interfaces*. *SIGART Newsletter*, 61, Feb. 1977, S.16-65
- WATERMAN77 Waterman, D. A.: *An introduction to production systems*. *AISB Newsletter*, 25, Jan. 1977, S. 7-10
- WECHSLER75 Wechsler, H.: *Applications of fuzzy logic to medical diagnosis*. *Proc. of the 1975 Intern. Sympos. on multiple-valued logic*. Indiana Univ., Bloomington, 1975, S. 162-174
- WERTHEIMER38 Wertheimer, M.: *Numbers and numerical concepts in primitive peoples*. Ellis, W. D. (ed.): *A source book of gestalt psychology*. N.Y.: Harcourt 1938
- WETTLER74 Wettler, M.: *Über die Struktur des semantischen Langzeitgedächtnisses*. Istituto per gli studi semantici e cognitivi, Working Paper No. 10, Castagnola Schweiz, 1974
- WILBER76 Wilber, M. B.: *A QLISP reference manual*. Stanford Research Inst. AI-Center, Techn. Note 118, März 1976
- WILKS73 Wilks, Y.: *An artificial intelligence approach to machine translation*. SCHANK/COLBY73, S. 114-151
- WILKS74 Wilks, Y.: *Natural language understanding systems within the AI paradigm: A survey and some comparisons*. Stanford Univ., AI-Lab., Memo 237, 1974
- WILKS75 Wilks, Y.: *Preference semantics*. Keenan, E. L. (ed.): *Formal semantics of natural language*. Cambridge: Univ. Press 1975, S. 320-348
- WILKS76 Wilks, Y.: *De minimis, or the archaeology of frames*. *AISB Summer Conf., Proc., Edinburgh, Juli 1976*, S. 355-365
- WILKS77 Wilks, Y.: *Methodological questions about artificial intelligence: approaches to understanding natural language*. *Journal of Pragmatics*, 1, 1977, S. 69-84
- WINOGRAD72 Winograd, T.: *Understanding natural language*. Edinburgh: Univ. Press 1972
- WINOGRAD73 Winograd, T.: *A procedural model of language understanding*. SCHANK/COLBY73, S. 152-186
- WINOGRAD74 Winograd, T.: *Five lectures on artificial intelligence*. Stanford Univ., AI-Lab., STAN-CS-74-459, Sept. 1974
- WINOGRAD75a Winograd, T.: *Breaking the complexity barrier again*. Nance, R. E. (ed.): *Programming languages-information retrieval*. *Proc. of ACM SIGPLAN-SIGIR interface meeting, 1973*, *SIGPLAN notes*, 10,1, Jan. 1975, S. 13-22

- WINOGRAD75b Winograd,T.: *Frame representations and the declarative/procedural controversy*. BOBROW/ COLLINS75, S. 185-210
- WINOGRAD76 Winograd, T.: *Computer memories: a metaphor for memory organization*. COFER76, S. 131-161
- WINSTON75 Winston, P. H.: *Learning structural descriptions from examples*. Winston, P. H. (ed.): *The psychology of computer vision*. N.Y.: McGraw-Hill 1975, S. 157-209
- WINSTON77 Winston, P. H.: *Artificial intelligence*. Reading: Addison-Wesley 1977
- WIRTH75 Wirth, N.: *Algorithmen und Datenstrukturen*. Stuttgart: Teubner 1975
- WITTGENSTEIN64 Wittgenstein, L.: *Tractatus logico-philosophicus*. Frankfurt: Suhrkamp 1964 (Original 1922)
- WITTGENSTEIN71 Wittgenstein, L.: *Philosophische Untersuchungen*. Frankfurt: Suhrkamp (Original: 1953)
- WITTIG75 Wittig,T.: *Semantische Analyse von Sätzen zur Erfassung eines Sachverhaltes*. Univ. Hamburg, Inst. f. Informatik, IfI-HH-B-24/75, 1975
- WITTIG76 Wittig, T.: *Auflösen von Mehrdeutigkeiten in einfachen semantischen Netzen*. LAUBSCH/ SCHNEIDER76, S. 244-257
- WOLF75 Wolf, R. G.: *A critical survey of many-valued logics 1966-1974*. Proc. of the 1975 Intern. Symp. on multiple-valued logic, Indiana Univ., Bloomington 1975, S. 468-474
- WOODS75 Woods, W. A.: *What's in a link*. Foundations for semantic networks. BOBROW/COLLINS75, S. 35-82
- WOODS77 Woods, W. A.: *A personal view of natural language understanding*. SIGART Newsletter,61, Feb. 1977, S. 17-20
- WOODSetal.72 Woods, W. / Kaplan, R. M. / Nash-Webber , B. : *The lunar sciences natural language information system. Final report*. BBN-Report 2378, Cambridge,Mass. 1972
- WUNDERLICH73 Wunderlich, D.: *Vergleichssätze*. Kiefer, F. / Ruwet, N. (eds.): *Generative grammar in Europe*. Dordrecht 1973, S. 629-672
- WUNDERLICH74 Wunderlich, D.: *Grundlagen der Linguistik*. Reinbeck: Rowohlt 1974
- ZADEH65 Zadeh, L. A.: *Fuzzy sets*. Information and Control, 8, 1965, S. 338-353
- ZADEH71 Zadeh, L. A.: *Quantitative fuzzy semantics*. Information Sc., 3, 1971, S. 159-176
- ZADEH72 Zadeh, L. A.: *A fuzzy-set-theoretic interpretation of linguistic hedges*. Journal of Cybernetics,2, 1972, S. 2-34
- ZADEH73 Zadeh, L. A.: *Outline of a new approach to the analysis of complex systems and decision processes*. IEEE Transac. on Systems, Man, and Cybernetics, SMC-3, 1, Jan. 1973, S. 28-44
- ZADEH74 Zadeh, L. A.: *A new approach to system analysis*. Marois, M. (ed.): *Man and computer*. Amsterdam: North-Holland 1974, S. 55-94
- ZADEH75 Zadeh, L. A.: *The concept of a linguistic variable and its application to approximate reasoning*. Information Sc., 8, 3, 1975, Teil 1, S. 199-249, 8, 4, 1975, Teil 2, S. 301-357, 9, 1, 1975, Teil 3, S. 43-30

- ZADEH76 *Zadeh, L. A.: A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. Intern. Journal of Man-Machine Studies, 8, 1976, S. 249-291*
- ZADEHetal.75 *Zadeh, L. A. / Fu, K. S. / Tanaka, K. / Shimura, M.: Fuzzy sets and their applications to cognitive and decision problems. N.Y.: Academic 1975*
- ZIFONUN75 *Zifonun, G.: ISLIB-Börse - eine Konstruktsprache (KS) zur formalen Repräsentation natürlichsprachlicher Information. Inst. f. Deutsche Sprache, Mannheim, Abt.: LDV, ISLIB-Info-I-3, Jan. 1975*