

NATÜRLICHSPRACHLICHE SYSTEME

EINE EINFÜHRUNG IN DIE SPRACHORIENTIERTE KI-FORSCHUNG¹

Wolfgang Wahlster
Forschungsstelle für Informationswissenschaft
und Künstliche Intelligenz
Universität Hamburg
Mittelweg 179
2000 Hamburg 13

VORWORT

Bei dem vorliegenden Aufsatz handelt es sich um eine einführende Darstellung wichtiger Fragestellungen, Methoden und Resultate des Forschungsgebietes 'Natürlichsprachliche Systeme'. Die Komplexität und die Breite des zu behandelnden Teilgebietes der Künstlichen Intelligenz (KI) machen es unmöglich, in einer kurzen Einführung auch nur annähernd alle für die sprachorientierte KI-Forschung relevanten Fragestellungen und Lösungsalternativen aufzuzeigen (z.B. umfaßt alleine eine gründliche Einführung in die syntaktische Analyse, wie sie Winograd 1983 bietet, rd. 640 Seiten). Dies mag auch als Grund dafür angesehen werden, daß Standardeinführungen in die Künstliche Intelligenz wie Winston 1977 und Nilsson 1981 das Teilgebiet 'Natürlichsprachliche Systeme' nicht in angemessener Weise behandeln.

In der vorliegenden Einführung werden nach einem Überblick zur Zielsetzung, Konzeption und historischen Entwicklung der sprachorientierten KI-Forschung nacheinander verschiedene Verfahren zur Realisierung der Verarbeitungsabschnitte 'Analyse', 'Auswertung', und 'Generierung' von natürlichsprachlichen Systemen erläutert, wobei, dem derzeitigen Forschungsstand entsprechend, der Schwerpunkt auf der Analyse natürlichsprachlicher Eingaben liegt. Oftmals mußten dabei extreme Vereinfachungen in der Darstellung und Verkürzungen in der Diskussion von Vor- und Nachteilen der vorgestellten Lösungsansätze vorgenommen werden, um dem Ziel einer kompakten, aber dennoch breit angelegten Einführung gerecht werden zu können. Es wurde versucht, den Leser in einigen Teilbereichen bis an die aktuellen Forschungsschwerpunkte heranzuführen. Auf eine Erläuterung von Verfahren, die zwar Meilensteine in der Entwicklung natürlichsprachlicher Systeme darstellen, aber heute nur noch von historischem Interesse sind, wurde weitgehend verzichtet. Bei der Selektion von Literaturhinweisen wurden leicht zugängliche Veröffentlichungen gegenüber der vielfach aktuelleren, aber für den Anfänger schlecht zugänglichen 'grauen' Literatur bevorzugt. Außerdem wurde bei der Auswahl von Beispielen deutschsprachigen Systemen der Vorrang gegeben. Obwohl versucht wurde, die Einführung so zu gestalten, daß möglichst wenig fachspezifische Vorkenntnisse vorausgesetzt werden müssen, ist es für eine sinnvolle Lektüre des vorliegenden Aufsatzes doch notwendig, daß der Leser über Grundkenntnisse in den Bereichen LISP, Prädikatenlogik, Repräsentation von Wissen sowie Automatentheorie und formale Sprachen verfügt.

Die Zielgruppe, an die sich diese Einführung wendet, besteht in erster Linie aus Informatikern, die sich in das Fachgebiet 'Künstliche Intelligenz' einarbeiten und dabei auch den Teilbereich 'Natürlichsprachliche Systeme' kennenlernen möchten. Da dem Informatiker in der Regel viele der in der sprachorientierten KI-Forschung gebräuchlichen sprachwissenschaftlichen Fachbegriffe unbekannt sein dürften, wurde versucht, die zum Verständnis der einschlägigen Literatur notwendigen linguistischen Grundbegriffe im Rahmen dieser Einführung zu vermitteln. Außerdem wendet sich diese Einführung auch an Linguisten und Psychologen, die Grundkenntnisse in Informatik haben und sich auch aus dem Blickwinkel der Künstlichen Intelligenz mit Problemen der natürlichsprachlichen Kommunikation beschäftigen möchten.

¹ Dieser Aufsatz entstand teilweise im Rahmen des Projektes HAM-ANS, das aus Mitteln des BMFT gefördert wird.

1. ZIELSETZUNG, KONZEPTION UND ENTWICKLUNG DER SPRACH-ORIENTIERTEN KI-FORSCHUNG

1.1. ZUR ZIELSETZUNG DES FORSCHUNGSGEBIETES

Innerhalb des Informatik-Fachgebietes Künstliche Intelligenz (Abk.: KI) stellt die sprachorientierte KI-Forschung, die sich mit dem Entwurf, der Implementierung und praktischen Erprobung natürlichsprachlicher Systeme sowie der Erforschung der damit verbundenen theoretischen Grundlagen beschäftigt, ein etabliertes und weiterhin expandierendes Teilgebiet dar, dessen Nutzen für hochentwickelte Industriegesellschaften unbestreitbar ist.

Die sprachorientierte KI-Forschung hat zwei *Hauptaufgaben*:

- (A₁) Die komplexen Informationsverarbeitungsprozesse, die dem Verstehen, der Produktion und dem Erwerb natürlicher Sprache zugrundeliegen, sollen mit informatischen Mitteln exakt beschrieben und erklärt werden.
- (A₂) An intelligentes Sprachverhalten gebundene Leistungen sollen maschinell verfügbar gemacht werden, und die Mensch-Maschine-Kommunikation soll durch die Entwicklung natürlichsprachlicher Systeme besser dem Menschen angepaßt werden.

In der sprachorientierten KI-Forschung wird das in (A₁) formulierte Erkenntnisinteresse wesentlich durch die mit (A₂) angesprochene ingenieurwissenschaftliche Zielsetzung der Konstruktion von Anwendungssystemen bestimmt. Die wichtigsten *gesellschaftlichen Auswirkungen* der mit (A₂) angestrebten Mensch-Maschine-Kommunikation in natürlicher Sprache sind:

- die Verbesserung der Arbeitsbedingungen für Arbeitsplätze, an denen auf Computerleistung zugegriffen werden muss, durch eine menschengerechte Kommunikationsform
- die Erweiterung der Einsatzmöglichkeiten von Computern und die Möglichkeit zu einer Demokratisierung des Informationszugangs durch Abbau der 'Sprachbarriere' zwischen Mensch und Computer.

Man unterscheidet in der sprachorientierten KI-Forschung zunächst grob drei Ebenen des Verstehens und der Generierung einer sprachlichen Äußerung :

- die *Syntax* als die Struktur der Äußerung
- die *Semantik* als die Bedeutung der Äußerung
- die *Pragmatik* als die Verwendung der Äußerung

Stellen wir uns ein natürlichsprachliches System als Komponente eines zukünftigen Industrieroboters vor, der durch natürlichsprachlichen Wissenserwerb von einem Ingenieur für die Montage eines neu entwickelten Motors umgerüstet wird und dabei die nacheinander eingegebene Montageanleitung schrittweise ausführt. Um auf die Eingabe (1) richtig

(1) BEN ¹: *Die große Flügelschraube muß den Kühler am Sockel halten.*

¹ Im weiteren werden Eingaben des Benutzers durch 'BEN' und Ausgaben des Systems durch 'SYS' markiert.

reagieren zu können, muß ein solches System die syntaktische Mehrdeutigkeit (Fachwort; *Ambiguität*; erste Lesart: 'Kühler soll am Sockel gehalten werden' oder zweite Lesart 'Kühler, welcher sich am Sockel befindet, soll gehalten werden') auf der semantischen Ebene auflösen (erste Lesart sinnvoller, da der Kühler an irgendetwas befestigt werden muß und die zweite Lesart wegen der Existenz eines einzigen Kühlers wenig plausibel ist) und auf der pragmatischen Ebene die in Form eines Aussagesatzes formulierte Äußerung als Handlungsanweisung ('Befestige den Kühler jetzt mit der großen Flügelschraube am Sockel') interpretieren können.

Aus der Perspektive der Pragmatik handelt es sich bei (1) um einen sog. *indirekten Sprechakt* (d.h. eine Äußerung, in der die vom Sprecher intendierte Wirkung auf den Hörer nicht grammatisch realisiert ist) und damit um ein Phänomen, für dessen algorithmische Bearbeitung in der sprachorientierten KI-Forschung erst vor kurzem erste Ansätze entwickelt wurden. Offensichtlich muß das System beim Verstehen von (1) auf der semantischen und pragmatischen Ebene in hohem Maße Wissen über den angesprochenen Sachbereich und die Dialogsituation ausnutzen.

Um die vom Ingenieur intendierte Handlung ausführen zu können, muß das System außerdem dazu in der Lage sein, die mit *dem Kühler*, *dem Sockel* und *der großen Flügelschraube* angesprochenen Gegenstände zu identifizieren.

Falls mehrere Kühler und Sockel in der Montagehalle vorhanden sind, müssen die entsprechenden, im Dialog zuletzt erwähnten oder am bisherigen Montageablauf beteiligten Gegenstände als Bezugsobjekte (Fachwort: *Denotate*) gewählt werden. Da es natürlich viele verschiedene Schrauben in der Montagehalle gibt, muß die Bedeutung des vagen Adjektivs *groß* analysiert werden, wobei auf die gespeicherte Erfahrungswerte für die Ausmaße von Schrauben (z.B. ist eine große Schraube kleiner als ein kleiner Kühler) Bezug genommen werden muß.

Schließlich muß das natürlichsprachliche System das Kompositum *Flügelschraube* auch bei erstmaligem Hören auf eine visuell präsente Schraube, die flügelähnliche Griffe aufweist, beziehen können.

Ein vielseitig verwendbares natürlichsprachliches System, das eine der menschlichen Verstehensleistung entsprechende Analyse einer so einfachen Äußerung wie (1) durchführen kann und dabei die erwähnten Fähigkeiten zur semantischen Auflösung von Mehrdeutigkeiten, zur Erkennung indirekter Sprechakte, zur Analyse vager Ausdrücke und zum Verstehen metaphorischer Wortbildungen ausnutzt, existiert bisher nicht. Obwohl für Teilaspekte jeder einzelnen der angegebenen Fähigkeiten in der KI-Literatur implementierte Beispielsysteme dokumentiert sind, ist man beim derzeitigen Forschungsstand wegen der außerordentlichen Komplexität der erforderlichen Systemkomponenten von einer Integration dieser Fähigkeiten in einem 'Universalsystem' noch weit entfernt.

1.2. DEFINIERENDE EIGENSCHAFTEN NATÜRLICHSPRACHLICHER SYSTEME

In der sprachorientierten KI-Forschung spricht man mit Bezug auf die drei im vorangegangenen Abschnitt eingeführten semiotischen Ebenen von *natürlichsprachlichen Systemen* (Abk.: NSS), wenn

- eine Teilmenge der in das System eingegebenen oder vom System ausgegebenen Nachrichten natürlichsprachlich codiert ist und
- zur Verarbeitung der Nachrichten syntaktische und semantisch-pragmatische Verfahren zur Analyse und Generierung natürlicher Sprache eingesetzt werden.

Die erste Bedingung sichert, daß auch solche Systeme als NSS bezeichnet werden können, die neben einer natürlichsprachlichen Ein- oder Ausgabe auch anders codierte Nachrichten mit ihrer Umgebung austauschen. Beispielsweise wird in einem laufenden KI-Projekt der Firma BBN ein NSS entwickelt, das auf

Eingaben des Benutzers wie (2), in der eine natürlichsprachliche Äußerung mit einer Zeigehandlung verknüpft ist, mit einer entsprechenden Veränderung der gezeigten Graphik und einer zusätzlichen natürlichsprachlichen Ausgabe (z.B. der Rückfrage 'Ist das genau genug') reagieren kann¹.

(2) BEN: *Entferne diesen* < der Benutzer zeigt z.B. mit einem Lichtgriffel auf ein graphisches Objekt, das er auf dem Sichtgerät sieht > *Zustand und zeige mir den Rest genauer.*

Die zweite Bedingung verhindert, daß Systeme, die lediglich einfachste Zeichenkettenoperationen über natürlichsprachlich codierten Texten durchführen (z. B . Texteditoren, sprachstatistische Systeme) als NSS aufgefaßt werden können.

SYS: *Hallo, mein Name ist GUS. Ich kann Ihnen beim Planen einer einfachen Flugreise helfen. Wohin möchten Sie?*

SEN: *Ich möchte am 28. Mai nach San Diego.*

SYS: *Wann möchten Sie von hier weg?*

SEN: *Um 10 Uhr muß ich in San Diego sein.*

SYS: *Wie wäre es mit dem Flug Nr. 102 der P.S.A., der um 9.15 Uhr ankommt?*

BEN: *Welches ist der nächste Flug?*

SYS: *A.C.- Flug Nr. 310, der um 8.30 Uhr hier losgeht und um 10.00 Uhr ankommt.*

SEN: *Ich nehme den ersten.*

SYS: *Auf welchen Namen soll ich die Reservierung vornehmen, bitte?*

SEN: *Dan Bobrow.*

SYS: *Ich habe folgenden Flug gebucht:*

P.S.A. Flug Nr. 102 am Mittwoch den 28. Mai von San José nach San Diego, der um 7.30 Uhr startet und 9.15 Uhr ankommt. An welchem Tag möchten Sie zurückkommen?

BEN: *Am Freitagabend.*

SYS: *Wie wäre es mit dem Flug, der um 19.45 Uhr losgeht?*

BEN: *Prima.*

Fig. 1: Beispieldialog mit dem System GUS

¹ In der sprachorientierten KI-Forschung wird langfristig durchaus angestrebt, die vielfältigen Möglichkeiten der Integration von Mimik, Gestik und anderen nicht-verbalen Handlungen in einen natürlichsprachlichen Dialog, wie sie der alltäglichen menschlichen Kommunikation zugrundeliegen, auch für die Mensch/Maschine-Kommunikation verfügbar zu machen.

TEXTEINGABE

BEN: *Peter wollte das Fahrrad von Uwe. Er ging zu Uwe und fragte ihn, ob er es ihm gibt. Uwe lehnte das ab. Peter sagte Uwe, er würde ihm dafür 20 DM geben aber Uwe war damit nicht einverstanden. Dann sagte Peter, er würde ihm den Arm brechen, wenn er es ihm nicht überlassen würde. Uwe gab ihm das Fahrrad.*

FRAGE-ANTWORT-MODUS

BEN: *Warum ging Peter zu Uwe?*

SYS: *Weil er sein Fahrrad haben wollte.*

BEN: *Warum gab Uwe Peter sein Fahrrad?*

SYS: *Weil er nicht verletzt werden wollte.*

BEN: *Was hatte es für Konsequenzen, daß Peter zu Uwe ging?*

SYS: *Dies ermöglichte ihm, Uwe um sein Fahrrad zu bitten.*

BEN: *Was waren die Konsequenzen, von Peters Frage an Uwe, ihm Uwes Fahrrad zu geben?*

SYS: *Uwe sagte ihm, daß Uwe ihm nicht sein Fahrrad geben würde.*

PARAPHRASIERUNGSMODUS

SYS: *Wer soll die Geschichte erzählen?*

BEN: *Peter*

SYS: *Ich wollte Uwes Fahrrad haben, also ging ich zu ihm und bat ihn darum. Er sagte mir, daß er es mir nicht geben würde. Also bat ich ihn, es mir für 20 DM zu verkaufen. Da sagte er, daß er es nicht vergeben würde. Ich sagte, wenn er es mir nicht gäbe, würde ich ihm seinen Arm brechen. Er gab mir sein Fahrrad.*

SYS: *Wer soll die Geschichte erzählen?*

BEN: *Uwe*

SYS: *Peter kam zu mir. Er fragte mich, ob ich ihm mein Fahrrad geben würde. Ich wollte das Fahrrad behalten, also sagte ich, daß ich es ihm nicht geben würde. Da bot er an, es mir für 20 DM abzukaufen. Ich wollte das Fahrrad behalten. Also sagte ich, daß ich es ihm nicht geben würde. Er sagte, wenn ich es ihm nicht gäbe, würde er mir meinen Arm brechen. Ich wollte nicht verletzt werden. Also gab ich ihm mein Fahrrad.*

Fig. 2: Beispieldialog mit dem System PAM

Dabei wird es als Evidenz für das semantisch-pragmatisch fundierte Sprachverstehen eines Systems angesehen, wenn das System

- Fragen über ein Sachgebiet oder einen Text beantwortet,
- einen aufgabenorientierten Dialog (z.B. Beratungsgespräch) führt oder Konversation betreibt,
- einen Text paraphrasiert, zusammenfaßt oder nacherzählt,
- sich gemäß natürlichsprachlich eingegebenen Anweisungen verhält oder danach handelt,
- einen Text in eine andere Sprache übersetzt.

Zur Veranschaulichung der vorangegangenen Abschnitte werden in den Figuren 1 und 2 Beispieldialoge¹ mit zwei natürlichsprachlichen Systemen wiedergegeben, deren Sprachverstehen vom Benutzer mit vier unterschiedlichen Methoden getestet werden kann.

Das System GUS (Genial Understanding System, Bobrow et al. 1977) führt mit dem Benutzer aufgabenorientierte Dialoge, in denen eine Flugreise des Benutzers geplant und die damit verbundenen Reservierungen vorgenommen werden (vgl. Fig. 1). Hier können die Flexibilität der Dialogführung und die Korrektheit der vom System ausgeführten Buchungen als Evidenzen für das Sprachverstehen des NSS gewertet werden.

Im System PAM (Plan Applier Mechanism, Wilensky 1978) kann der Benutzer nach Eingabe einer Kurzgeschichte das System z.B. nach Begründungen für die in der Kurzgeschichte beschriebenen Handlungen fragen (vgl. Fig. 2). Das Textverstehen von PAM wird durch Antworten belegt, die in dieser Form nicht explizit im eingegebenen Text enthalten sind. Beispielsweise wird in der ersten Systemantwort mit der Konjunktion *weil* ein Zusammenhang zwischen den ersten beiden Sätzen des eingegebenen Textes hergestellt, den der Text in seiner konkret vorliegenden Form (Fachwort: *Oberflächenstruktur*) nicht enthält. In der zweiten Systemantwort tritt mit *verletzt* ein Adjektiv auf, das ebenfalls nicht in der Oberflächenstruktur des Textes vorkommt und dessen Verwendung in der Begründung des Systems sich nur auf Schlußfolgerungen (Fachwort: *Inferenzen*) stützen kann. Im Paraphrasierungsmodus kann PAM den in der Kurzgeschichte beschriebenen Handlungsablauf aus der Perspektive der beteiligten Akteure nacherzählen (vgl. Fig. 2), wobei die Nacherzählung wiederum Aussagen enthält, die als Evidenzen für eine semantisch-pragmatische Verarbeitung des eingegebenen Textes gewertet werden können.

An dieser Stelle ist eine Warnung vor der Gefahr der Übergeneralisierung von an Beispieldialogen demonstrierten Systemfähigkeiten angebracht. Im vorangegangenen Abschnitt wurde nur von Evidenzen und nicht von Beweisen für Sprachverstehen gesprochen, weil Rückschlüsse von Beispielen des Ein-/Ausgabeverhaltens eines NSS (Fachwort: *Performanz*) auf die sprachlichen, kommunikativen und kognitiven Fähigkeiten eines Systems nur in sehr begrenztem Umfang möglich sind. Wissenschaftlich fundierte Methoden des Tests (vgl. v. Hahn 1978) und der Evaluierung (vgl. Krause 1982) von NSS liegen bisher erst ansatzweise vor. Auf keinen Fall sollte aus der korrekten Behandlung eines sprachlichen Phänomens in einem Beispieldialog der Schluß gezogen werden, daß das entsprechende NSS dieses Phänomens in allen Ausprägungen korrekt behandelt (vgl. auch Tennant 1981, S.10/11, 164/165 und Morik 1982, S. 9). Die sicherste, aber zugleich auch aufwendigste Methode, zu einer realistischen Einschätzung der Fähigkeiten eines NSS zu kommen, ist zur Zeit immer noch die Detailanalyse der Verarbeitungskomponenten eines NSS, die zeigt, auf welche Weise Performanzleistungen zustande kommen. Ein Ziel dieses Aufsatzes ist es, dem Leser das zu einer solchen Beurteilung notwendige Grundwissen zu vermitteln. Dazu gehört zunächst die Definition einiger weiterer Grundbegriffe.

¹ Es handelt sich um Auszüge aus tatsächlichen Interaktionsprotokollen, die vom Englischen ins Deutsche übersetzt wurden.

1.3. GRUNDBEGRIFFE DER SPRACHORIENTIERTEN KI-FORSCHUNG

Wie bereits die einführende Diskussion des Beispielsatzes (1) zeigte, ist Sprachverstehen ohne ein gewisses Maß an Allgemeinwissen und Kenntnissen über die besprochene Situation und den angesprochenen Gegenstandsbereich (Fachwort: *Diskursbereich*, oder: *Diskurswelt*, *Weltausschnitt*, *Weltmodell*) nicht möglich. Natürlichsprachliche Systeme werden daher heute grundsätzlich als *wissensbasierte Systeme* konstruiert.

Dabei wird unter *Wissen* in der KI 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. Obwohl die Wahl einer problemgerechten Darstellung von Daten eine allgemeine Fragestellung innerhalb der Informatik ist, stellt sich das Problem der Repräsentation von Wissen bei der Konstruktion von NSS verschärft, da solche Systeme auf einen umfangreichen Wissensfundus zurückgreifen müssen, der aus vielen heterogen strukturierten Einheiten bestehen kann.

In der KI wurden verschiedene Hilfsmittel zur systematischen Repräsentation von Wissen entwickelt, die man als *Wissensrepräsentationssprachen* bezeichnet. Die durch eine Wissensrepräsentationssprache definierten Ausdrücke nennt man *Repräsentationskonstruktionen*. Die Gesamtheit an Wissen, die einem KI-System zur Verfügung steht, nennt man *Wissensbasis*. Die Wissensbasis gliedert sich in *Wissensquellen*, die wiederum aus einzelnen *Wissenseinheiten* bestehen. Wissen, das sich auf anderes Wissen innerhalb der Wissensbasis bezieht, nennt man *Metawissen*. Es sind im allgemeinen mehrere Metaebenen zu unterscheiden.

Man unterscheidet in der sprachorientierten KI-Forschung die sprachexterne Bedeutung eines sprachlichen Ausdrucks, die seinen Bezug zum Diskursbereich charakterisiert und in der *Referenzsemantik* behandelt wird, von seiner sprachinternen Bedeutung, die als Relation des Ausdrucks zu anderen Ausdrücken aufgefaßt wird und im Rahmen der *Inhaltssemantik* beschrieben wird.

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 *Kühler* wird in der Referenzsemantik die Menge von Individuen betrachtet, die in einem bestimmten Kontext als Kühler bezeichnet werden (vgl. Beispielsatz(1)).

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. Die Extension ist also der Wert der Intension bei einem bestimmten Index.

Beispiele für in der Inhaltssemantik erfaßte Relationen zwischen sprachlichen Ausdrücken sind die starke Bedeutungsähnlichkeit (Fachwort: *Synonymie*, z.B. Fleischer - Metzger - Schlachter) und die Gegensätzlichkeit von Bedeutungen (Fachwort: *Antonymie*, z.B. faul - fleißig). Zusätzliches begriffliches Wissen wird in der Inhaltssemantik durch sog. *Bedeutungspostulate* erfaßt, die in NSS meist durch Inferenzregeln realisiert werden (z.B. $ROT(X) \rightarrow FARBIG(X)$).

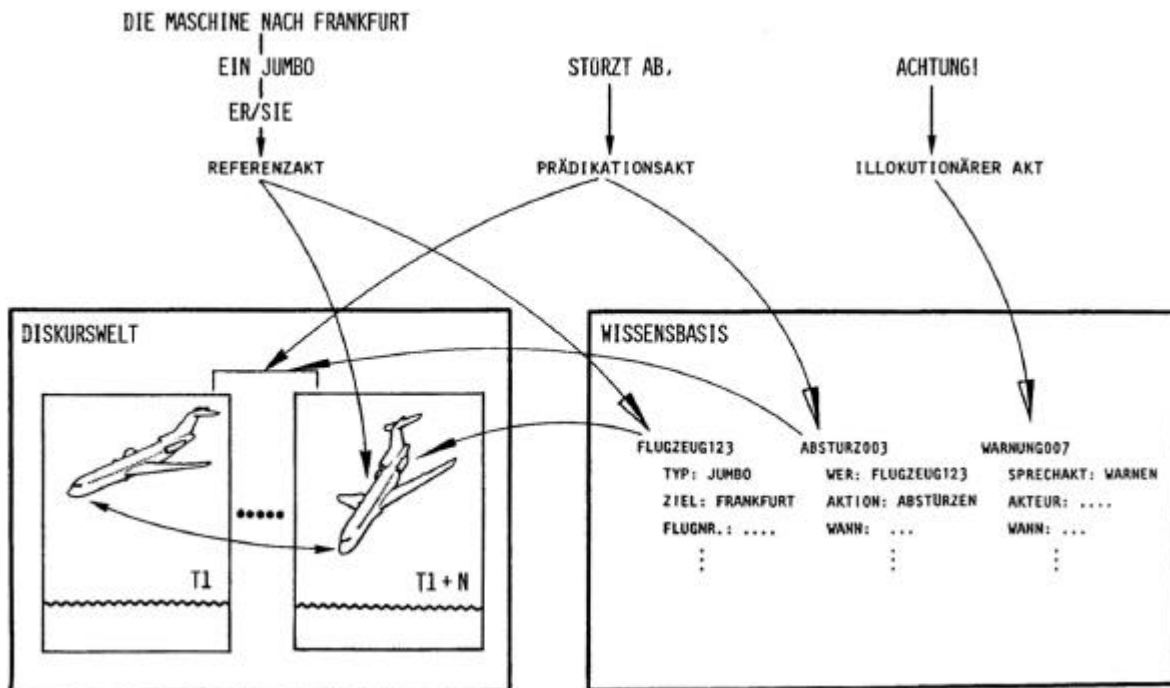


Fig. 3: Referenzsemantik in KI-Systemen am Beispiel *Die Maschine nach Frankfurt stürzt ab, Achtung!*

Aus pragmatischer Sicht wird Sprechen als eine Form von Handeln betrachtet, das wie nicht-sprachliches Handeln zur Verwirklichung von Absichten und Zielen (Fachwort: *Intentionen*) dient. Dabei wird jede Äußerung einem bestimmten Typ von sprachlicher Handlung (Fachwort: *Sprechakt*) zugeordnet. Ein Sprechakt wie *Die Maschine nach Frankfurt stürzt ab, Achtung!* muß von einem NSS in einen *propositionalen Akt* (d.h. dem, was 'gesagt' wird) und einen *illokutionären Akt* (d.h. dem, was 'gemeint' ist) untergliedert werden (vgl. Fig. 3). Der propositionale Akt wird weiter gegliedert in *Referenzakte* (d.h. das Identifizieren von Denotaten in der Diskurswelt) und *Prädikationsakte* (d.h. das Charakterisieren). Der Ausdruck *Achtung!* ist wie *Ich warne Dich davor, daß ...* ein expliziter Indikator für den vollzogenen illokutionären Akt (Fachwort: *performativer Ausdruck*).

In Figur 3 werden die referenzsemantischen Verhältnisse in einem natürlichsprachlichen KI-System veranschaulicht. Diese zeichnen sich dadurch aus, daß neben den referentiellen Beziehungen zwischen sprachlichen Ausdrücken und der Diskurswelt (Pfeile mit breiter Spitze in Figur 3) sowohl die Relationen zwischen sprachlichen Ausdrücken und den in der Wissensbasis enthaltenen symbolischen Repräsentationskonstruktionen für deren Denotate¹ (Pfeile mit halb gefüllter Spitze in Figur 3) als auch Relationen zwischen den in der Wissensbasis enthaltenen Repräsentationskonstruktionen und den Objekten und Vorgängen in der Diskurswelt (Pfeile mit schmaler Spitze in Figur 3) berücksichtigt werden müssen.

¹ Der von *Achtung* wegführende Pfeil nimmt eine Sonderstellung ein, da er auf eine Repräsentationskonstruktion für einen erkannten Sprechakt verweist.

1.4. FORSCHUNGSLEITLINIEN UND ARBEITSSCHWERPUNKTE

Das Paradigma der sprachorientierten KI-Forschung kann durch die Kombination von vier *Leitlinien* charakterisiert werden (vgl. Wahlster 1982):

- (L₁) Angestrebt wird eine vollständig operationalisierte, extreme Form der Referenzsemantik, die bis auf die sensorische Ebene 'durchgeschaltet' wird. Stets wird mit einer expliziten, maschinell handhabbaren und vollständigen Repräsentation der Diskurswelt gearbeitet.
- (L₂) Angestrebt wird eine ganzheitliche Modellierung der Interaktion zwischen dem Sprachverstehen, der Sprachproduktion und dem Spracherwerb zugrundeliegenden Prozessen und anderen für die Sprachverwendung relevanten kognitiven und sozialen Prozessen.
- (L₃) Angestrebt wird eine prozeßorientierte Rekonstruktion von sprachlichem Verhalten, wobei Sprachgebrauch prozedural als Ergebnis kommunikativer und kognitiver Prozesse beschrieben wird.
- (L₄) Angestrebt wird die Modellierung des instrumentellen Charakters von Sprache, da in KI-Systemen Sprache als Werkzeug in Arbeitsprozessen dienen soll.

Für jede einzelne dieser in zufälliger Reihenfolge aufgelisteten Leitlinien gilt, daß die damit verbundenen Zielvorstellungen nicht nur für die grundlagenorientierte Perspektive (vgl. (A₁) in Abschnitt 1.1.), sondern auch für die Anwendungsperspektive (vgl. (A₂) in Abschnitt 1.1.) der sprachorientierten KI-Forschung von Bedeutung sind.

Der wissenschaftliche Erfolg der von diesen Leitlinien ausgehenden Forschungsrichtung ist sowohl vom Fortschritt in grundlegenden KI-Bereichen wie der Repräsentation von Wissen und der Organisation komplexer kognitiver Systeme als auch von den Ergebnissen anderer Arbeitsgebiete der KI wie Bildverstehen und Automatische Deduktion abhängig. Außerdem sind die genannten Aufgaben ohne interdisziplinäre Zusammenarbeit mit den Fachwissenschaften Linguistik und Psychologie, die beide wichtige Grundlagen liefern, nicht sinnvoll zu bearbeiten. Neben der KI gibt es mindestens noch drei weitere Forschungsgebiete, die wichtige Beiträge zur Entwicklung von NSS leisten: die Linguistische Datenverarbeitung (LDV), die Informationswissenschaft (IW) und die Kognitionswissenschaft (KW).

Zur Zeit gelten als Hauptarbeitsgebiete der sprachorientierten KI-Forschung die ersten drei der im folgenden genannten *Forschungsschwerpunkte* (vgl. Wahlster 1981a):

- Frage-Antwort-Systeme und Dialogsysteme
- Textverstehende Systeme
- Verstehen gesprochener Sprache
- Natürlichsprachliche Wissensaufnahme
- 'Abstract'-Generierung
- Korrekturlesen von Texten
- Sprachliche Bildbeschreibung
- Automatische Übersetzung
- Natürlichsprachliche Programmierung

Das bisher erfolgreichste Arbeitsgebiet ist die Konstruktion von Frage-Antwort- und Dialogsystemen, die oft als sog. *natürlichsprachliche Schnittstellen* einen besonders komfortablen Zugang zu den in der folgenden Liste zusammengestellten Systemklassen ermöglichen:

- Datenbanksysteme und Methodenbanksysteme
- Wissensbasierte Experten- und Beratungssysteme
- Tutorielle KI-Systeme
- Theorembeweiser
- Szenenanalysesysteme
- Graphik-Systeme

Natürlichsprachliche Systeme zeichnen sich durch *zielgesteuertes Verhalten* aus, wobei das Hauptziel meist aus einer Hilfestellung für den Systembenutzer besteht, z.B. bei der

- Auskunftserteilung (z.B. HAM-RPM, vgl. v. Hahn et al. 1980)
- Beratung (z.B. GUS, vgl. Bobrow et al. 1977)
- Kooperation bei der Lösung einer Aufgabe (z.B. BBN-Graphik-System, vgl. Brachmann et al. 1979)
- Ausbildung (z.B. SCHOLAR, vgl. Collins et al. 1975)

In letzter Zeit wurden auch experimentelle NSS entwickelt, deren Verhalten zusätzlich durch solche vorgegebenen Ziele gesteuert wird, die keine reine Hilfsfunktion des Systems bewirken, solche Ziele sind z.B.

- Informationsgewinnung durch das System (z.B. KLAUS, vgl. Haas/Hendrix 1980)
- Verkauf (z.B. Hotelreservierungssituation in HAM-RPM, vgl. Jameson et al. 1980)

Selbstverständlich werden anfangs jegliche 'Systemziele' entweder bereits beim Entwurf eines KI-Systems fest vorgegeben oder beim Systemstart explizit als Parameter übergeben (z.B. in POLITICS, vgl. Carbonell 1980, das verschiedene 'politische Zielvorstellungen verfolgen' kann).

Die Breite des Spektrums der gewählten Diskursbereiche soll anhand deutschsprachiger Systeme durch folgende, alphabetisch nach Systemnamen geordnete Liste verdeutlicht werden:

- ATN-SIC (vgl. Metzging 1981): Wegebeschreibungen
- BACON (vgl. Habel et al. 1980): medizinische Daten, Kongreßreisen
- Erlanger System für gesprochene Sprache (vgl. Hein 1981): Fahrplanauskünfte
- FAS-80 (vgl. Witschas et al. 1980) : Dokumentation einer Programmbibliothek
- HAM-ANS (vgl. Nebel/Marburger 1982): Straßenverkehrsszene, Hotelreservierung, Fischereidaten
- KIPUS (vgl. Fauser 1980): Arbeitsmarktberichte
- NATAN (vgl. Liebisch 1980): Meldungen über militärische Kampfhandlungen
- PLIDIS (vgl. Kolvenbach et al. 1979): Abwasserüberwachung
- USL (vgl. Ott 1979): Personaldaten, Schulnoten

Nachdem in den vorangegangenen Abschnitten einige Elemente der theoretisch-methodischen Grundkonzeption und des Begriffsinventars der sprachorientierten KI-Forschung eingeführt wurden, soll im nächsten Abschnitt die historische Entwicklung dieses Arbeitsgebietes skizziert werden, bevor dann in den weiteren Kapiteln KI-Verfahren zur Konstruktion natürlichsprachlicher Systeme beschrieben werden.

1.5. ZUR HISTORISCHEN ENTWICKLUNG DER SPRACHORIENTIERTEN KI-FORSCHUNG

Bereits seit dem Mittelalter haben immer wieder einzelne Wissenschaftler, die sich mit Problemen der Logik, Sprachwissenschaft, Sprachphilosophie und Sprachpsychologie beschäftigten, das Ziel verfolgt, Aspekte der menschlichen Sprachfähigkeit algorithmisch zu erfassen. Als in den vierziger Jahren die ersten Computer verfügbar wurden, begann man daher schon bald, sprachverarbeitende Programme (z.B. Lexikonsuche, Flexionsanalyse) zu entwickeln.

1.5.1. Die Anfänge der maschinellen Sprachverarbeitung zu Beginn der fünfziger Jahre

Häufig wird das Jahr 1949, in dem Warren Weaver ein Memorandum zur *automatischen Übersetzung* verfaßte, als Beginn der maschinellen Sprachverarbeitung angesetzt. Die mit viel Enthusiasmus und großem finanziellen Aufwand begonnenen Projekte zur maschinellen Übersetzung (Abk.: MÜ), die mit einfachen Wörterbuchzugriffen, Textsubstitutionen und Umstellungstransformationen ohne Berücksichtigung der Semantik Qualitätsübersetzungen anstrebten, konnten die in sie gesetzten hohen Erwartungen nicht erfüllen. Übersetzungsergebnisse wie *The vodka is strong but the meat is rotten* als Rückübersetzung des automatisch ins Russische übertragenen Satzes *The spirit is willing but the flesh is weak* werden noch immer gerne als Beispiel für die fehlgeschlagene Wort-für-Wort-Übersetzung belächelt.

Die heute aus der Sicht der sprachorientierten KI-Forschung trivial wirkende Einsicht, daß eine Qualitätsübersetzung ohne ein auf Weltwissen basierendes Verstehen des zu übersetzenden Textes unmöglich ist, die Bar-Hillel 1960 gegen den damaligen MÜ-Ansatz vorbrachte, führte 1966 im sog. ALPAC-Report zum offiziellen Eingeständnis der fehlgeschlagenen Bemühungen und zur vorläufigen Einstellung der Forschungsförderung auf diesem Gebiet in den USA.

Neben den Arbeiten zur MÜ begann man bereits in den fünfziger Jahren, Computer auch als Hilfsmittel für die sprachwissenschaftliche Forschung einzusetzen, wobei zunächst die sprachstatistische Auswertung und philologische Aufbereitung großer Textmengen (z.B. als Indices und Konkordanzen) und die maschinelle Unterstützung von Linguisten bei der Konstruktion von umfangreichen Lexika und Grammatiken im Vordergrund stand.

Dieser Zweig der maschinellen Sprachverarbeitung, der auch heute noch als Teil der Linguistischen Datenverarbeitung von Bedeutung ist, wird in dieser Übersicht nicht weiterverfolgt, da er keinen wesentlichen Beitrag zu den Fragestellungen der sprachorientierten KI-Forschung leistet.

Ausgehend von den frühen Arbeiten Chomskys zur Formalisierung der Syntax natürlicher Sprachen wurden Anfang der sechziger Jahre Parsing- und Generierungsalgorithmen u.a. für *kontextfreie Grammatiken* und *Transformationsgrammatiken* entwickelt, die z.T. sehr komplexe Sätze analysieren bzw. durch zufällige Auswahl eines Ableitungspfades syntaktisch korrekte, aber oft unsinnige Sätze generieren konnten (vgl. auch Fig. 5).

1.5.2. Die ersten natürlichsprachlichen Systeme Mitte der sechziger Jahre

Mitte der sechziger Jahre entstanden die ersten Systeme, die sich auf die Frage der *'semantischen Informationsverarbeitung'* beim Sprachverstehen konzentrierten (z.B. BASEBALL, SIR, STUDENT, vgl. Fig. 4). Diese Systeme markieren den Beginn der sprachorientierten KI-Forschung. Im Gegensatz zu den bis dahin rein syntaktisch orientierten Systemen weisen sie erstmals alle definierenden Merkmale für natürlichsprachliche KI-Systeme (vgl. Abschnitt 1.2.) auf.

Um die komplexen Fragen der Semantik überhaupt in Angriff nehmen zu können, schränkte man den Sachbereich, in denen diese frühen KI-Systeme ihr 'Sprachverstehen' unter Beweis stellen sollten, drastisch ein (z.B. Verwandtschaftsbeziehungen in SAD-SAM, Spiele der amerikanischen Baseball-Liga in BASEBALL). Außerdem beschränkte man sich auf wenige sehr einfache syntaktische Muster (z.B. 24

Satzmuster in SIR), die ohne großen Parsingaufwand in eine auf die jeweilige Aufgabenstellung abgestimmte semantische Repräsentationssprache (z.B. Lineare Gleichungssysteme in STUDENT, einfache Datenbankabfragen in BASEBALL) abgebildet werden konnten.

- 1949-1960 ANFÄNGE DER MASCHINELLEN SPRACHVERARBEITUNG: MASCHINELLE ÜBERSETZUNG UND LINGUISTISCHE DATENVERARBEITUNG
- 1960-1970 BEGINN DER SPRACHORIENTIERTEN KI-FORSCHUNG: EINFACHE FRAGE-ANTWORT-SYSTEME
- 1963 SAD-SAM (Lindsay), BASEBALL (Green)
 1966 DEACON (Craig), ELIZA (auch: DOCTOR, Weizenbaum), SYNTHES (Simmons et al.)
 1968 TLC (Quillian), SIR (Raphael), STUDENT (Bobrow), CONVERSE (Kellog)
- 1970-1980 WISSENSBASIERTE EXPERIMENTALSYSTEME UND NATÜRLICHSPRACHLICHE DATENBANKSCHNITTSTELLEN
- DIALOGSYSTEME
- 1972 SHRDLU (Winograd)
 1977 GUS (Bobrow et al.), PAL (Sidner et al.)
- NATÜRLICHSPRACHLICHE SCHNITTSTELLEN
- 1972 LUNAR (Woods et al.)
 1972-76 RENDEZVOUS (Codd), REL (Thompson), REQUEST (Plath)
 1977 LIFER (Hendrix), INTELLECT (vormals: ROBOT, Harris), PLANES (Waltz et al.)
 CO-OP (Kaplan)
- TEXTVERSTEHENDE UND TEXTGENERIERENDE SYSTEME
- 1975 MARGIE (Schank et al.), SAM (Schank et al.)
 1976-79 TALE-SPIN (Meehan), PAM (Wilensky), FRUMP (DeJong)
- 1980-1990 ANWENDUNGSORIENTIERTE ZUGANGSSYSTEME UND MODELLE KOMPLEXER KOMMUNIKATIONSLEISTUNGEN
- ROBUSTE, ANWENDUNGSORIENTIERTE DIALOGSYSTEME
 - INTEGRATION NATÜRLICHSPRACHLICHER KOMPONENTEN IN EXPERTENSYSTEME, BILDVERSTEHENDE SYSTEME UND GRAPHIKSYSTEME
 - PARTNERMODELLIERUNG FÜR DIE SPRECHAKTPLANUNG UND -ERKENNUNG
 - NATÜRLICHSPRACHLICHER WISSENSERWERB

Fig. 4: Entwicklungsphasen und Meilensteine der Systementwicklung in den USA

Bei diesen meist für Frage-Antwort-Situationen konzipierten Systemen gab es zwei Untersuchungsschwerpunkte: die Entwicklung von Verfahren zur effizienten Überführung einer eingegebenen Frage in eine semantische Repräsentationskonstruktion und die Suche nach möglichst allgemeinen Verfahren zur Beantwortung einer Benutzerfrage (z.B. durch Traversierung eines semantischen Netzes oder prädikatenlogische Beweisverfahren mit Antwortprädikat). Da die einfachen Repräsentationsformalisten meist sehr speziell auf die jeweilige Problemstellung des KI-Systems zugeschnitten waren und bei der Überführung der natürlichsprachlichen Eingaben komplexe sprachliche oder kommunikative Fähigkeiten (z.B. Erkennen von Anaphora, Auflösen von Ambiguitäten) unberücksichtigt blieben, konnten die für diese Systeme verwendeten Analyseverfahren kaum auf andere Aufgabenstellungen übertragen werden. Dagegen haben die in der Anfangsphase entwickelten Such- und Inferenzverfahren zur Fragebeantwortung wegen ihrer größeren Allgemeinheit die weitere Entwicklung der sprachorientierten KI-Forschung nachhaltig beeinflußt. Allerdings waren diese Verfahren noch sehr aufwendig (große Suchräume, Gefahr der kombinatorischen Explosion), da außer einem Minimum an Faktenwissen kaum andere Wissensquellen zur Steuerung der semantischen Auswertung verwendet wurden. Außerdem war das Wissen dieser Systeme meist nicht explizit in einer Wissensbasis gespeichert, sondern implizit in der Funktionsweise der Verarbeitungsprozeduren enthalten (z.B. die Transitivität der Teilmengenrelation als Teil des allgemeinen Suchverfahrens für ein semantisches Netz), wodurch eine Ergänzung oder Änderung des vom System verwendeten Wissens kaum möglich war.

In vielen Einführungs- und Überblicksaufsätzen zur sprachorientierten KI-Forschung werden diese frühen natürlichsprachlichen Systeme dargestellt und verglichen (z.B. in Barr/ Feigenbaum 1981, Tennant 1981, Waltz 1981, Winograd 1972, Winograd 1974). Die stark *systemorientierte* historische Perspektive dieser Aufsätze soll hier nicht aufgegriffen werden, sondern durch eine *methodisch-thematische* Betrachtungsweise abgerundet werden (vgl. Fig. 5¹).

Fig. 4, die eine systemorientierte Entwicklungsübersicht enthält, wird daher durch Fig. 5 ergänzt, in der statt der üblichen 'Systemgeschichte' eine 'Ideengeschichte', skizziert wird².

Nicht erfaßt in den Figuren 4 und 5 ist die Abhängigkeit des dargestellten Entwicklungsprozesses von den Fortschritten in der Hardware und Software, in den einschlägigen Nachbardisziplinen (z.B. Linguistik und Psychologie) und in anderen Teilgebieten der KI (z.B. Repräsentation von Wissen, KI-Programmiersprachen). Beispielsweise wären die natürlichsprachlichen KI-Systeme von heute schon mangels geeigneter Hardware und Software (z.B. Magnetband als einziger Sekundärspeicher, Fehlersuche anhand von binär codierten Speicherausügen) Anfang der sechziger Jahre nicht realisierbar gewesen.

Man beachte, daß Fig. 4 lediglich den Entwicklungsverlauf in den USA erfaßt. In Europa und Japan verlief die Entwicklung trotz der eindeutigen Führungsrolle amerikanischer Wissenschaftler häufig anders³. So wurde in Europa und Japan die Forschung zur syntaxbasierten MU auch nach 1966 weitergeführt und die Anfangsphase der sprachorientierten KI-Forschung in der Dekade 1960-1970 fast vollständig 'übersprungen'.

¹ In Fig. 5 wurde versucht, die Methoden möglichst konzise durch die entsprechenden Fachbegriffe zu charakterisieren und verwandte methodische Ansätze unter einem Begriff zusammenzufassen (z.B. lexikalische Dekomposition in der Präferenzsemantik von Wilks und der konzeptuellen Dependenz von Schank).

² Den Jahreszahlen in den Figuren 4 und 5 liegen jeweils die ersten allgemein zugänglichen Veröffentlichungen zugrunde. Selbstverständlich gab es in vielen Fällen schon zwei bis drei Jahre vorher unveröffentlichte Dissertationen, Institutsberichte oder Ideenskizzen in Konferenzberichten. Es handelt sich bei den Zeitangaben also lediglich um ungefähre Entstehungsdaten. Aus Gründen der Übersichtlichkeit wurde auf Literaturangaben verzichtet, zumal alle für eine Einführung relevanten Literaturhinweise in anderen Teilen dieses Aufsatzes enthalten sind.

³ Im Rahmen der vorliegenden Einführung kann auf die historische Entwicklung der sprachorientierten KI-Forschung in Europa und Japan nicht eingegangen werden.

- 1960 - Parser für kontextfreie Grammatiken
- 1962 - Generierung mit kontextfreien Grammatiken
- 1963 - Generierung mit Dependenzgrammatiken
- 1964 - Parser für Dependenzgrammatiken
- 1965 - Parser für Transformationsgrammatiken
- 1966 - Parsing als Pattern-Matching
 - Generierung durch Instantiierung von Satzmustern
- 1969 - Generierung mit Transformationsgrammatiken
- 1970 - ATN-Parsing
- 1972 - ATN-basierte Generierung
- 1973 - Chart-Parsing
- 1975 - Lexikon-basiertes Parsing
 - Diskriminationsnetze zur Lexikon-basierten Generierung
- 1976 - Semantisches Parsing
 - 'Island'-Parsing
- 1977 - Wortexperten-Parsing
- 1978 - Inkrementelle Generierung mit prozeduraler Grammatik
- 1980 - Kaskadierte ATNs
 - Deterministisches Parsing
 - Generierung durch Produktionensysteme
- 1981 - Mehrfachstrategie-Parsing

a) Analyse- und Generierungsverfahren

- 1968 - Prädikatenlogik als semantische Repräsentationssprache und inferenz-basierte Antwortgenerierung
 - 'Prozedurale Semantik' als Überführung einer natürlichsprachlichen Eingabe in eine Datenbank-anfrage
 - Antwortgenerierung als Suchprozeß in einfachen semantischen Netzen
- 1972 - Prinzip der referenzsemantischen Verankerung durch 'Miniwelten'
 - Integration metakommunikativer Dialogteile wie Präsuppositions zurückweisungen und Klärungsdialoge
 - Anaphoraauflösung durch Dämon-Prozeduren
- 1973 - Prinzip der lexikalischen Dekomposition und Tiefenkasusanalyse
 - Prinzip des Script-gesteuerten Textverstehens
- 1977 - Explizite Darstellung des Fokus u.a. zur Ellipsenrekonstruktion
 - Verwendung von Dialogschemata u.a. zur Organisation des Sprecherwechsels in aufgabenorientierten Dialogen
- 1978 - Erkennung und Planung von Sprechakten mithilfe eines expliziten Partnermodells

b) Verfahren zur semantisch-pragmatischen Verarbeitung

Fig. 5: Ein Überblick zur Entwicklung wichtiger Methoden zur Konstruktion natürlichsprachlicher Systeme

Erst unter dem Eindruck der weltweiten Anerkennung für Winograds Dialogsystem SHRDLU wandten sich Anfang der siebziger Jahre einige der bis dahin durch Ziele und Methoden der Linguistischen Datenverarbeitung geprägten europäischen und japanischen Forschungsgruppen dem Paradigma der sprachorientierten KI-Forschung zu.

1.5.3. Die Etablierung und Ausweitung des Forschungsgebietes seit Beginn der siebziger Jahre

SHRDLU (vgl. Fig. 4) ist einer der Meilensteine der sprachorientierten KI-Forschung, da hier erstmals ein umfangreiches Bündel sprachlicher, kommunikativer und kognitiver Fähigkeiten in einem System mit eindrucksvoller Performanzleistung vereint wurden. In SHRDLU wurden natürlichsprachige Benutzereingaben aus dem Diskursbereich der inzwischen klassischen MIT-Blockwelt in Prozeduren der KI-Programmiersprache MICRO-PLANNER übersetzt, deren Evaluation je nach Eingabetyp (Frage, Aussage, Befehl) Suchprozesse, Änderungen in der Wissensbasis oder Aktionen in der Blockwelt auslöste. Allerdings zeigte sich, daß viele der in diesem experimentellen System implementierten Algorithmen, z.B. zur Analyse und Generierung von Pronomen oder zur Beantwortung metakommunikativer Fragen, speziell auf die zugrundeliegende Dialogsituation und den Diskursbereich zugeschnitten waren. Außerdem ließen sich aus den auf der Programmebene angesiedelten Systembeschreibungen nicht immer präzise wissenschaftliche Hypothesen über die modellierten Aspekte menschlichen Dialogverhaltens ableiten.

Einen vergleichbaren Einfluß, wie er von SHRDLU für die weitere Entwicklung der Grundlagenforschung im Bereich der natürlichsprachlichen Dialogsysteme ausging, hatte das ebenfalls Anfang der siebziger Jahre entwickelte System LUNAR auf die anwendungsorientierte Forschung im Bereich der natürlichsprachlichen Schnittstellen zu Datenbanken.

Das unter der Leitung von Woods entwickelte System LUNAR (ursprüngliche Bezeichnung: LSNLIS für Lunar Science Natural Language Information System) beantwortete natürlichsprachliche Anfragen von Geologen über Daten zu Gesteinsproben der Apollo-11-Expedition, die in einer aus einer einzigen siebenstelligen Relation mit 13000 Elementen bestehenden Datenbank gespeichert waren. In LUNAR wurde das von Woods 1970 veröffentlichte Parsing-Konzept der Augmented Transition Networks (Abk. ATNs, vgl. auch Fig. 5), das mit verschiedenen Erweiterungen und Variationen bis heute das am häufigsten verwendete Parsing-Konzept blieb, erfolgreich in einem größeren Rahmen erprobt.

Wie für SHRDLU so erwies sich auch für LUNAR eine Erweiterung oder Anpassung an einen neuen Anwendungsbereich als unmöglich. Beide Systeme waren Prototypen mit kurzer Lebensdauer, die bewiesen, daß Dialogsysteme und natürlichsprachliche Schnittstellen 'prinzipiell machbar' sind, und damit viele weiterführende Forschungsaktivitäten auslösten. Bereits fünf Jahre nach LUNAR konnten dann 1977 mit LIFER und INTELLECT die ersten produktreifen, kommerziell einsetzbaren natürlichsprachlichen Schnittstellen vorgelegt werden (vgl. Fig. 4).

Wichtige Impulse für die gesamte sprachorientierte KI-Forschung (z.B. zum robusten Parsing, Island-Parsing, Blackboard-orientierten Systemarchitekturen, vgl. Fig. 5) ergaben sich aus dem Forschungsförderungsprogramm der ARPA (1971-1976, vgl. Fig. 4) zum Thema 'Verstehen kontinuierlich gesprochener Sprache'.

Grundlegend für die Konstruktion textverstehender Systeme sind die Mitte der siebziger Jahre in der Gruppe um Schank entstandenen Arbeiten (vgl. Fig. 4). Das Prinzip der lexikalischen Dekomposition der Semantik komplexer Konzepte in eine kleine Auswahl wohldefinierter semantischer Primitive und die Verfahren der script-gesteuerten Textanalyse gehören inzwischen zum klassischen Methodenrepertoire der sprachorientierten KI-Forschung (vgl. Fig. 5).

Zusammenfassend läßt sich feststellen, daß in den siebziger Jahren wichtige Grundkonzepte des Entwurfs, der Konstruktion und Implementation natürlichsprachlicher Systeme gefunden wurden (natural language interface engineering, vgl. auch Fig. 4). Die aus der Anwendungsperspektive wichtigen Probleme der

Robustheit, Flexibilität und Anpassbarkeit von natürlichsprachlichen Systemen an unterschiedliche Diskursbereiche konnten bis zum Ende der siebziger Jahre allerdings nicht abschließend gelöst werden. Außerdem stellte man nach Abschluß der Arbeiten an zahlreichen Experimentalsystemen Ende der siebziger Jahre fest, daß natürlichsprachliche Systeme selbst in stark eingeschränkten aufgabenorientierten Dialogen eine nach den Anfangserfolgen von SHRDLU und LUNAR unerwartete Fülle aufwendiger sprachlicher, kommunikativer und kognitiver Leistungen erbringen müssen (z.B. Umgang mit metaphorischem Sprachgebrauch, indirekten Sprechakten, auf den Gesprächspartner abgestimmte Dialogstrategien).

Dies bedeutet für die Forschung der achtziger Jahre (vgl. Fig. 4), daß sie einerseits versuchen muß, Modelle dieser komplexen Kommunikationsleistungen zu erstellen, wobei Sprachverhalten im größeren Zusammenhang mit anderem zielgerichtetem Verhalten (z.B. Handeln, Problemlösen, Planen, Lernen) zu sehen ist. Andererseits muß sich die anwendungsorientierte KI-Forschung der Herausforderung stellen, flexibel einsetzbare und robuste Zugangssysteme zu anderen wissensbasierten Systemen zu entwickeln.

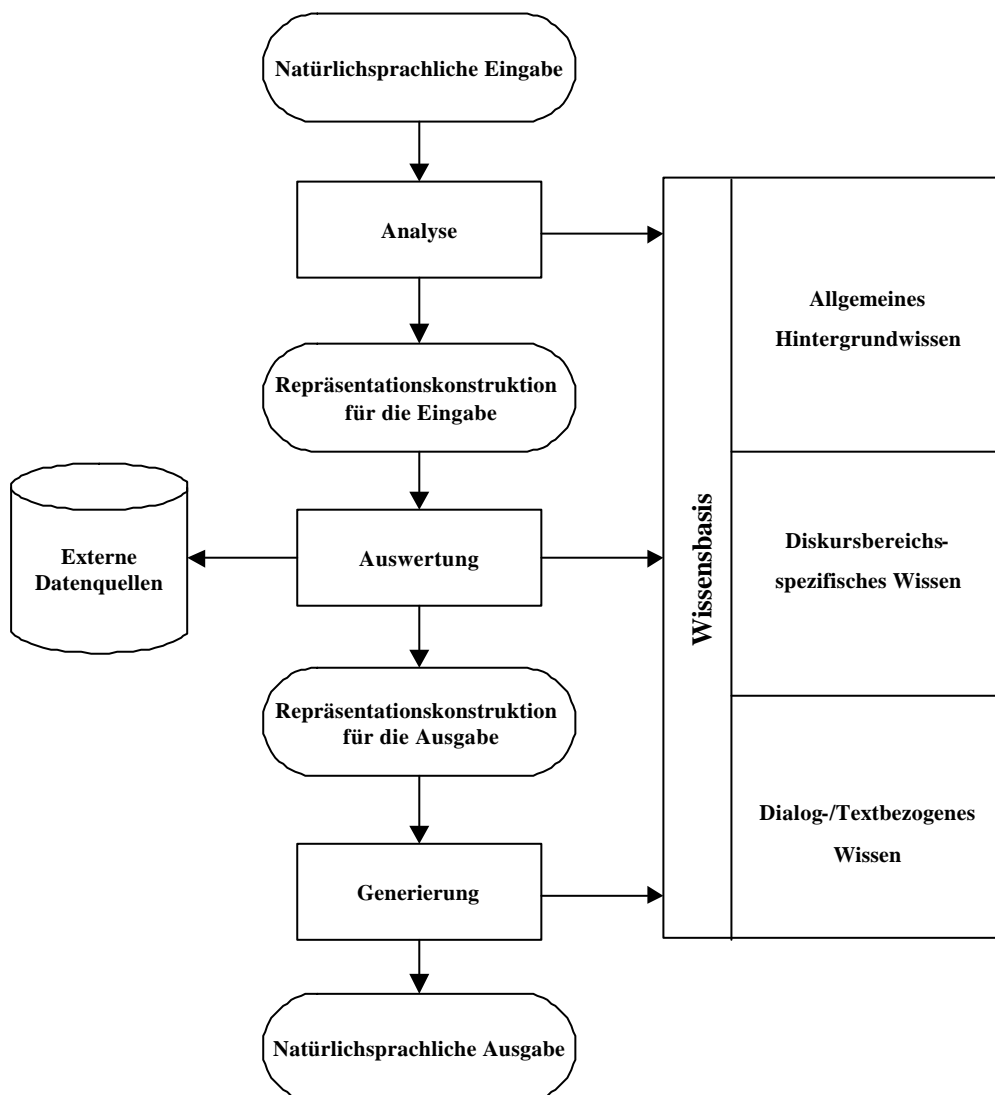


Fig. 6: Die Grobstruktur eines natürlichsprachlichen Systems

2. VERARBEITUNGSPHASEN UND WISSENSQUELLEN IN NATÜRLICH- SPRACHLICHEN SYSTEMEN

Die in Fig. 6 dargestellte Grobstruktur eines NSS besteht aus den drei großen Verarbeitungsabschnitten¹ *Analyse*, *Auswertung* und *Generierung*, nach denen im weiteren auch die Darstellung der einzelnen KI-Verfahren gegliedert werden soll. Ein schriftlich oder mündlich eingegebener natürlichsprachlicher Satz (im Falle eines Frage-Antwort- oder Dialogsystems) oder Text (im Falle eines textverstehenden Systems) wird im Verarbeitungsabschnitt 'Analyse' in einen Ausdruck einer Wissensrepräsentationssprache überführt, der für eine Auswertung geeignet ist. Je nach der von einem NSS geforderten Verstehensleistung (vgl. Abschnitt 1.2) kann es sich bei der Auswertung z.B. um das Auffinden der konzeptuellen Struktur der Antwort auf eine eingegebene Frage oder um die Zusammenfassung eines eingegebenen Textes handeln. Das Ergebnis der Auswertung wird dann wiederum in einem Ausdruck einer Wissensrepräsentationssprache² codiert, der im Verarbeitungsabschnitt 'Generierung' in eine natürlichsprachliche Ausgabe überführt wird.

Es gibt auch Ausprägungen des in Fig. 6 dargestellten Schemas, in denen der Verarbeitungsabschnitt 'Generierung' nicht realisiert ist (z.B. in NSS, in denen die Auswertung in der Planung einer physischen Operation oder dem Auffinden eines Dokumentes besteht). Umgekehrt können auch die Verarbeitungsabschnitte 'Analyse' und 'Auswertung' fehlen, wenn das NSS lediglich als Sprachgenerator eingesetzt werden soll (z.B. zur natürlichsprachlichen Beschreibung der Analyseergebnisse eines bildverstehenden Systems).

2.1 ZUM AUFBAU DER WISSENSBASIS VON NATÜRLICHSPRACHLICHEN SYSTEMEN

Die Wissensbasis eines NSS läßt sich in drei große Bereiche gliedern: das allgemeine Hintergrundwissen, das diskursbereichspezifische Wissen und das dialog-/textbezogene Wissen (vgl. Fig. 6).

Das *allgemeine Hintergrundwissen* eines NSS besteht als diskursbereichsunabhängiges Wissen aus einer systemspezifischen Auswahl folgender Wissensquellen:

- Wort-Lexikon
- Syntagmen-Lexikon
- Morphologisches Wissen
- Grammatisches Wissen
- Begriffliches Wissen
- Allgemeine Inferenz- und Metainferenzregeln
- Regeln der Dialog- bzw. Textgestaltung

Im *Syntagmen-Lexikon* sind u.a. feste Redewendungen (z.B. Wie geht's?) und aus zusammenhängenden Wortgruppen bestehende sprachliche Einheiten (Fachwort: *Mehrwortlexeme*, z.B. *Bundesrepublik Deutschland*) erfaßt. Das *morphologische Wissen* umfaßt Flexionsinformation (u.a. Deklinations- und Konjugationsformen) und Wortbildungsinformation. Im *begrifflichen Wissen* eines NSS wird die sprachinterne Bedeutung (vgl. Abschnitt 1.3.) sprachlicher Ausdrücke festgelegt. In *Metainferenzregeln* werden Strategien zur Steuerung von Inferenzprozessen codiert.

¹ In jedem Verarbeitungsabschnitt wirken mehrere, im weiteren noch genauer zu beschreibende Verarbeitungs-komponenten zusammen, die jeweils in sehr unterschiedlichen Systemarchitekturen (vgl. auch Abschnitt 2.2.) organisiert sein können.

² Diese Repräsentationssprache muß nicht mit der zur Codierung der Eingabe verwendeten identisch sein.

In NSS, in denen die Ein-/Ausgabe nicht schriftlich sondern in Form von kontinuierlich gesprochener Sprache erfolgt, kommen als allgemeines Hintergrundwissen noch zusätzlich *akustisch-phonetisches Wissen* (zur Segmentierung des Sprachsignals in Laute bzw. zur Lautsynthese) und z.T. auch *phonologisch-prosodische Wissensquellen* hinzu (z.B. Regeln der Koartikulation, d.h. wie sich Laute durch benachbarte Laute verändern, und Wissen über die Funktion von Betonung, Tonhöhe und Sprechgeschwindigkeit als nichtlautliche Bedeutungsträger).

Beispiele für *diskursbereichsspezifische Wissensquellen* sind:

- Referentielles Wissen
- Spezielle Inferenzregeln
- A priori Partnermodelle
- Spezielle Dialogstrategien bzw. Lese-/Schreibstrategien

Das *referentielle Wissen* besteht aus referenzsemantischen Relationen (vgl. Abschnitt 1.3.). In *Partnermodellen* ist Information über das beim Dialogpartner vermutete Vorwissen und seine voraussichtlichen Dialogziele gespeichert. A priori Partnermodelle enthalten Wissen über prototypische Benutzerklassen (z.B. 'Tourist' und 'Geschäftsreisender' in einem Hotelreservierungssystem), das es dem NSS ermöglicht, sein Verhalten auf das jeweilige Benutzerprofil einzustellen.

Als Beispiele für *dialog-/textbezogene Wissensquellen*, in denen während der Dialog- bzw. Textanalyse temporär Information gespeichert wird, seien genannt:

- Syntaktische und semantische Gedächtnisse
- Fokus
- Partnermodell
- Inferenzgedächtnis

In den *syntaktischen* und *semantischen Gedächtnissen* werden Resultate von syntaktischen und semantischen Verarbeitungskomponenten abgelegt, auf die im weiteren Dialog bzw. Text zurückgegriffen werden muß (z.B. bei der Vervollständigung eines syntaktisch unvollständigen Eingabesatzes durch Rückgriff auf die semantische Struktur des vorangegangenen Satzes).

Die als *Fokus* bezeichnete Wissensquelle enthält Information über die jeweiligen thematischen Schwerpunkte eines Dialog-/Textabschnittes (vgl. Grosz 1977). Diese Wissensquelle dient u.a. zur Auflösung von Mehrdeutigkeiten in der Benutzereingabe, zur sprachlichen Markierung des inneren Dialog-/Textzusammenhangs (Fachwort: *Kohärenz*, vgl. v. Hahn 1979) sowie zur Steuerung von Such- und Inferenzprozessen. Im Inferenzgedächtnis wird eine formale Beschreibung der vom NSS durchgeführten Inferenzen gespeichert, um später mithilfe der Erklärungskomponente (vgl. Wahlster 1981b) dem Benutzer inferenzbasiertes Systemverhalten erklären zu können.

Neben der Wissensbasis, auf die in allen Verarbeitungsabschnitten zugegriffen werden muß, werden im Verarbeitungsabschnitt 'Auswertung' in vielen NSS auch externe Datenquellen verwendet (vgl. Fig. 6). Beispiele für solche externen Datenquellen sind:

- in einer Datenbank gespeicherte Massendaten (z.B. Kursbuch der Bundesbahn)
- in einer Methodenbank gespeicherte Verfahrensbeschreibungen (z.B. statistische Verfahren).
- visuelle oder andere sensorische Information (z.B. Ergebnisse eines bildverstehenden Systems).

Häufig besteht das wichtigste Entwurfsziel eines NSS sogar darin, dem Benutzer einen komfortablen und effizienten Zugang zu der in den externen Datenquellen enthaltenen Information zu ermöglichen.

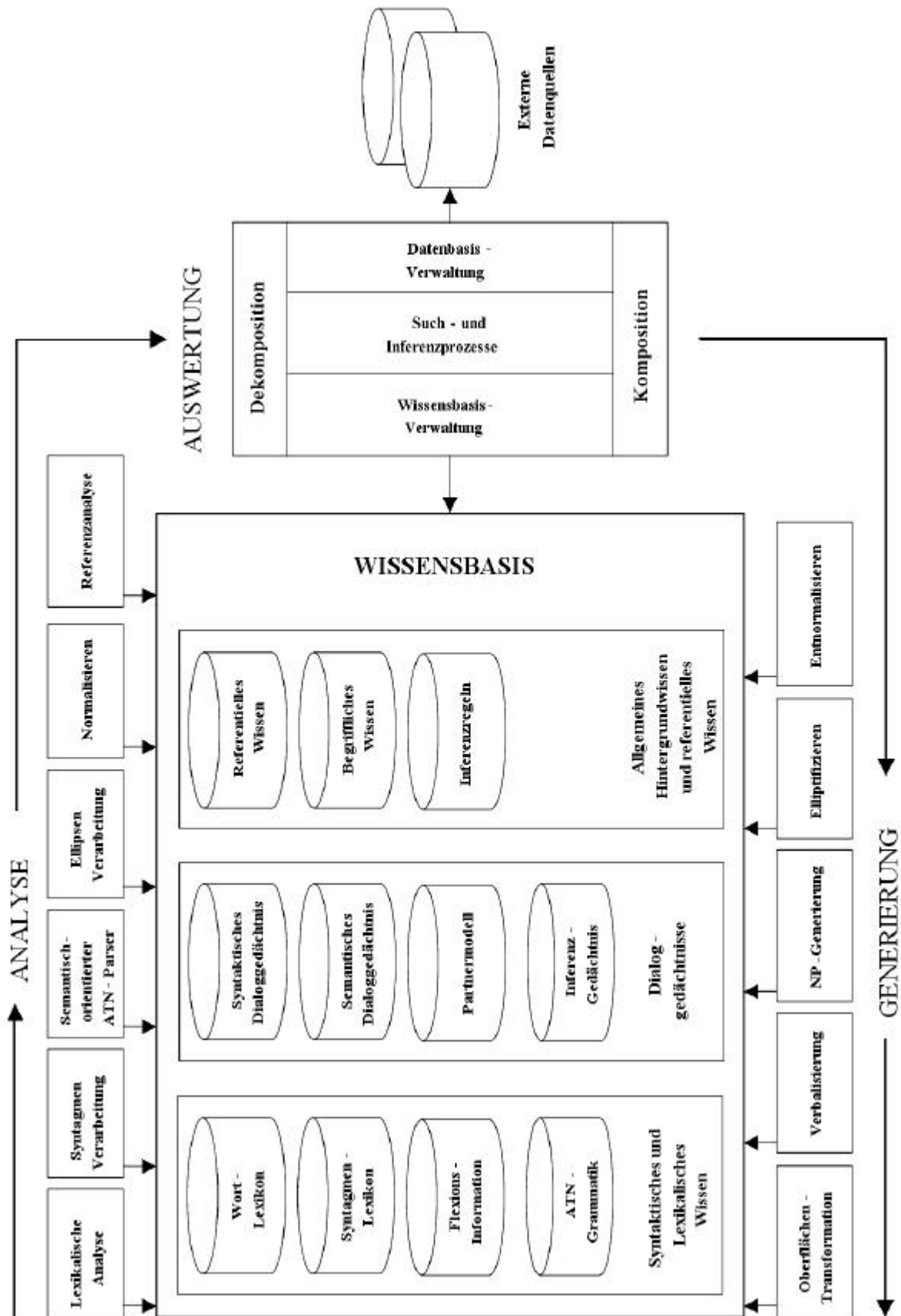


Fig. 7: Verarbeitungskomponenten und Wissensquellen des Systems HAM-ANS

Es gibt bisher kein NSS, in dem alle im vorangegangenen Abschnitt genannten Wissensquellen integriert sind. Vielmehr unterscheiden sich NSS durch eine jeweils spezifische Kombination von Wissensquellen. In einigen Systemen sind Wissensquellen, die nach der obigen Klassifikation im allgemeinen unabhängig vom Diskursbereich sind, speziell auf die Diskurswelt abgestimmt. Beispielsweise ist in NSS mit sog. *semantischen Parsern* (z.B. SOPHIE, PLANES, vgl. auch Kapitel 4) das lexikalische und grammatische Wissensdiskursbereichsspezifisch, so daß das Lexikon und die Grammatik bei der Adaptierung des NSS an einen neuen Diskursbereich ausgetauscht werden muß.

Als Beispiel für eine Ausprägung des in Fig. 6 dargestellten Schemas zeigt Fig. 7 (nach Nebel/Marburger 1982) die Verarbeitungskomponenten und Wissensquellen einer neueren Version des von uns entwickelten Systems HAM-ANS (Hamburger anwendungsorientiertes natürlichsprachliches System). Auf eine Darstellung des Kontroll- und Datenflusses zwischen den einzelnen Verarbeitungskomponenten wurde dabei verzichtet (vgl. Kapitel 4-6), da zuvor die verschiedenen in NSS verwendeten Systemarchitekturen eingeführt werden sollen.

2.2. ALTERNATIVE SYSTEMARCHITEKTUREN VON NATÜRLICHSPRACHLICHEN SYSTEMEN

Fig. 8 bietet einen Überblick zu den fünf wichtigsten *Systemarchitekturen*, die für die Konstruktion von NSS eingesetzt werden¹. Aus Gründen der Übersichtlichkeit wird in Fig. 8 die Struktur der einzelnen Verarbeitungsmodelle jeweils anhand von drei mit A, B und C bezeichneten Verarbeitungskomponenten verdeutlicht, wobei sich jedes der skizzierten Modelle ohne weiteres für n Verarbeitungskomponenten verallgemeinern lässt. Pfeile mit fettgedrucktem Schaft markieren in Fig. 8 die Eingabe der zu verarbeitenden Information.

Die bisher für NSS am meisten verwendeten Modelle sind das *Phasen-Modell* (z.B. durch LUNAR realisiert, vgl. Woods et al. 1972) und das *hierarchische Modell* (z.B. durch SHRDLU realisiert, vgl. Winograd 1972). Im Gegensatz zu Phasen-Modellen und hierarchischen Modellen sind *Kaskaden-Modelle* (z.B. durch RUS realisiert, vgl. Bobrow/Webber 1980), *Blackboard-Modelle* (z.B. durch HEARSAY realisiert, vgl. Reddy et al. 1976) und *heterarchische Modelle* (z.B. in Wortexperten-Parsern realisiert, vgl. Small 1981) auch für die Parallelverarbeitung geeignet.

Im Kaskaden-Modell² können analog zu einer Wasser-Kaskade die Kaskadenstufen A, B, C parallel arbeiten, sobald der 'Informationsfluß' die Komponente C erreicht hat. Im Blackboard-Modell tauschen die nebenläufig arbeitenden Komponenten A, B und C über eine Nachrichten-Tafel Information aus (vgl. Fig. 8). Die nebenläufig und unabhängig voneinander arbeitenden Komponenten eines heterarchischen Modells versenden und empfangen Botschaften aufgrund definierter Kontrollbeziehungen.

Komplexe Systemarchitekturen für NSS entstehen dadurch, daß erstens jede in einem der genannten Verarbeitungsmodelle integrierte Komponente selbst wieder aus vielen Modulen besteht, die nach einem anderen Verarbeitungsmodell organisiert sein können, und zweitens in verschiedenen Verarbeitungsabschnitten unterschiedliche Systemarchitekturen gewählt werden können (z.B. die Analyse als Kaskaden-Modell, die Auswertung als heterarchisches Modell und die Generierung als Phasen-Modell).

Typisch für die KI als experimentelle Wissenschaft ist, daß während der Entwicklungsphase eines NSS die Systemarchitektur einzelner Verarbeitungskomponenten mehrfach geändert wird, um empirisch eine optimale Strategie zu ermitteln. Beispielsweise wurde in unseren, Projekten HAM-RPM (vgl. v. Hahn et al. 1980) und HAM-ANS mehrfach eine mit wenigen Ausnahmen auskommende, heterarchische Strukturierung

¹ Eine vergleichbare Darstellung von vier Systemarchitekturen findet man in Görz 1979, S. 5

² Dieses Modell wurde in der sprachorientierten KI-Forschung erstmals in Form von sog. kaskadierten Augmented Transition Networks realisiert (vgl. Woods 1980 und Kapitel 3).

einer Komponente nach einer genauen Analyse der anwendungstypischen Daten- und Kontrollflüsse durch ein Phasen-Modell ersetzt.

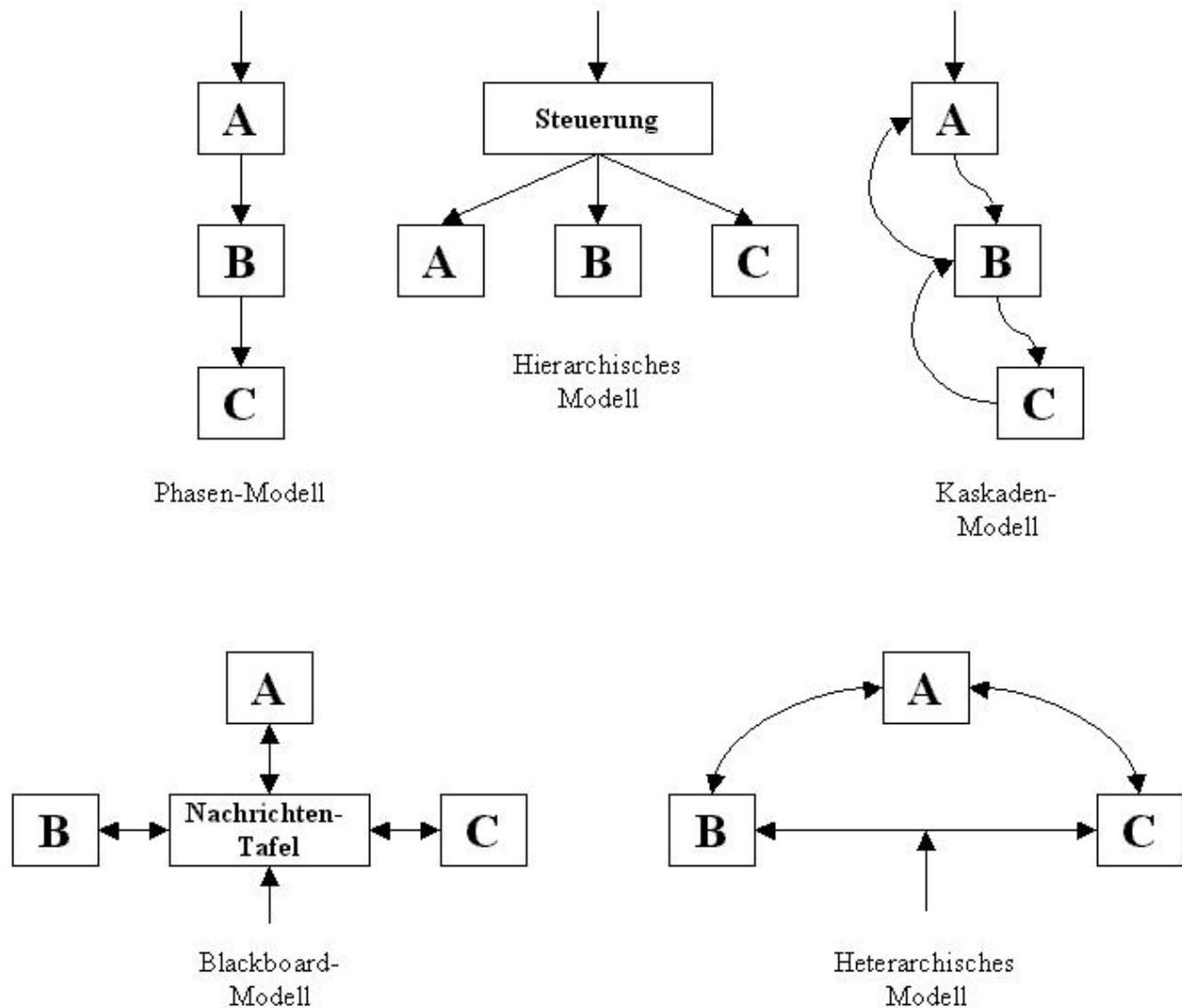


Fig. 8: Alternative Systemarchitekturen von natürlichsprachlichen Systemen

3. VERFAHREN ZUR SYNTAKTISCHEN ANALYSE NATÜRLICHER SPRACHE

In der sprachorientierten KI-Forschung wird häufig der gesamte Verarbeitungsabschnitt 'Analyse' unter dem Begriff *'Parsing'* zusammengefaßt. Gegenüber der in der Informatik aus dem Compilerbau bekannten Definition von Parsing als die Erzeugung eines Ableitungsbaums für eine bzgl. einer formalen Grammatik wohlgeformten Eingabekette, hat der Begriff in der KI also eine erhebliche Bedeutungserweiterung erfahren. Bevor grammatisches Wissen zur Analyse einer natürlichsprachlichen Eingabe eingesetzt werden kann, muß die Eingabe zumindest morphologisch und lexikalisch voranalysiert sein. In Abhängigkeit von der gewählten Systemarchitektur (vgl. Abschnitt 2.2.), wird die morphologische und lexikalische Analyse entweder gleich zu Beginn für die gesamte Eingabekette durchgeführt oder sie wird inkrementell und mit der syntaktischen und semantischen Analyse verzahnt ausgeführt.

3.1. DIE KOMPONENTEN ZUR MORPHOLOGISCHEN UND LEXIKALISCHEN ANALYSE

Da die in NSS eingesetzten Verfahren zur morphologischen und lexikalischen Analyse hauptsächlich aus der Linguistischen Datenverarbeitung übernommen wurden und bisher nur wenige für das Paradigma der sprachorientierten KI-Forschung typische Methoden (z.B. im Bereich der Analyse von Syntagmen und Wortbildungen) bekannt sind, werden die für diesen Bereich der Analyse entwickelten Ansätze im folgenden nur kurz skizziert. Ein weiterer Bereich, der aus Platzgründen in der vorliegenden Einführung nicht in angemessener Weise behandelt werden kann, ist die für die Verarbeitung kontinuierlich gesprochener Eingaben notwendige *akustisch-phonetische* Analyse zusammen mit der

Segmentierung des Sprachsignals und der *Bildung von Worthypothesen* (vgl. z.B. Lea 1980).

Die morphologische und lexikalische Analyse erfolgt sowohl auf der Ebene von einzelnen Wörtern als auch von zusammenhängenden Wortgruppen (Fachwort: *Syntagmen*). Die Analyse von einzelnen Wörtern kann folgende Verarbeitungskomponenten umfassen:

- Lemmatisierung und Flexionsanalyse
- Synonym-Ersetzung
- Korrektur von Tipp- und Übertragungsfehlern
- Wortbildungsanalyse
- Klärungsdialog und natürlichsprachlicher Wissenserwerb

Ziel der *Lemmatisierung* ist es, zu jeder eingegebenen Wortform einen Lexikoneintrag (Fachwort: *Lemma*) zu finden. Neben grammatischer Information (z.B. Wortart, Flexionsklasse) kann ein Lexikoneintrag auch semantische Information (z.B. Kasusrahmen für Verben und Selektionsrestriktionen, vgl. Kapitel 4) und bei gesprochener Eingabe auch phonologische Information enthalten. Beispielsweise wird in (3) für die Worterkennung die Artikulationsweise von *groß* in einer mithilfe des ASCII-Zeichensatzes codierten Lautschrift als 'stimmhaftes G, Zäpfchen-R, geschlossener langer Vokal O, stimmloser

(3) GROSS ((AUSSPRACHE ".G.R.OH:.S.") (WORTART ADJ))

Konsonant S' angegeben und dem Lemma die Wortklasse Adjektiv zugeordnet.

Prinzipiell lassen sich zwei Arten der Lexikonorganisation unterscheiden: *Vollformenlexika* und *Grundformenlexika*. Im Vollformenlexikon sind alle Formen¹, in denen ein Wort auftreten kann, gespeichert (z.B. *parken, parkt, geparkte* usw.), so daß viel Speicherplatz aber wenig Verarbeitungsaufwand für die Lemmatisierung benötigt wird. Im Grundformenlexikon sind nur Grundformen gespeichert, so daß eingegebene Wörter durch morphologische Analyseprozeduren, in denen u.a. Endungen und Präfixe (z.B. *ge* in *geparkt*) abgetrennt und Umlaute rückgängig gemacht werden (z.B. *fällt* - *fallen*), zunächst auf Grundformen reduziert werden müssen. Diese für das Deutsche z.T. sehr aufwendigen Verfahren führen zu einer erheblichen Reduktion des Lexikonumfangs. Doch auch für ein einigermaßen vollständiges Grundformenlexikon der deutschen Gegenwartssprache ergeben sich Massendaten, wie die sechsbändige Ausgabe des Duden mit ihren ca. 500 000 Lexikoneinträgen zeigt. Die bisher in NSS verwendeten Lexika liegen in der Größenordnung zwischen 1000 und 20 000 Einträgen und sind damit vergleichsweise sehr klein, wenn man bedenkt, daß im Deutschen alleine rd. 20.000 Verben unterschieden werden können. Auch wenn das Wort-Lexikon diskursbereichsspezifisch gestaltet wird (z.B. für ein Expertensystem im Bereich der Chemie), können sich noch sehr große Lexika ergeben, da für Fachsprachen z.B. für die Chemie über 100 000 Fachausdrücke bekannt sind.

¹ Im Deutschen können z.B. für ein Verb im Mittel ca. 16 Formen unterschieden werden.

Unregelmäßig gebildete oder nur mit hohem Analyseaufwand reduzierbare Wortformen werden häufig auch bei der Verwendung eines Grundformenlexikons als Lemma aufgenommen. Beispielsweise wird im Lexikoneintrag (4) die Grundform (Fachwort: *Stamm*) und die Flexionsinformation '3. Person Singular Präsens' angegeben. Häufig kann die weitere Verarbeitung dadurch vereinfacht werden, daß eine Wortform durch eine im Lexikon

(4) IST ((STAMM SEIN) (FLEXION ((PRES (SG (3)))))).

verzeichnete synonyme Wortkette ersetzt wird (vgl. (5)).

(5) AM ((SYNONYM (AN DEM)))

Kann einem eingegebenen Wort auch nach der morphologischen Analyse kein Lexikoneintrag zugeordnet werden, so wird die in einigen NSS vorgesehene Komponente zur *Korrektur von Schreib- und Übertragungsfehlern* angestoßen (vgl. auch Husmann/Genzmann 1982). Wie ein Vergleich der fehlerhaften Eingabe (6) mit der korrigierten Version (7) zeigt,

(6) *Es führt überhaupt ein Weg daran vorbei, daß ien Tippfehler ien der Eingabbe erkannt wird.*

(7) *Es führt überhaupt kein Weg daran vorbei, daß ein Tippfehler in der Eingabe erkannt wird.*

sind dazu Permutationen, Löschungen und Ergänzungen von Zeichen notwendig. Beispiel (6) demonstriert auch, daß eine umfassende Korrektur nicht nur auf Ähnlichkeitsvergleichen (Eingabbe → Eingabe) beruhen kann, sondern *kontextsensitiv* auch syntaktisches und semantisches Wissen verwenden muß (*ien* → *ein* oder *in*). Besonders kompliziert sind Fälle, in denen erst eine semantische oder pragmatische Analyse der gesamten Äußerung ergibt, daß ein Eingabefehler vorliegen muß, da die fehlerhafte Wortform zufällig einem Lemma entspricht (z.B. *ein* → *kein*).

Obwohl zur Steigerung der *Robustheit* von Parsern z.B. in LIFER (Hendrix 1977), PARRY (Parkison et al. 1977) und FLEXP (Hayes/Carbonell 1981) spezielle Verfahren zur Tippfehlerkorrektur vorgesehen sind, existiert bisher kein NSS mit einer wissensbasierten Fehleranalyse, durch die (6) in (7) überführt werden kann.

Trotz korrekter Schreibweise kann natürlich auf für sog. *ad-hoc* Wortbildungen des Benutzers wie *Prominentenzimmer* in (9) kein Lexikoneintrag gefunden werden. Durch

(8) SYS: *Von unseren prominenten Gästen werden die Zimmer mit Alsterblick bevorzugt.*

(9) BEN: *Und was kostet so ein Prominentenzimmer?*

Segmentierung der Wortbildung und eine sich anschließende Lemmatisierung der Segmente kann die Komponente zur Wortbildungsanalyse unter Verwendung von morphologischem Wissen solchen ad-hoc Wortbildungen eine Wortklasse und Flexionsangaben zuweisen. Die Semantik von Wortbildungen wird erst in einer späteren Verarbeitungsphase bei der Überführung der Eingabe in eine semantische Repräsentationssprache rekonstruiert (vgl. Hoepfner 1982a).

Falls eine Lemmatisierung nach dem Versuch einer Fehlerkorrektur und Wortbildungsanalyse nicht gelingt, kann das NSS versuchen, in einem *Klärungsdialog* mit dem Benutzer (vgl. (10)-(12)), Information über das unbekannte Wort zu erhalten. Besonders problematisch für die lexikalische Analyse sind im Lexikon nicht enthaltene Eigennamen, da

(10) BEN: *Kann man sich in Ihrem Hotel leger kleiden?*

(11) SYS: *Was heißt 'leger'?*

(12) BEN: *Dasselbe wie ungezwungen.*

Rückfragen wie *Was heißt Handke?* auf Eingaben wie (13) blockiert werden müssen. Im

(13) BEN: *Mein Name ist Peter Handke. Ich suche ein Zimmer.*

vorliegenden Fall kann durch die Erkennung des Eigennamens als Teil der Wendung *Mein Name ist <Eigennamen>* eine Wortklassenzuweisung durchgeführt werden.

Wendungen werden während der *Syntagmen-Verarbeitung* erkannt, die folgende Prozesse umfassen kann (vgl. v. Hahn et al. 1980)

- Ersetzung von Mehrwortlexemen
- Verarbeitung von Idiomen
- Interpretation konventionalisierter Sprechakte.

Die Syntagmen-Verarbeitung beruht im wesentlichen auf der Anwendung von Produktionen, die aus einem Pattern und einer Transformationsanweisung bestehen und im Syntagmen-Lexikon gespeichert sind. Die Produktionen werden über ein Schlüsselwort angesprochen, das einem obligatorischen Bestandteil des Syntagmas entspricht. Beispielsweise wird das Mehrwortlexem *Bundesrepublik Deutschland* durch Anwendung von (14) in *BRD* überführt.

(14) (DEFPROP BUNDESREPUBLIK
(ERSETZEN (BUNDESREPUBLIK DEUTSCHLAND) (BRD))
IDIOME)

Auf die gleiche Weise können *feste Wortverbindungen* (z.B. *in Bewegung setzen* → *starten*) oder *Redewendungen* (z.B. *auf die Palme bringen* → *ärgern*) ersetzt werden. Schliesslich können Wortfolgen, denen fest ein illokutionärer Akt zugeordnet ist (z.B. *Auf Wiederhören* oder *Wie kommst Du darauf?*), durch Zugriff auf das Syntagmen-Lexikon ohne weitere syntaktische und semantische Analyse pragmatisch interpretiert werden (z.B. 'Werte Verabschiedungsprozedur aus' oder 'Rufe Erklärungskomponente auf').

Die von der lexikalischen und morphologischen Analyse erzeugten kanonischen Wortformen und Flexionsangaben werden dann vom Parser unter Verwendung grammatischen Wissens weiterverarbeitet.

3.2. EINE TYPOLOGIE FÜR PARSER

Bevor im folgenden einige in der sprachorientierten KI-Forschung entwickelte Parsing Verfahren eingeführt werden, sollen zunächst die wichtigsten Kriterien zur Klassifikation von Parsern für natürliche Sprache genannt werden (vgl. Wahlster 1979).

Neben den aus der theoretischen Informatik und dem Compilerbau bekannten Unterscheidungen

- Top-down-Parser vs. Bottom-up-Parser
- Deterministischer Parser vs. nicht-deterministischer Parser
- Parser mit Tiefensuche vs. Parser mit Breitensuche

sind zur Einordnung von Parsern für natürliche Sprache noch weitere Kriterien geeignet. Durch das in der KI entwickelte Verfahren des *Insel-Parsing* (vgl. Fig. 9) kommt zu der üblichen Unterscheidung von zwei möglichen Verarbeitungsrichtungen noch eine weitere Alternative hinzu, so daß sich insgesamt die folgenden vier zusätzlichen Unterscheidungskriterien ergeben:

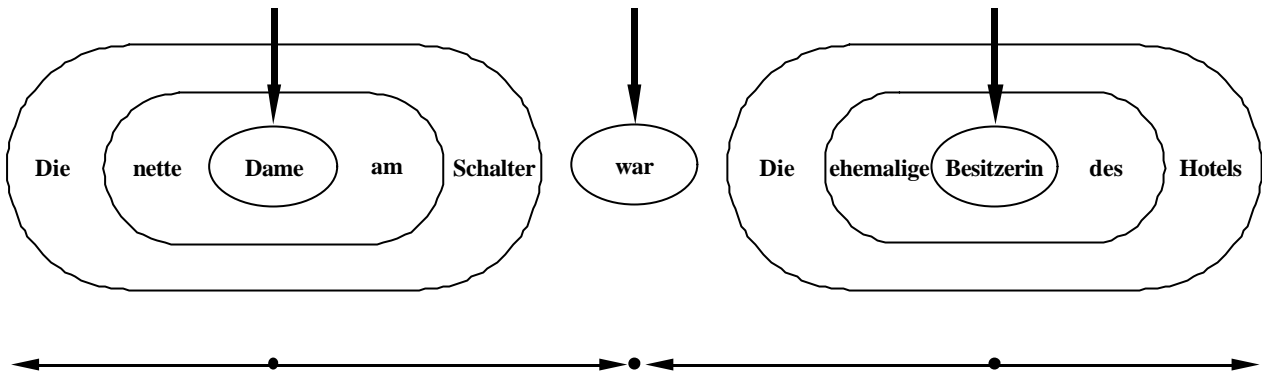


Fig. 9: Bidirektionales Insel-Parsing

- Verarbeitungsrichtung:

- Links-nach-Rechts-Parser

der natürlichen Leserichtung für das Deutsche entsprechende und am häufigsten verwendete Verarbeitungsrichtung

- Rechts-nach-Links-Parser

selten verwendete, aber z.B. für das Arabische natürliche Verarbeitungsrichtung

- Insel-Parser

besonders zum Parsing von gesprochener Sprache geeignete Technik (vgl. z.B. Bates 1980). Ausgehend von sicher erkannten Wort-Inseln (z.B. Nomen in einer NP und Verb, vgl. Fig. 9) in der Mitte der Eingabe wird solange nach links und rechts expandiert, bis eine Inselkette die gesamte Eingabe überdeckt (vgl. Fig. 9).

- Typ der erzeugten Repräsentationskonstruktion:

- Syntaktisch orientierte Parser

Ziel: Aufbau einer syntaktischen Strukturbeschreibung

- Semantisch orientierte Parser

Ziel: Aufbau einer semantischen Repräsentationskonstruktion

- Art der zur Wahl zwischen alternativen Ableitungen oder zur Blockierung von Pfaden verwendeten Information:
 - Syntaktisch gesteuerte Parser
Syntaktische Tests zur Wahl zwischen alternativen Ableitungen oder zur Blockierung von Pfaden
 - Semantisch gesteuerte Parser
Syntaktische und semantische Tests zur Wahl zwischen alternativen Ableitungen und zur Blockierung von Pfaden
 - Semantisch-pragmatisch gesteuerte Parser
Syntaktische, semantische und pragmatische Tests zur Wahl zwischen alternativen Ableitungen und zur Blockierung von Pfaden
- Interpretation der terminalen und nicht-terminalen Symbole der Grammatik:
 - Syntaktische Grammatik
Terminale und nicht-terminale Symbole der Grammatik werden als syntaktische Kategorien interpretiert.
 - Semantische Grammatik
Ein Teil der terminalen und nicht-terminalen Symbole der Grammatik wird als semantische Kategorien interpretiert.
 - Pragmatische Grammatik
Ein Teil der terminalen und nicht-terminalen Symbole der Grammatik wird als pragmatische Kategorie (z.B. Sprechakttyp) interpretiert.

Da für NSS meistens *Parser-Generatoren* verwendet werden, die erst zusammen mit einer als Parameter übergebenen Grammatik einen Parser ergeben, ist es sinnvoll, auch das zuletzt genannte Unterscheidungsmerkmal, das sich auf Grammatiken bezieht, zur Klassifikation von Parsern zu verwenden. Wird nämlich z.B. eine semantische Grammatik an einen Parser-Generator übergeben, so kann der daraus resultierende Parser als 'semantischer Parser' klassifiziert werden.

3.3. DIE REPRÄSENTATION VON GRAMMATISCHEM WISSEN IN AUGMENTED TRANSITION NETWORKS

Grammatisches Wissen wird in den meisten natürlichsprachlichen Systemen der Künstlichen Intelligenz (KI) in Form von ATNs (*erweiterte Übergangnetzwerke*, engl.: Augmented Transition Networks) repräsentiert. Die Grundformen des ATN-Formalismus (vgl. Woods 1970), die im folgenden eingeführt wird, wurde in zahlreichen Anwendungsprojekten in verschiedenen Richtungen z.T. stark erweitert und in eine komfortable Software-Umgebung (z.B. ATN-Trace, ATN-Editor, ATN-Compiler) eingebettet (vgl. die beiden Bände von Christaller und Metzger 1979, 1980).

Ein ATN kann als eine endliche Menge von endlichen Übergangsdiagrammen dargestellt werden. Jedes Übergangsdiagramm ist ein bewerteter gerichteter Graph mit benannten Knoten (als Zustände interpretiert), Kanten (als Zustandsübergänge interpretiert), einem ausgezeichneten Startzustand und einer Menge von ausgezeichneten Endzuständen. Das nichtterminale Symbol der Grammatik, mit dem der Startzustand eines Übergangsdiagrammes benannt ist, dient als Bezeichnung für das Übergangsdiagramm (im folgenden auch Netzwerk genannt). Jeder Zustandsübergang hat die Form ($\langle \text{Kantentyp} \rangle \langle \text{Argumente} \rangle^{1-n} \langle \text{Aktion} \rangle^{0-n}$). Durch eine spezielle Interpretation der Kantentypen SEEK (in anderen Implementationen auch als PUSH

bezeichnet) und SEND (in anderen Implementationen auch als POP bezeichnet) werden Übergänge zwischen Netzwerken und eine rekursive Auswertung von Netzwerken realisiert. Das erste Argument von SEEK entspricht dem *Startzustand* eines Netzwerkes. Stößt der ATN-Interpreter auf eine SEEK-Kante, so wird die aktuelle Konfiguration als oberster Eintrag in einem Kellerspeicher abgelegt und damit begonnen, das in der SEEK-Kante spezifizierte Netzwerk zu durchlaufen. Wird in diesem Netzwerk ein durch eine wegführende SEND-Kante markierter *Endzustand* erreicht, so wird das oberste Element des Kellerspeichers gelesen und reduziert, um das vorher unterbrochene Durchlaufen des aufrufenden Netzwerkes mithilfe der im Kellerspeicher gefundenen Konfigurationsbeschreibung fortzusetzen. In einem ATN gilt eine Eingabe als akzeptiert, sobald versucht wird, den leeren Kellerspeicher zu reduzieren, wenn der Eingabezeiger am Ende der Eingabe steht.

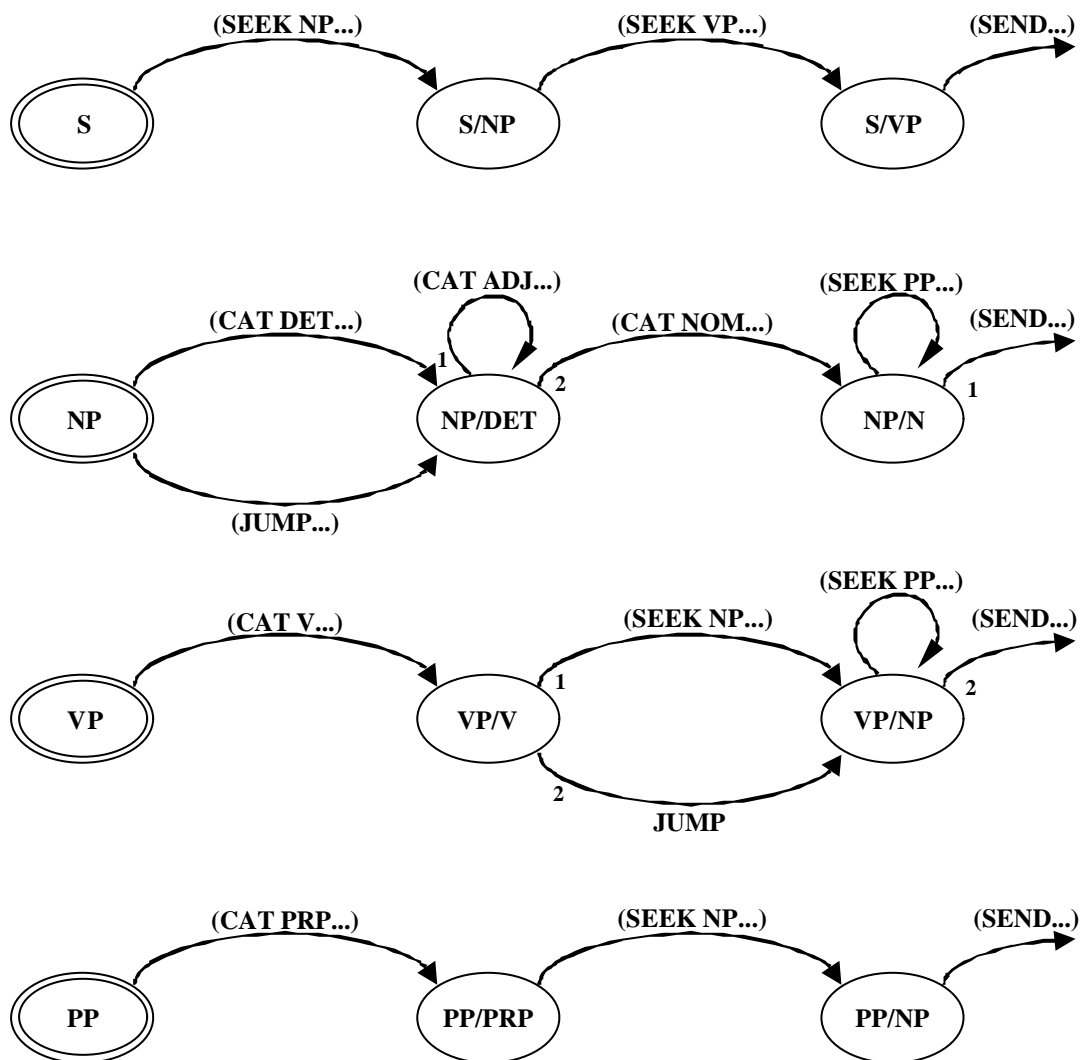


Fig. 10: Eine einfache ATN-Grammatik

Neben dem SEEK/SEND-Mechanismus besteht eine weitere Besonderheit des ATN-Konzeptes gegenüber den bekannten Übergangsdiagrammen endlicher Automaten darin, daß im Aktionsteil eines Zustandsübergangs Lese-/Schreiboperationen über einer unbeschränkten Menge von *zusätzlichen Registern* durchgeführt werden können. Die Register werden bezogen auf ein einzelnes Netzwerk wie lokale Variable

behandelt, die nur über spezielle Zugriffsoperationen mit einer Spezifikation des Bezugsnetzwerkes von einem anderen Netzwerk aus zugänglich sind.

Ein sehr kleiner Sprachausschnitt des Deutschen wird durch das in Fig. 10 dargestellte ATN beschrieben, in dem die Aktionsteile der Kantenbewertungen aus Gründen der Übersichtlichkeit nicht spezifiziert wurden. In diesem Beispiel für ein ATN werden die präterminalen Wortklassensymbole DET für Determinans (synonym: Artikel), ADJ für Adjektiv, N für Nomen, V für Verb und PRP für Präposition benutzt. Die nichtterminalen Symbole S für Satz, NP für Nominalphrase, VP für Verbalphrase und PP für Präpositionalphrase dienen zur Bezeichnung syntaktischer Konstituenten.

Zusätzlich zu den bereits eingeführten SEEK und SEND-Kanten enthält das ATN in Fig. 10 weitere Kanten von Typ CAT und JUMP. Bei der Auswertung einer Kante vom Typ CAT prüft der ATN-Interpreter zunächst, ob das aktuelle Eingabewort ein Element derjenigen Wortklassen ist, die als erstes Argument der CAT-Kante spezifiziert wurden. Falls diese Überprüfung erfolgreich verläuft, wird der Eingabezeiger auf das nächste Wort der Eingabekette weiterbewegt. Anderenfalls wird versucht, einen Zustandsübergang mithilfe einer anderen Kante zu erreichen, oder es muß ein *Backtracking-Prozeß* ausgelöst werden. JUMP bewirkt einen Zustandsübergang, ohne dabei den Eingabezeiger zu verändern. Startzustände sind in Fig. 10 durch doppelte Ellipsen und Endzustände durch wegführende SEND-Kanten markiert. Falls von einem Zustand mehr als eine Kante wegführt, wird durch Zahlen an den Kanten die Reihenfolge spezifiziert, nach der vom ATN-Interpreter versucht wird, die entsprechenden Zustandsübergänge zu realisieren.

Bei einer syntaktischen Analyse des ambigen Satzes (15) (vgl. auch Abschnitt 1.1.)

(15) *Die große Flügelschraube hält den Kühler am Sockel*

mithilfe der in Fig. 10 dargestellten ATN-Grammatik können beide Lesarten des Satzes akzeptiert werden. Zunächst wird für die Teilkette *Die große Flügelschraube hält den Kühler* in beiden Fällen die Zustandsfolge (16) durchlaufen.

(16) S-NP-NP/DET-NP/DET-NP/N-S/NP-VP-VP/V-NP-NP/DET-NP/N

Da vom Zustand NP/N aus zunächst die SEND-Kante verfolgt wird, lautet die zur Akzeptanz der Eingabeführende Fortsetzung der Zustandsfolge (16):

(17) VP/NP-PP-PP/PRP-NP-NP/DET-NP/N-PP/NP-VP/NP-S/VP

Falls im Rahmen einer Breitensuche oder durch Auslösung eines Backtracking-Prozesses in der von S/VP wegführenden SEND-Kante nach alternativen Erfolgspfaden gesucht wird, ergibt sich als alternative Fortsetzung der Zustandsfolge (16)¹ die Folge:

(18) PP-PP/PRP-NP-NP/DET-NP/N-PP/NP-NP/N-VP/NP-S/VP

Während durch (18) die Präpositionalphrase am Sockel in eine Nominalphrase eingebettet wird, bildet diese Präpositionalphrase in (17) die dritte Konstituente der Verbalphrase. Das Analyseergebnis hängt also davon ab, ob die Präpositionalphrase durch die von NP/N oder von VP/NP wegführende (SEEK PP...)-Kante verarbeitet wird.

¹ Durch Vertauschung der Auswertungsreihenfolge für die von NP/N wegführenden Kanten ändert sich die Reihenfolge, in der beide Lesarten gefunden werden.

Die von dem ATN in Fig. 10 akzeptierten Beispielsätze (19)-(21) zusammen mit

- (19) *Die Maschine arbeitet*
- (20) *Fritz schläft*
- (21) *Der Drucker mit der schönen kleinen Schrift steht hinter verschlossenen Türen*

den zurückgewiesenen Sätzen (22)-(25) verdeutlichen den begrenzten Umfang des erfaßten Sprachausschnittes. Während Satz (25) tatsächlich syntaktisch nicht

- (22) *Die Maschine arbeitet nicht*
- (23) *Schläft Fritz*
- (24) *Hinter verschlossenen Türen steht der Drucker mit der schönen kleinen Schrift*
- (25) **Der mit der schönen Schrift hinter Türen Drucker den kleinen steht verschlossenen*

wohlgeformt ist, müßten die Sätze (22)-(24) von einer umfassenderen Grammatik analysiert werden.

In starkem Kontrast zu den in diesem Abschnitt eingeführten ATN-Parsern, die bei der Tiefensuche nach einem Ableitungspfad häufig Backtracking verwenden, stehen *deterministische Parser* wie PARSIFAL (vgl. Marcus 1980) und YAP (vgl. Church 1980). Durch die Berücksichtigung von Restriktionen, denen vermutlich auch die Parsing-Prozesse bei einem menschlichen Hörer unterliegen, wird mit der Entwicklung deterministischer Parser versucht, extrem effiziente aber dennoch möglichst leistungsfähige syntaktisch orientierte Parser zu konstruieren. Deterministische Parser haben wie der menschliche Hörer Schwierigkeiten mit sog. *Garden Path* ('Holzweg') Sätzen wie (26), in denen einer Teilkette wie (27) zunächst eine Lesart zugeordnet wird, die sich bei der Verarbeitung der

- (26) *Bei der Geldübergabe wird er erkannt und verhaftet sofort den Verdächtigen.*
- (27) *Bei der Geldübergabe wird er erkannt und verhaftet.*

restlichen Eingabekette als falsch erweist. Church berücksichtigt in YAP zusätzlich zur *Determinismus-Hypothese* von PARSIFAL noch die *Hypothese vom festen Speicherplatzbedarf*. Dadurch können z.B. Relativsätze wie (28) mit einer auch für den menschlichen Hörer unakzeptablen Einbettungstiefe von YAP nicht akzeptiert werden. Die

- (28) *Der Hund, welcher die Katze, welche die Maus jagte, anbellte, heißt Waldi.*

bisher implementierten deterministischen Parser sind trotz ihrer Effizienz für den Einsatz in anwendungsorientierten NSS noch nicht geeignet, da in ihnen bisher eine allgemeine Behandlung von Konjunktionen, Präpositionalergänzungen und lexikalischen Mehrdeutigkeiten fehlt (vgl. auch Winograd 1983, S. 410).

3.4. DIE REALISIERUNG DES BLACKBOARD-MODELLS IN CHART-PARSERN

Zur Realisierung des Verarbeitungsabschnitts 'Analyse' in Form eines Blackboard-Modells (vgl. Abschnitt 2.2.) wird als Informationstafel häufig eine als *Chart* bezeichnete Datenstruktur eingesetzt (vgl. Fig. 11), in der alle Zwischenergebnisse der Analyse abgelegt werden. Eine Chart ist ein Netzwerk aus Knoten, zwischen denen die einzelnen Wörter der Eingabekette stehen, und bewerteten Kanten, welche die überspannte Teilkette der Eingabe als syntaktische Konstituente kennzeichnen. Fig. 11 zeigt einen Auszug¹

¹ Aus Gründen der Übersichtlichkeit sind u.a. die Wortklassenkanten in Fig. 11 nicht enthalten.

aus einer Chart, wie sie für Beispielsatz (15) bei Anwendung einer Breitensuche auf das in Fig. 10 dargestellte ATN entsteht.

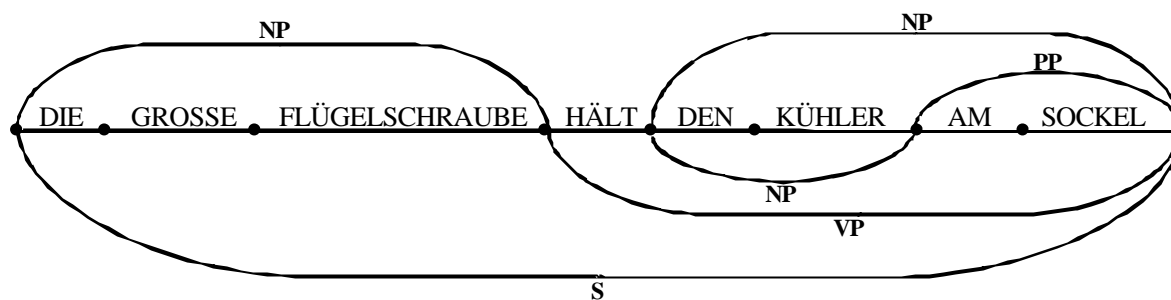


Fig. 11: Beispiel für eine Chart

Gegenüber der üblichen Darstellung syntaktischer Strukturen durch Ableitungsbäume hat die Chart u.a. den Vorteil, daß für mehrdeutige Strukturen nicht zwei vollständige Bäume mit z.T. identischen Teilstrukturen erstellt werden, sondern wie in Fig. 11 gemeinsame Teilstrukturen der Lesarten in der Chart verschmolzen werden. Um eine Chart durch ein ATN aufbauen zu lassen, werden im Aktionsteil der SEND-Kanten Chart-Kanten aufgebaut, die den Namen des durch SEND verlassenen Übergangsdiagramms tragen und die durch das Übergangsdiagramm verarbeitete Teilkette überspannen.

Auf die Chart greifen im Sinne des Blackboard-Modells oft nicht nur syntaktische sondern auch semantische Prozesse zu, z.B. um semantisch nicht plausible Lesarten frühzeitig auszublenden. Falls die verwendete Grammatik viele Sackgassen oder Mehrdeutigkeiten enthält, können die notwendigen Backtracking-Prozesse durch Zugriff auf die in der Chart gespeicherten Analyseergebnisse von wohlgeformten Teilketten (in der ATN-Literatur häufig als *well-formed-substring table* (Abk.: WFST) bezeichnet) beschleunigt werden. Beispielsweise braucht das PP-Netzwerk in Fig. 10 zur Analyse der Präpositionalphrase am Sockel bei einer Suche nach sämtlichen möglichen Lesarten nur einmal durchlaufen zu werden, falls das Analyseergebnis nach dem ersten erfolgreichen Durchlauf gespeichert wurde.

Alle Parser, die eine Chart verwenden, werden in der sprachorientierten KI-Forschung zur Gruppe der Chart-Parser zusammengefaßt. Der erste Chart-Parser wurde als Bottom-up-Analysator für kontextfreie Grammatiken von Kay implementiert (der sog. MIND-Parser, vgl. Kay 1973).

Als *Software-Werkzeuge* zur Entwicklung von Chart-Parsern wurden die Systeme GSP (General Syntactic Processor, vgl. Kaplan 1973) und GLP (General Linguistic Processor, vgl. Görz/Beckstein 1980) entworfen. Da diese beiden Systeme nicht nur hinsichtlich des Lexikons und der Grammatik sondern auch in bezug auf die Parsing-Strategie (z.B. Top-downs vs. Bottom-up, Breitensuche vs. Tiefensuche, links-nach-rechts vs. rechts-nach-links vs. bidirektional) parametrisiert sind, bieten sie sich besonders zur *Emulation* und zum Vergleich verschiedener Parser an¹.

Neben einer Chart verfügen GSP und seine Erweiterung GLP als eine zweite Datenstruktur über eine *Agenda* von Koroutinen, die durch einen *zentralen Monitor* generiert und verwaltet werden. In GSP und GLP können durch Veränderungen an einem *Scheduler* zur Prozeß-synchronisation und an einem *Selektor* zur Auswahl der zu aktivierenden Koroutinen verschiedene Parsingstrategien für Ein- oder Mehrprozessorsysteme realisiert werden.

¹Es gibt u.a. GSP-Emulatoren für den MIND-Parser, den LNR-Parser und den ELI-Parser (vgl. Eisenstadt 1979). Auch der Parser des Systems GUS (vgl. Abschnitt 1.2.) wurde mithilfe von GSP entwickelt.

3.5. DIE VERZÄHNUNG VON VERARBEITUNGSEBENEN IN KASKADIERTEN ATNs

Eine andere Möglichkeit, die verschiedenen Komponenten des Verarbeitungsabschnitts 'Analyse' als Koroutinen zu realisieren, bieten die sog. *kaskadierten ATNs* (Abk.: CATN, vgl. Woods 1980). Ein CATN besteht aus einer endlichen Folge von hintereinandergeschalteten ATNs $M_1, M_2, \dots, M_{n-1}, M_n$, wobei für $1 \leq j \leq n-1$ die Ausgabe von M_j jeweils die Eingabe von M_{j+1} bildet. Die Eingabe von M_1 ist die Eingabe in die gesamte Kaskade (vgl. auch Abschnitt 2.2.) und die Ausgabe von M_n ist die Ausgabe der gesamten Kaskade. Ausgaben eines ATNs M_j mit $1 \leq j \leq n-1$ werden durch die Aktion TRANSMIT bewirkt. Eine TRANSMIT-Aktion in M_j übergibt ihr Argument als Eingabe an M_{j+1} . Falls eine Eingabe in M_{j+1} mit $j \geq 1$ nicht akzeptiert wird, so wird in M_j ein Backtracking-Prozeß ausgelöst. Eine Eingabe in ein CATN gilt als akzeptiert, wenn alle ATNs der Kaskade ein leeres Eingabeband aufweisen und sich gleichzeitig in einem ihrer Endzustände befinden.

Die Funktionsweise eines CATNs wird im folgenden am Beispiel einer zweistufigen Kaskade erläutert (vgl. Fig. 12), welche die Sprache $L_1 = \{a^n b^n c^n \mid n \geq 1\}$ akzeptiert (vgl. Woods 1980, S. 9). L_1 ist eine kontextsensitive Sprache, die nicht durch eine kontextfreie Grammatik erzeugt werden kann. Aus der theoretischen Informatik ist bekannt, daß kontextfreie Sprachen gegenüber Durchschnittsbildung nicht abgeschlossen sind. Mithilfe der beiden kontextfreien Sprachen $L_2 = \{a^n b^n c^m \mid n, m \geq 1\}$ und $L_3 = \{a^m b^n c^n \mid n, m \geq 1\}$ kann daher $L_1 = L_2 \cap L_3$ eine nicht kontextfreie Sprache gebildet werden.

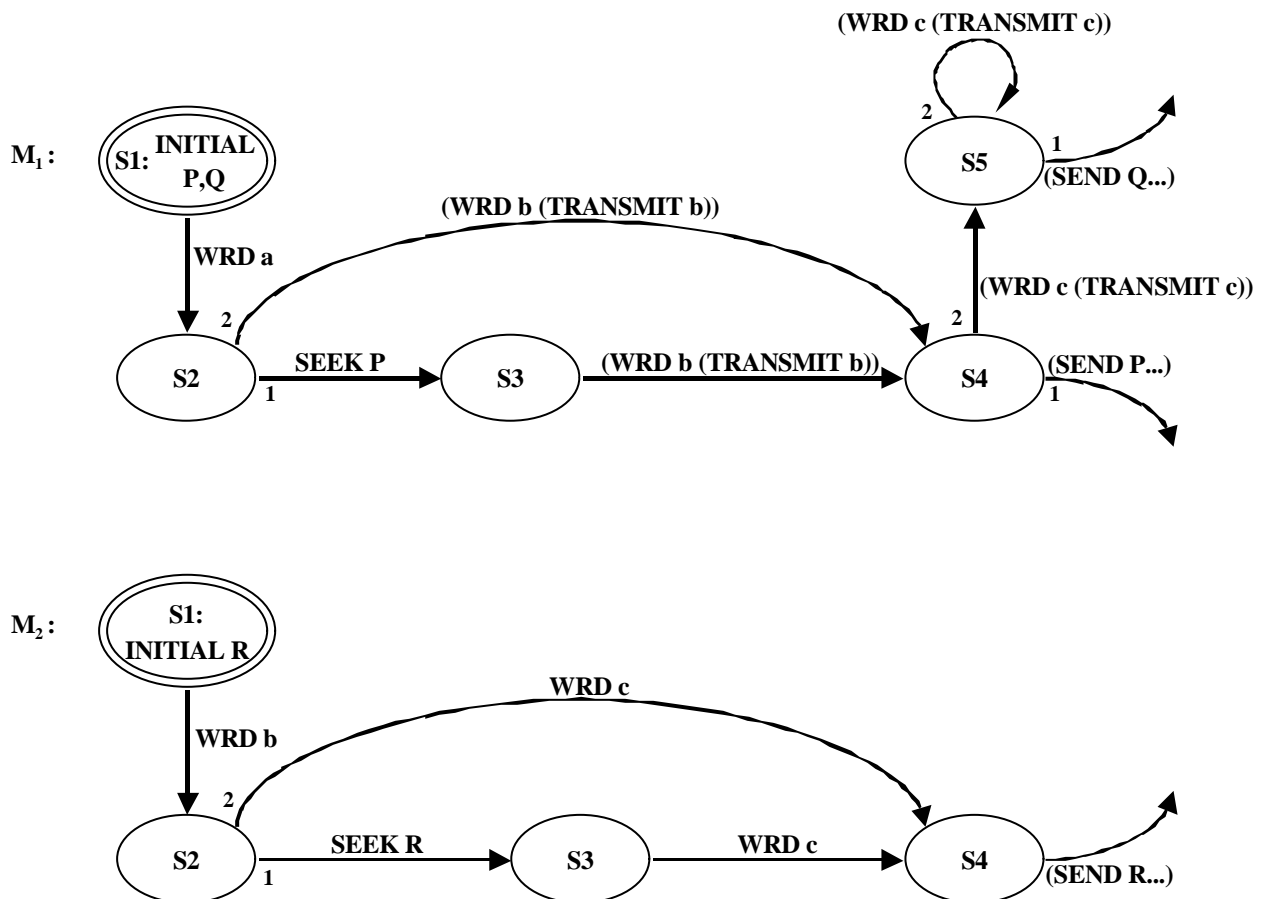


Fig. 12: Beispiel einer zweistufigen ATN-Kaskade

In dem für CATNs verwendeten ATN-Formalismus werden getrennte Bezeichnungen für den Startzustand eines Netzwerkes und die von dem Netzwerk akzeptierte Konstituente verwendet, so daß z.B. die Erkennung unterschiedlicher Konstituenten (in Fig. 12: P und Q) von dem gleichen Startzustand (in Fig. 12: S1) ausgehen kann. Die SEND-Kanten müssen dann allerdings immer den Namen der erkannten Konstituente enthalten, um verschiedene Endzustände eindeutig erkannten Konstituenten zuordnen zu können (vgl. SEND P, SEND Q in Fig. 12).

Durch M_1 wird die Grammatik G_1 und durch M_2 die Grammatik G_2 realisiert:

$$G_1: Q \rightarrow RS_5, S_5 \rightarrow cS_5, S_5 \rightarrow c, P \rightarrow aPb, P \rightarrow ab$$

$$G_2: R \rightarrow bRc, R \rightarrow bc$$

Es gilt $L(G_1) = L_2$, aber $L(G_2) = \{b^n c^n \mid n \geq 1\} = L_3$. Der Unterschied des CATN gegenüber der formalen Durchschnittsbildung mit L_2 und L_3 beruht darauf, daß eine TRANSMIT-Aktion mit den von M_1 verarbeiteten 'a'-Symbolen für die Erkennung von L_1 durch die gesamte Kaskade überflüssig ist, da alle bzgl. L_1 an 'a' geknüpften Bedingungen bereits von M_1 geprüft werden. Die Zeichenkette $aabbcc \in L$ wird durch die in Fig. 12 dargestellte Kaskade folgendermaßen akzeptiert:

$$M_1:S1-S2-S1-S2-S4, M_2:S1-S2, M_1:S3-S4, M_2:S1-S2, M_1:S5, M_2:S4-S3, M_1:S5, M_2:S4$$

Da bereits ein ATN, in dem Register verwendet werden, bzgl. des akzeptierten Sprachumfangs die Mächtigkeit einer Turing-Maschine besitzt, kann durch die Organisation von ATNs in einer Kaskade natürlich keine weitere Steigerung der Ausdrucksstärke erreicht werden. Vielmehr besteht der Vorteil von CATNs bei der Konstruktion von NSS darin, daß die Verzahnung morphologischer, syntaktischer, semantischer und pragmatischer Prozesse im Verarbeitungsabschnitt 'Analyse' in einem klar strukturierten und effizienten Verarbeitungsmodell erfaßt werden kann. Jeder Analyseebene wird dabei genau eine Stufe der Kaskade zugeordnet. Im Gegensatz zu nicht kaskadierten ATNs wird in CATNs Redundanz bei der Wissensrepräsentation dadurch verhindert, daß z.B. auf der semantischen Stufe nur genau eine Stelle existiert, an der z.B. Zeitangaben semantisch überprüft werden, während in einem nicht kaskadierten ATN die gleiche semantische Prüfung mit jedem einzelnen Zustand, in dem eine als Zeitangabe interpretierbare syntaktische Konstituente erkannt wird, verknüpft werden muß.

Das Kaskaden-Modell hat sich im RUS-Parser (vgl. Bobrow 1978), der mit einer zweistufigen Syntax/Semantik-Kaskade arbeitet, bereits bewährt. Ausgehend von den Erfahrungen beim Einsatz von CATNs zur Analyse von Dialogprotokollen wurde eine gegenüber der ursprünglichen Version von Woods stark verallgemeinerte Version von CATNs entwickelt (vgl. Christaller/Metzing 1982):

- Innerhalb einer Kaskade können Stufen übersprungen werden. Dadurch ergibt sich u.a. die Möglichkeit, *phatische Ausdrücke* wie *oh* und *aha* von der lexikalischen Stufe unter Umgehung der syntaktischen und semantischen Stufe direkt an die pragmatische Stufe weiterzugeben.
- Nicht jede Stufe einer Kaskade muß als ATN realisiert werden. Dadurch wird es in einem Kaskaden-Modell möglich, ATNs nur auf denjenigen Analysestufen einzusetzen, für die sie zur Wissensrepräsentation geeignet sind.

4. VERFAHREN ZUR SEMANTISCHEN ANALYSE NATÜRLICHER SPRACHE

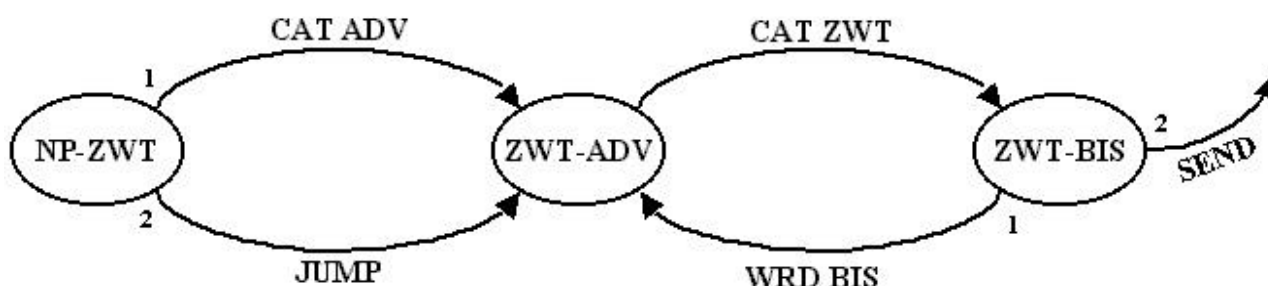
Im vorangegangenen Kapitel wurden Parser zunächst schwerpunktmäßig unter dem Aspekt des *Akzeptierens* bzw. *Zurückweisens* eingegebener Wortfolgen betrachtet. Da die Hauptaufgabe eines Parsers in einem NSS aber in der Überführung der Eingabe in eine für die Auswertung geeignete Repräsentationskonstruktion besteht, soll im folgenden anhand eines Beispiels aus einem *semantisch orientierten ATN-Parser*¹ ein einfaches Verfahren zur Realisierung der Überführungsfunktion erläutert werden.

4.1 DIE ERZEUGUNG VON REPRÄSENTATIONSKONSTRUKTIONEN DURCH SEMANTISCH ORIENTIERTE PARSER

Fig. 13 (a) zeigt einen winzigen Ausschnitt aus einem semantisch orientierten ATN, das in unserem System HAM-ANS eingesetzt wird, um bereits lexikalisch und morphologisch analysierte Eingaben in die oberflächennahe, logik-orientierte Repräsentationssprache SURF zu überführen (vgl. Jameson et al. 1981). Die Kanten des ATN enthalten die Wortklassensymbole ADV für Adverb und ZWT für Zahlwort. Bei der Auswertung des Kantentyps WRD durch den ATN-Interpreter wird überprüft, ob das aktuelle Wort mit dem ersten Argument der RD-Kante identisch ist.

In Teil (b) von Fig. 13 werden zu jedem in Teil (a) dargestellten Zustand die wegführenden Kanten mit den in der graphischen Darstellung nicht enthaltenen Aktionsteilen in der typischen Listendarstellung wiedergegeben. Durch Aktionen der Form (TO <Zustandsname>) wird zu dem spezifizierten Zustand übergegangen und der Eingabezeiger auf das nächste Wort der Eingabekette weiterbewegt. Bei der Auswertung der Form (VERIFY <LISP form>) durch den ATN-Interpreter wird das Argument von VERIFY als Prädikat gedeutet, dessen Evaluierung zu NIL die Blockierung des betreffenden Zustandsübergangs bewirkt (vgl. Zustand ZWT-BIS in Fig. 13 (b)).

Beim Durchlaufen des Teilnetzes wird mithilfe von ATN-Registern nacheinander eine *semantische Repräsentationskonstruktion* zusammengesetzt. Durch die Aktionen SETR und GETR² können ATN-Register gesetzt bzw. gelesen werden. Der Wert der Variablen \$ ist stets das aktuell zu analysierende Wort aus der Eingabekette. Im Register R-ZHECKE wird ein semantisch als Präzisions- bzw. Unschärfeoperator wirkendes Adverb (Fachwort: *Hecke*) gespeichert. Im Register R-NUMBER wird eine dem eingegebenen Zahlwort entsprechende natürliche Zahl³ oder ein der eingegebenen Formulierung entsprechendes Intervall abgelegt.



(a) Graphische Darstellung eines Teilnetzes zur Verarbeitung von Zahlwörtern

¹ Aus Platzgründen wird auf die Möglichkeiten zur Erzeugung einer expliziten syntaktischen Strukturbeschreibung (vgl. Bates 1980), wie sie in syntaktisch orientierten Parsern angestrebt wird, nicht näher eingegangen. Die entsprechenden Verfahren arbeiten analog zu dem hier skizzierten Vorgehen bei der direkten Erzeugung einer semantischen Repräsentationskonstruktion.

² Durch Makroexpansion wird '#A' in '(GETR A)' überführt.

³ %Zahlen-1 ist eine Assoziationsliste, in der diese Entsprechungen für die häufigsten Zahlwörter codiert sind.

```

NP-ZWT
(CAT ADV (SETR R-ZHECKE $) (TO ZWT-ADV))
(JUMP ZWT-ADV))

ZWT-ADV
((CAT ZWT
  (SETR R-NUMBER (COND [(NULL #R-NUMBER)(CDR (ASSOC $ %ZAHLEN-1))])
    [T (BUILD RANGE
      LOWER _#R-NUMBER
      UPPER (CDR (ASSOC $ %ZAHLEN-1))]))))
  (TO ZWT-BIS)))

ZWT-BIS
((WRD BIS (VERIFY (ATOM _#R-NUMBER)) (TO ZWT-ADV))
  (SEND (COND [_#R-ZHECKE
    (BUILD (QUANTITY HEDGE)
      HEDGE _#R-ZHECKE
      NUMBER-WORD _#R-NUMBER)]
    [T _#R-NUMBER])))

```

(b) Linearisierte Darstellung des semantisch-orientierten ATN-Teilnetzes

```

<quantity>
  ::= (qt-h: <hedge> <number-word>) ;HEDGE
  ::= <number-word>.
  ::= <range>.

<range>
  ::= (r: <number-word> ;LOWER <number-word> ;UPPER)

<hedge>
  ::= ETWA | GENAU | FAST | UNGEFAEHR | IN/ ETWA | NUR

<number-word>
  ::= <number>.
  ::= VIEL | EINIG | MEHRER | ETLICH | WENIG | E- | EPL | KEIN

<number>
  ::= (*LOCAL (*R ?N (NUMBERP !N))) ;FINITE
  ::= T ;INFINITE

```

(c) Auszug aus der syntaktischen Definition der semantischen Repräsentationssprache

```

(DREI ZWT)           ➔ 3
((UNGEFAEHR ADV) (ZWEI ZWT)) ➔ (qt-h: UNGEFAEHR 2)
((ZWEI ZWT) (BIS PRP) (SIEBEN ZWT)) ➔ (r: 2 7)
((ETWA ADV) (ZWEI ZWT) (BIS PRP) (DREI ZWT)) ➔ (qt-h: ETWA (r: 2 3))

```

(d) Beispiele für die Überführung in Repräsentationskonstruktionen

Fig. 13: Zur Funktionsweise eines semantisch orientierten ATN-Parsers

Die Funktion BUILD (vgl. die Zustände ZWT-ADV und ZWT-BIS in Fig. 13 (b)) dient zum Aufbau von Repräsentationskonstruktionen gemäß einer in erweiterter Backus-Naur-Form definierten Syntax. Fig. 13 (c) zeigt einen Auszug aus der syntaktischen Definition der Repräsentationssprache SURF, die in genau dieser Form von BUILD gelesen wird, um wohlgeformte SURF-Ausdrücke zu konstruieren.

Teil (d) von Fig. 13 enthält vier Beispiele für die ATN-gesteuerte Überführung in SURF-Ausdrücke, von denen das letzte kurz erläutert werden soll. Ausgehend vom Zustand NP-ZWT wird zunächst ETWA als Hecke erkannt und das Zahlwort ZWEI im Register R-NUMBER als 2 abgelegt. Über die WRD-Kante führt der Weg wiederum zum Zustand ZWT-ADV, von dem aus durch eine CAT-Kante das Zahlwort DREI verarbeitet wird. Dabei wird durch BUILD das Intervall (r: 2 3) aufgebaut. Anschließend wird die vom Zustand ZWT-BIS wegführende WRD-Kante durch die VERIFY-Aktion blockiert (R-NUMBER ist kein Atom mehr), so daß die SEND-Kante ausgewertet wird. Dabei wird ein Quantity-Ausdruck vom Typ 'Hedge' aufgebaut, dessen zwei Konstituenten (vgl. Fig. 13 c)) aus den Registerinhalten von R-ZHECKE und R-NUMBER zusammengesetzt werden (vgl. Fig. 13 (b)). Der von dem dargestellten Teilnetz erzeugte SURF-Ausdruck wird dann an das aufrufende ATN-Netzwerk weitergegeben, wo er in eine größere Repräsentationskonstruktion eingebettet wird.

4.2. EFFIZIENTES, DISKURSBEREICHSABHÄNGIGES PARSING MIT SEMANTISCHEN GRAMMATIKEN

Um in einem semantisch orientierten Parser die Überführung in eine Repräsentationskonstruktion zu vereinfachen, wurde für den schnellen Aufbau natürlichsprachlicher Schnittstellen mit stark eingeschränktem Diskursbereich das Konzept der *semantischen Grammatik* entwickelt.

Fig. 14 zeigt ein semantisches ATN mit zahlreichen WRD-Kanten, mit CAT-Kanten, die neben den üblichen lexikalischen Kategorien wie DET und ADJ auch die *diskursbereichsspezifische* semantische Kategorie ORT (z.B. können im Lexikon *Hamburg* und *Teisendorf* als Orte klassifiziert sein) enthalten, und mit SEND-Kanten für semantische Konstituenten wie PERSON (z.B. <Vorname> <Familiennamen> vs. <Titel> <Familiennamen> vs. <Anrede> <Familiennamen>) und ZEITPUNKT (z.B. <Wochentag> vs. <DATUM> vs. <Feiertag>). Die Nominalphrasen (29) und (30) sind Beispiele für Strukturen, die von dem in Fig. 14 dargestellten semantischen ATN akzeptiert werden.

(29) *eine abenteuerliche Reise von Dr. Otto Unblutig nach Teisendorf am Gründonnerstag*

(30) *die Reise von Hamburg nach Teisendorf über Hannover, Nürnberg, München am 15.3.82*

In anwendungsorientierten Systemen wie SOPHIE (Sophisticated Instructional Environment, vgl. Burton 1976), HWIM (Hear What I Mean, vgl. Woods et al. 1976) und PLANES (Programmed Language-based Enquiry System, vgl. Waltz 1978) konnten mit semantischen Grammatiken *hohe Verarbeitungsgeschwindigkeiten* beim Parsing erzielt werden (z.B. beträgt die Analysezeit für eine typische Eingabe in SOPHIE 0,1 - 1 Sekunde). Die im System LIFER (Language Interface Facility with Ellipsis and Recursion, vgl. Hendrix 1977) verfügbaren Hilfsmittel für den schnellen Aufbau semantischer Grammatiken haben sich in mehreren Anwendungsprojekten bewährt, in denen Bearbeiter mit guten Kenntnissen der Diskurswelt und ausreichenden Linguistik-Kenntnissen in wenigen Wochen einsatzfähige Parser erstellten.

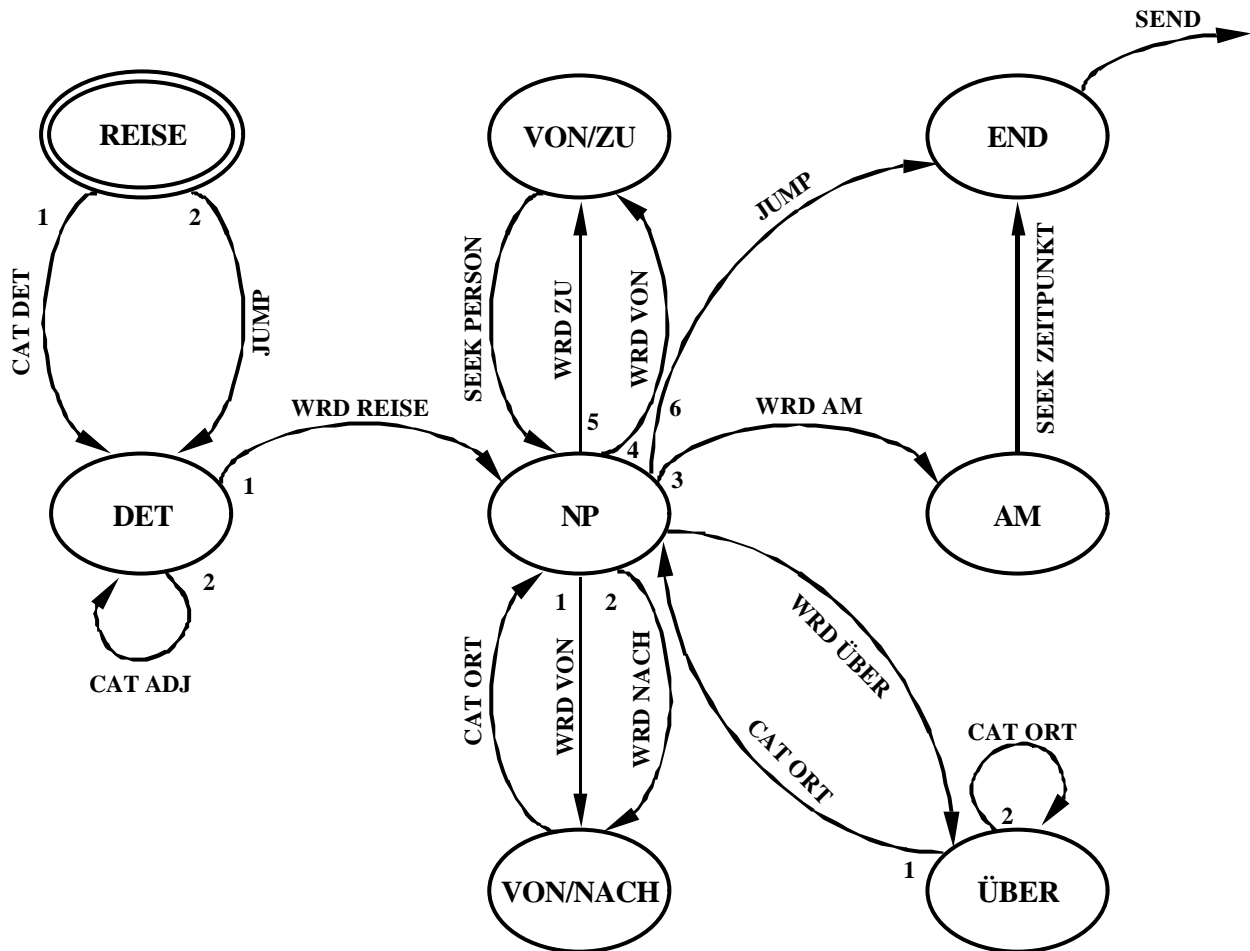


Fig. 14: Auszug aus einem semantischen ATN

Den Vorteilen einer einfachen Überführung in eine semantische Repräsentationskonstruktion stehen zwei Nachteile semantischer Grammatiken gegenüber:

- Semantische Grammatiken sind in der Regel wesentlich umfangreicher (und damit schwieriger zu verändern) als syntaktische Grammatiken, die den gleichen Sprachausschnitt erfassen. Beispielsweise wurde für HWIM eine semantische Grammatik mit 448 Zuständen, 881 Kanten und 2280 Aktionen entwickelt, die bzgl. der Vielfalt der akzeptierten Strukturen beschränkter war als eine für das gleiche System, entwickelte Grammatik mit 83 Zuständen, 202 Kanten und 386 Aktionen (vgl. Bates 1980).
- In semantischen Grammatiken wird grammatisches Wissen diskursbereichsspezifisch codiert, d.h. sie sind nur einmal verwendbar, da sie kaum an andere Diskursbereiche adaptiert werden können.

4.3. DIE SEMANTISCHE STEUERUNG DER ANALYSE DURCH KASUSRAHMEN

Eine Möglichkeit, Parser durch die Berücksichtigung semantischer Restriktionen effizient zu steuern, ohne dabei wie bei der Verwendung semantischer Grammatiken diskursbereichsabhängig zu werden, stellen die

sog. *kasusgesteuerten Parser* dar. In diesen Parsern werden sog. *Kasusrahmen*, die im Lexikon des Systems mit Verben assoziiert sind, zur Analysesteuerung eingesetzt. Im Kasusrahmen werden die mit der Verwendung eines Verbs obligatorisch und fakultativ vorhandenen semantischen Relationen zu anderen Bestandteilen einer Äußerung spezifiziert. Beispiele für solche semantischen Relationen sind:

- Agent/Actor: Urheber einer Tätigkeit oder Handlung
(z.B. Peter zerschlug das Glas.)
- Objective/Object: Von einer Handlung betroffene oder einbezogene Größe
(z.B. Peter zerschlug das Glas.)
- Instrumental/Instrument: Mittel der Handlung
(z.B. Peter zerschlug das Glas mit dem Hammer.)
- Source: Ausgangspunkt, Quelle oder Material einer Handlung
(z.B. Peter entfloh aus dem Gefängnis.)
- Goal: Zielpunkt der Handlung
(z.B. Peter entkam nach Südamerika.)

Solche semantischen Relationen werden als *Tiefenkasus* (vgl. Fillmore 1968) bezeichnet im Gegensatz zu den üblichen syntaktischen Unterscheidungen zwischen Nominativ, Akkusativ, Dativ usw., die als sog. *Oberflächenkasus* lediglich zur sprachlichen Markierung von Tiefenkasus dienen. In der sprachorientierten KI-Forschung werden sehr verschieden definierte Inventare von Tiefenkasus verwendet, wobei die Anzahl der unterschiedenen Tiefenkasus zwischen 7 und 20 schwankt (vgl. auch Winograd 1983, S. 311 ff.).

```
(agent: ( 0
         (d-o: OR
          (lambda: x1 (af-a: ISA x1 FAHRZEUG))
          (lambda: x1 (af-a: ISA x1 MENSCH))

source:
         ( F
          (lambda: x1 (af-a: ISA x1 VERKEHRSWEG))

goal:
         ( F
          (lambda: x1 (af-a: ISA x1 VERKEHRSWEG))
```

Fig. 15: Beispiel für einen Kasusrahmen

Der in Fig. 15 dargestellte Kasusrahmen für das Verb *Einbiegen*, der mithilfe der semantischen Repräsentationssprache SURF formuliert ist (vgl. auch Abschnitt 4.1.), wird im System HAM-ANS zur Kasussteuerung eines semantisch orientierten ATN-Parsers eingesetzt (vgl. Hoepfner 1982b). Neben den obligatorischen (0) und fakultativen (F) Tiefenkasus enthält der Kasusrahmen auch sog. *Selektionsrestriktionen*, die eine Konstituente erfüllen muß, um die entsprechende Kasusrolle übernehmen zu können. Beispielsweise muß in einem Ereignis vom Typ 'Einbiegen' der Agent ein Fahrzeug oder ein Mensch sein, bei Source und Goal muß es sich um Verkehrswege handeln (vgl. Fig. 15). Ohne die Verwendung solcher Selektionsrestriktionen kann ein NSS z.B. nicht erkennen, daß sowohl in (31) als auch in (32) *Edison* (und

nicht der Glühlampe) die Tiefenkasusrolle Agent zugewiesen werden muß.

(31) *Edison erfand die Glühlampe.*

(32) *Die Glühlampe erfand Edison.*

Fig. 16 zeigt einen Ausschnitt aus einem kasusgesteuerten ATN (vgl. Hoepfner 1982b), in dem die übliche Top-down-Strategie des ATN-Interpreters bottom-up durch den Kasusrahmen des Vollverbs modifiziert wird. Dabei wird für jeden im Kasusrahmen spezifizierten Tiefenkasus ein ATN-Register angelegt. Die Steuerung besteht nun z.B. darin, daß nur noch nach Präpositionalphrasen (z.B. für die Tiefenkasusrolle Source und Goal) gesucht wird, sobald durch VERIFY-Aktionen über dem Agent- und Objective-Register festgestellt wurde, daß alle durch Nominalphrasen realisierbaren Tiefenkasusrollen abgedeckt sind.

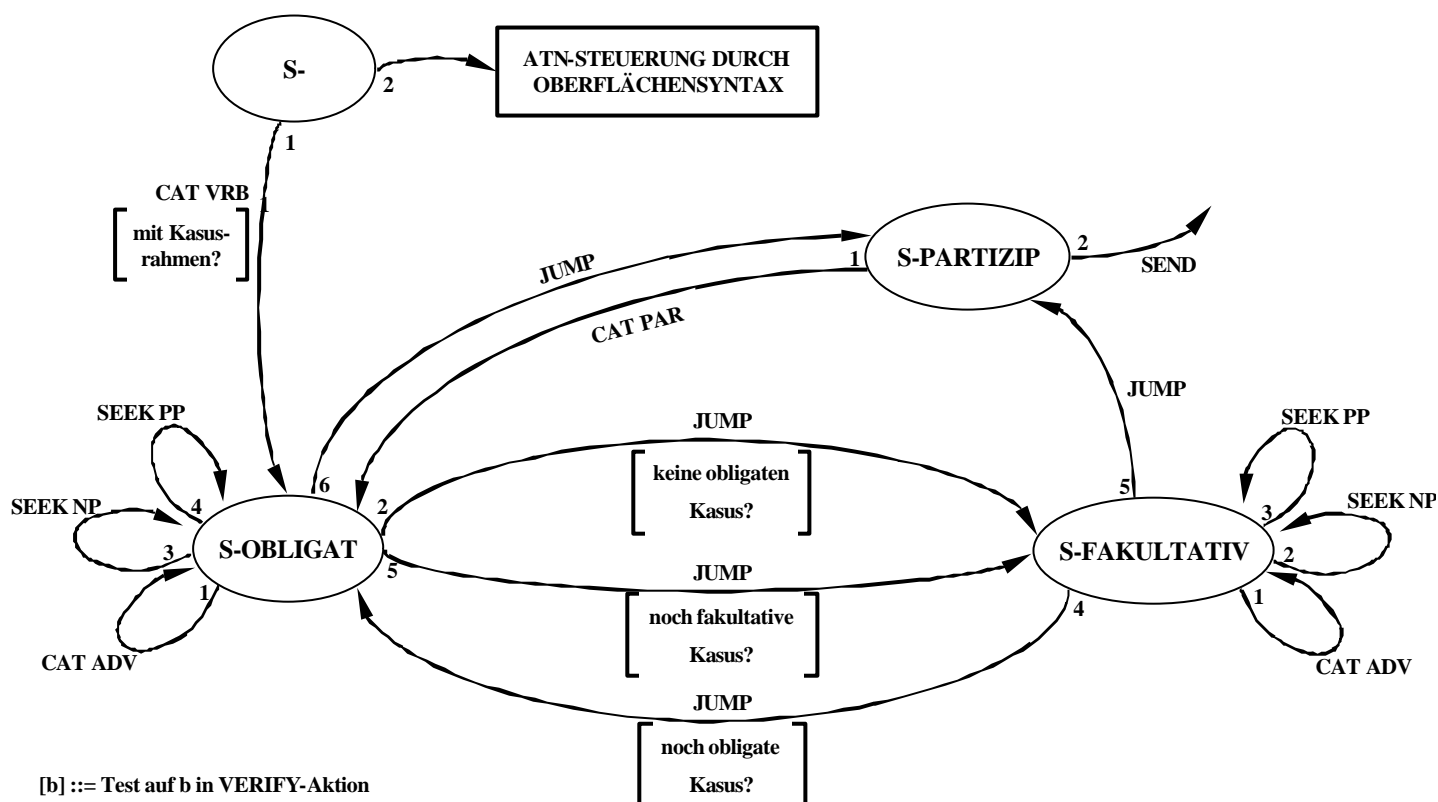


Fig. 16: Ausschnitt aus einem kasusgesteuerten ATN

Gesteuert durch den in Fig. 15 angegebenen Kasusrahmen können mithilfe des in Fig. 16 dargestellten ATNs (PAR steht für die Wortklasse Partizip) z.B. folgende Sätze analysiert werden:

(33) *Das Auto biegt ein.*

(34) *Der Betrunkene ist in eine Einbahnstraße eingebogen.*

(35) *Aus der Schlüterstraße ist danach der Polizist in die Einbahnstraße eingebogen*

4.4. DER VERZICHT AUF GRAMMATISCHE REGELN IN LEXIKON-BASIERTEN PARSERN

Die Verwendung von im Lexikon gespeicherter Information zur Steuerung der Analyse wurde in einer Serie von Parsern (MARGIE-Parser, vgl. Riesbeck 1975, ELI-Parser, vgl. Schank/Riesbeck 1976, CA-Parser, vgl. Birnbaum/Selfridge 1979), die alle an der Yale University entwickelt wurden, verallgemeinert und zum alleinigen Parsingprinzip erhoben. Bei diesen semantisch-pragmatisch gesteuerten und semantisch orientierten Parsern, die natürlichsprachliche Eingaben in sog. *konzeptuelle Dependenzgraphen* (vgl. Kapitel 5) überführen, spricht man auch von *lexikon-basierten Parsern* (z.T. auch von 'grammatikfreien' Parsern), da sie keine als grammatisches Wissen gespeicherten Regeln sondern ausschließlich Lexikoneinträge verwenden.

In lexikon-basierten Parsern soll Sprachverstehen als *erwartungsgesteuerter Prozeß* modelliert werden, d.h. daß beim Hören eines Wortes jeweils semantisch motivierte Erwartungen und Annahmen über nachfolgende Wörter oder bereits gehörte Wörter ausgelöst werden.

In NSS mit lexikon-basierten Parsern wird jedem Lexikoneintrag eine Menge von *Produktionen* zugeordnet, die jeweils aus einem *Bedingungs-* und einem *Aktionsteil* bestehen. Die wichtigsten Datenstrukturen, die vom Parser verwendet werden, sind eine Konzeptliste (C-LIST), in der die erzeugten Zwischenstrukturen gespeichert werden und die semantische Repräsentationskonstruktion sukzessive aufgebaut wird, und eine Liste der aktivierten Produktionen (R-LIST), die nach jedem Lesen eines Eingabewortes ausgewertet wird. Der Parsingalgorithmus kann folgendermaßen zusammengefaßt werden:

1. Lies von links nach rechts das nächste Wort der Eingabekette ein. Falls kein Wort mehr vorhanden ist, wird die Prozedur beendet.
2. Füge die mit dem eingegebenen Wort assoziierten Produktionen der R-LIST hinzu.
3. Werte die Produktionen in R-LIST aus.
4. Gehe zurück zu Schritt 1.

Im Bedingungsteil der Produktion kann geprüft werden, ob

- ein bestimmtes Wort oder eine bestimmte Konstituente vorhanden ist oder
- die in C-LIST gespeicherten Strukturen bestimmte semantische Bedingungen oder Ordnungsrelationen erfüllen.

Der Aktionsteil der Produktionen bewirkt:

- das Hinzufügen einer konzeptuellen Struktur an das Ende der C-LIST oder
- die Ersetzung eines Platzhalters in einer konzeptuellen Struktur durch eine andere Struktur oder
- die Speicherung von Produktionen in R-LIST oder
- die Löschung von Produktionen in R-LIST.

Anhand einer Beispielanalyse für den Satz *Fritz aß einen Apfel* soll im folgenden das Verfahren des lexikon-basierten Parsings in seinen Grundzügen erläutert werden. Zunächst wird das Wort *Fritz* eingelesen. Die mit diesem Wort assoziierte Produktion lautet:

P1: Bedingung: T

Aktion: Füge die Struktur (PP CLASS (MENSCH) NAME (FRITZ)) der C-LIST hinzu.

Durch die Auswertung von P1 in Schritt 3 des Parsingalgorithmus wird die angegebene Struktur in C-LIST gespeichert. Als nächstes wird das Wort *aß* eingelesen, mit dem die Produktionen P2, P3 und P4 assoziiert sind. Nach Schritt 3, in dem die Aktionsteile

P2: Bedingung: T

Aktion: Füge die Struktur (INGEST ACTOR (NIL) OBJECT (NIL) TIME (IMPERFEKT) der C-LIST hinzu.

P3: Bedingung: Gibt es vor der INGEST-Struktur auf der C-LIST eine konzeptuelle Struktur X vom Typ MENSCH?

Aktion: Ersetze in der INGEST-Struktur den Platzhalter NIL in der ACTOR-Position durch X.

P4: Bedingung: Gibt es nach der INGEST-Struktur auf der C-LIST eine konzeptuelle Struktur Y vom Typ NAHRUNG?

Aktion: Ersetze in der INGEST-Struktur den Platzhalter NIL in der OBJECT-Position durch Y.

von P2 und P3 ausgeführt werden, enthält C-LIST die Struktur (36). Bisher wurde die

```
(36) (INGEST ACTOR (PP CLASS (MENSCH) NAME (FRITZ))
      OBJECT (NIL)
      TIME (IMPERFEKT))
```

Erwartung noch nicht erfüllt, daß der Eingabesatz auch Angaben darüber enthält, was Fritz aß (vgl. P4). Als nächstes Wort wird einen eingelesen und P5 der R-LIST hinzugefügt. Da der Bedingungsteil von P5 nicht erfüllt ist, wird als nächstes das Wort

P5: Bedingung: Wurde eine neue Struktur Z an das Ende der C-List angehängt?

Aktion: Markiere Z als indefinite Beschreibung.

Apfel eingelesen und die Produktion P6 der C-LIST hinzugefügt.

P6: Bedingung: T

Aktion: Hänge die Struktur (PP CLASS (NAHRUNG) TYPE (APFEL)) an die C-LIST an.

Nun werden die Produktionen P6, P5 und P4 ausgeführt. Da kein weiteres Wort in der Eingabekette vorhanden ist, wird als semantische Repräsentationskonstruktion für *Fritz aß einen Apfel* die Struktur (37) ausgegeben. Innerhalb der Theorie der konzeptuellen

```
(37) (INGEST ACTOR (PP CLASS (MENSCH) NAME (FRITZ))
      OBJECT (PP CLASS (NAHRUNG) TYPE (APFEL) REF (INDEF)))
      TIME (IMPERFEKT))
```

Dependenz wird INGEST als sog. *semantisches Primitiv* (vgl. Kapitel 5) für den Handlungstyp 'Einverleiben eines Objektes durch ein Lebewesen' verwendet.

Das skizzierte Verfahren ermöglicht einerseits eine hohe Flexibilität bei der Mischung von Bottom-up- und Top-down-Verarbeitung, die z.B. bei der Auflösung von lexikalischen Mehrdeutigkeiten Vorteile bietet, weist aber andererseits folgende Nachteile auf:

- Da syntaktische Regularitäten nicht in getrennten Regeln erfaßt werden, entsteht eine hohe Redundanz im Lexikon, da die gleichen Regularitäten in vielen verschiedenen Lexikoneinträgen wiederholt werden müssen.
- Das für den Parsing-Prozeß verwendete Wissen ist schlecht lesbar.
- Jede Ergänzung und Änderung eines Lexikoneintrages kann viele, z.T. schwer vorhersehbare Seiteneffekte auf die Verarbeitung anderer Wörter auslösen.

Ein weiterer lexikon-basierter Parser ist der *Wortexperten-Parser* (vgl. Rieger/Small 81). Der Wortexperten-Parser¹, dessen Aufbau einer heterarischen Systemarchitektur (vgl. Abschnitt 2.2.) entspricht, ist semantisch orientiert und semantischpragmatisch gesteuert. Mit jedem Wort sind als Wortexperte Prozeßbeschreibungen assoziiert, durch die kontextsensitiv die passende Lesart des Wortes und seine funktionale Rolle in der Eingabekette bestimmt werden können. Das Einlesen eines Wortes aus der Eingabekette führt zur Aktivierung des assoziierten Wortexperten. Die aktivierten Wortexperten tauschen untereinander solange Nachrichten aus, bis eine eindeutige semantische Repräsentationskonstruktion für die gesamte Eingabe gefunden wurde.

4.5. WEITERE SEMANTISCH-PRAGMATISCHE PROZESSE IM VERARBEITUNGS-ABSCHNITT 'ANALYSE'

4.5.1. Die Rekonstruktion der Semantik von elliptischen Eingaben

In den vorangegangenen Abschnitten sind wir zunächst davon ausgegangen, daß der Benutzer nur syntaktisch vollständige natürlichsprachliche Äußerungen in ein NSS eingibt. Tatsächlich sind aber viele natürlichsprachliche Äußerungen *elliptisch*, d.h. es werden Wörter oder Satzteile ausgelassen, von denen der Sprecher annimmt, daß sie vom Hörer mitverstanden werden. Da die Effizienz der natürlichsprachlichen Kommunikation entscheidend von der Möglichkeit zur Verwendung von Ellipsen abhängt, wurden in der sprachorientierten KI-Forschung verschiedene Verfahren zur Verarbeitung elliptischer Eingaben entwickelt (z.B. enthalten die NSS LIFER, PLANES, GUS und HAM-ANS eine Komponente zur Verarbeitung von Ellipsen). Da die Aufgabe des Verarbeitungsabschnitts 'Analyse' darin besteht, eine vollständige semantische Repräsentationskonstruktion für die Eingabe zu erzeugen, muß in der Komponente zur *Ellipsenverarbeitung* (vgl. auch Fig. 7) die Semantik einer unvollständigen Eingabe unter Verwendung aller verfügbaren Wissensquellen rekonstruiert werden.

Aus der Linguistik ist eine Vielzahl unterschiedlicher Typen von Ellipsen bekannt, (vgl. z.B. Hussmann/Genzmann 1982), von denen im Rahmen dieser Einführung nur die in NSS besonders wichtigen Fälle angesprochen werden können. Eng verknüpft mit der Verwendung von Konjunktionen und Komparationen sind die sog. *syntaktischen Ellipsen*, die weiter in *Vorwärtsellipsen* (vgl. (38)) und *Rückwärtsellipsen* (vgl. (39)) untergliedert werden können. In (1) muß ein Links-nach-rechts-Parser den syntaktischen Kontext vorwärts

¹ Im Projekt TOPIC wird ein Wortexperten-Parser als Komponente eines Kondensierungssystems für deutschsprachige Texte entwickelt (vgl. Hahn et al. 1982)

(38) *Der große und der mittlere Sitzungssaal sind leider schon reserviert.*

(39) *Die Eltern bekommen das Doppelzimmer und der Sohn das Einzelzimmer.*

und in (39) rückwärts durchsuchen, um die fehlende Konstituente (in (38): *Sitzungssaal*, in (39): *bekommt*¹) ergänzen zu können. Die in der KI-Literatur beschriebenen Verfahren zur Verarbeitung syntaktischer Ellipsen beruhen auf verschiedenen Erweiterungen von Parsing- Algorithmen (u.a. durch Konjunktionen ausgelöste Generierung zusätzlicher ATN-Kanten, in denen bereits erkannte Konstituenten optional sind (vgl. Kwasny 1980), die allerdings bisher jeweils nur einen Teil dieser Kategorie von Ellipsen korrekt verarbeiten.

Im Gegensatz zu syntaktischen Ellipsen können die sog. *kotextuellen Ellipsen* nur durch Zugriff auf die als text- bzw. dialogbezogene Wissensquellen aufgebauten semantischen Dialoggedächtnisse (vgl. Abschnitt 2.1.) aufgelöst werden. Wie die Texteingabe (40) und die Dialogsequenz (41) - (43) veranschaulichen, müssen textverstehende Systeme und Dialogsysteme auf die semantische Struktur vorausgegangener Eingaben zugreifen, um den Teil der Bezugsäußerung zu finden, der durch die Ellipse ersetzt wird.

(40) BEN: *In den Familienzimmern stehen vier Betten. In den anderen Zimmern stehen jeweils nur zwei.*

(41) BEN: *Gibt es ein Radio im Zimmer?*

(42) SYS: *Ja, natürlich.*

(43) BEN: *Und ein Telephon?*

Man bezeichnet diese Form der kotextuellen Ellipse auch als *Ersetzungsellipse*. Andere wichtige Typen von kotextuellen Ellipsen sind die *Wiederholungsellipse* (vgl. (45)) und die *Erweiterungsellipse* (vgl. (46)). Die Verarbeitung kotextueller Ellipsen

(44) SYS: *Ein Zimmer mit Alsterblick kostet 120 DM.*

(45) BEN: *120 DM?*

(46) BEN: *Pro Nacht?*

setzt voraus, daß ein semantisch orientierter Parser auch für unvollständige Sätze eine semantische Repräsentationskonstruktion erzeugen kann. Ein Verfahren zur Ellipsenverarbeitung besteht dann darin, zunächst für den Bezugssatz alle zulässigen Elliptifizierungen zu erzeugen und nacheinander unter Verwendung eines Ähnlichkeitsmaßes und Gewichtungsfaktoren für semantische Strukturelemente mit der Repräsentationskonstruktion der Ellipse zu vergleichen. Wenn eine mögliche Elliptifizierung ein ausreichendes Maß an Ähnlichkeit mit einer vorliegenden Ersetzungsellipse aufweist, wird in der semantischen Repräsentationskonstruktion der ermittelte Teil des Bezugssatzes durch die Repräsentationskonstruktion für die Ellipse ersetzt (vgl. auch Jameson /Wahlster 1982).

¹ Dieses Beispiel zeigt auch, daß die elidierte Konstituente nicht in allen Merkmalen mit der Bezugsform übereinstimmen muß (z.B. Verb in 3. Person Singular vs. Plural).

4.5.2. Die referenzsemantische Analyse anaphorisch gebrauchter Pronomen und Nominalphrasen

Neben der Ellipsenverarbeitung stellt die referenzsemantische Analyse *anaphorisch gebrauchter Pronomen* und *Nominalphrasen*, die ähnlich wie Ellipsen Verweise auf den Kontext enthalten und damit kohärenzstiftend wirken, eines der Hauptprobleme der semantisch-pragmatischen Analyse dar. Grundvoraussetzung für die Auflösung anaphorischer Referenzen ist, daß von NSS als Teil der Wissensquelle *Fokus* (vgl. Abschnitt 2.1.) ein sog. *Vorerwähnheitsgedächtnis* angelegt wird, in dem nacheinander die Bezeichner für alle Denotate textuell vorausgegangener Beschreibungen zusammen mit Angaben über Numerus, Genus sowie syntaktischer Funktion und semantischer Struktur der verwendeten Beschreibung abgelegt werden. Zur Auflösung eines Pronomens wird dann solange rückwärts durch das Vorerwähnheitsgedächtnis gesucht, bis ein Eintrag für ein Denotat gefunden wird, der zumindest bzgl. Numerus und Genus mit dem Pronomen übereinstimmt. Wenn wir

(47) SYS: *Die Lampe steht neben dem Schreibtisch.*

(48) BEN: *Ist er groß?*

(49) BEN: *Ist sie groß?*

annehmen, daß als Seiteneffekt der referenzsemantischen Analyse von (47) LAMPE2 und SCHREIBTISCH1 zusammen mit den entsprechenden Angaben im Vorerwähnheitsgedächtnis abgelegt werden, ergibt die durch das Auftreten von Pronomen in (48) und (49) ausgelöste Suche im Vorerwähnheitsgedächtnis, daß *er* auf SCHREIBTISCH1 und *sie* auf LAMPE2 bezogen wird.

Während nach Hobbs 1976 im Englischen¹ durchschnittlich rd. 85% der vorkommenden Pronominalreferenzen mithilfe des skizzierten Verfahrens auf recht einfache Weise korrekt aufgelöst werden können, sind die restlichen Fälle nur mit relativ hohem Aufwand und unter Verwendung komplexer Heuristiken referenzsemantisch analysierbar. Beispielsweise versagt das beschriebene Verfahren (*er* und *ein Hund* werden für koreferent gehalten), bei der Auflösung von *er* in (51), wenn nicht zusätzlich noch *Selektionsrestriktionen* ('Nur Menschen, nicht aber Tiere können bewaffnet sein') überprüft werden. In (52) können die Selektionsrestriktionen des Verbs *bellen* dazu verwendet werden, die allein

(50) SYS: *Unser Nachtportier hat einen Hund bei sich.*

(51) BEN: *Ist er bewaffnet?*

(52) BEN: *Bellt er häufig?*

aufgrund von Genus und Numerus nicht eindeutig bestimmbare Referenz auf den Hund zu erkennen. Daß auch die Auswertung von Selektionsrestriktionen nicht immer zur Bestimmung des intendierten Referenzobjektes ausreicht, sondern z.B. auch mit Wissen über Verhaltensmuster verknüpfte *Erwartungswerte* berücksichtigt werden müssen, zeigt ein Vergleich von (53) und (54). Aufgrund von Inferenzregeln, die auf alltäglichen Erfahrungen

(53) *Doris wollte in die Disco, aber Mutter sagte, sie sei zu jung.*

(54) *Doris wollte in den Zoo, aber Mutter sagte, sie habe keine Zeit.*

¹ Für das Deutsche dürfte dieser Wert noch höher liegen, da durch Auswertung von Genusinformation noch häufiger als im Englischen Mehrdeutigkeiten verhindert werden können (z.B. lautet die engl. Übersetzung von (48) und (49) gleich, da jeweils das Pronomen *it* verwendet wird.)

basieren, hält der Hörer es für wahrscheinlich, daß sich *sie* in (53) auf Doris und in (54) auf die Mutter bezieht.

Schließlich sei darauf hingewiesen, daß in einigen Fällen auch die Berücksichtigung der subjektiven Wahrscheinlichkeiten nicht ausreicht, sondern zusätzlich die *Relevanz* der möglichen Lesarten für den Dialog bzw. Text als Kriterium herangezogen werden muß. Fauser 1979 demonstriert dies am Beispielpaar (55) - (56), in dem er jeweils als koreferent zu Karl verstanden wird, obwohl in (56) die subjektive Wahrscheinlichkeit dafür,

(55) *Hans beschimpft Karl. Er gerät daraufhin in Wut.*

(56) *Hans beschimpft Karl. Er gerät daraufhin nicht in Wut.*

in Wut zu geraten, für Hans höher ist als für Karl. Nur falls ein NSS mit Wissen über Handlungsschemata ausgestattet wird (z.B.: wenn jemand beleidigt wird, reagiert er anschließend darauf, vgl. auch Kapitel 5), kann es ableiten, daß Karls Reaktion auf die Beleidigung im Zentrum des Interesses steht, und *er* in beiden Fällen auf Karl beziehen.

Obwohl fast alle neueren NSS über eine Komponente zur Pronomenauflösung verfügen und dabei einige der genannten Heuristiken verwenden, ist bisher kein System bekannt, da in Realzeit alle eingegebenen Pronomen korrekt auflöst. In anwendungsorientierten Dialogsystemen wird daher bei Schwierigkeiten mit der Pronomenauflösung von der Möglichkeit eines *Klärungsdialogs* Gebrauch gemacht, in dem das System zusammen mit dem Benutzer eine Klärung der Pronominalreferenz herbeiführt (vgl. auch v. Hahn et al. 1980).

Zur Auflösung *anaphorischer Nominalphrasen* muß der Fokus eines NSS mehr als die im Vorerwähnheitsgedächtnis gespeicherten Angaben enthalten (vgl. auch Webber 1978, Sidner 1979). Um die Nominalphrase *die Vorderräder* in (57) anaphorisch als die Vorderräder

(57) *Der weiße Mercedes parkt falsch. Die Vorderräder versperren den Radweg.*

des weißen Mercedes deuten zu können, muß die Komponente zur Referenzanalyse dafür sorgen, daß zusätzlich zum Verweis auf das Denotat für *der weiße Mercedes* im Vorerwähnheitsgedächtnis die Vorderräder zusammen mit anderen wichtigen Teilen des angesprochenen Autos in den Fokus gebracht werden. Eine solche *Ausweitung des Fokus* beruht auf begrifflichem Wissen über Bestandteile von Objekten und damit assoziierten Inferenzregeln (z.B. Transitivität der Teil-von-Relation), die zum allgemeinen Hintergrundwissen eines NSS gehören (vgl. Abschnitt 2.1.).

4.5.3. Zur Präsuppositionsanalyse und generischen Interpretation von definiten Nominalphrasen

Auf die vielfältigen und teilweise noch ungelösten Probleme bei der Referenzanalyse nicht-anaphorischer definiten und indefiniten Nominalphrasen kann im Rahmen dieser Einführung aus Platzgründen nicht eingegangen werden. Stattdessen wird in Kapitel 6 als Beispiel für semantisch-pragmatische Prozesse im Verarbeitungsabschnitt 'Generierung' die Erzeugung definiten und indefiniten Nominalphrasen behandelt. Auf zwei der wichtigsten Problemkreise im Bereich der Referenzsemantik, die in dieser Form nur bei der Analyse und nicht bei der Generierung von Deskriptionen auftreten, sei dennoch kurz hingewiesen.

Die Verwendung einer nicht-anaphorischen definiten Nominalphrase wie *die berühmte Forschungsstelle am Mittelweg* in (58) ist mit der impliziten Voraussetzung (Fachwort:

(58) *Liegt Ihr Hotel in der Nähe der berühmten Forschungsstelle am Mittelweg?*

Präsupposition) des Sprechers verbunden, daß in der Diskurswelt ein Objekt existiert, auf das die angegebene Beschreibung zutrifft. Außerdem verbindet der Sprecher mit der Äußerung der definiten Nominalphrase u.a. die Annahme, daß der Hörer das intendierte Objekt kennt und es aufgrund der angegebenen Beschreibung eindeutig identifizieren kann. Es ist daher eine wesentliche Aufgabe der Referenzanalyse so früh wie möglich die mit definiten Nominalphrasen verbundenen Präsuppositionen zu überprüfen und ggf. zurückzuweisen (vgl. (59)) oder in einem Klärungsdialog eine Disambiguierung der Referenz herbeizuführen (vgl. (60), (61)).

(59) *Ich kenne keine berühmte Forschungsstelle am Mittelweg.*

(60) *Welche Forschungsstelle meinen Sie?*

(61) *Meinen Sie die Forschungsstelle KI?*

Durch die frühzeitige Präsuppositionsüberprüfung kann verhindert werden, daß im Falle von *Präsuppositionsverletzungen* unnötiger Verarbeitungsaufwand durch aufwendige Such- und Inferenzprozesse bei der von vornherein zum Scheitern verurteilten Auswertung der Eingabe entsteht oder gar aufgrund einer gescheiterten Auswertung eine für den Benutzer irreführende Antwort erzeugt wird (z.B. *Nein* als Antwort auf (58)). Eng verbunden mit der frühzeitigen Präsuppositionsprüfung ist als zweites wichtiges Problem die *generische Interpretation* von Nominalphrasen mit bestimmtem Artikel. Es muß z.B. verhindert werden, dass die Referenzanalyse für die Eingabe (62) eine Rückfrage wie *Welches moderne Hotel meinen Sie?* generiert, da die Nominalphrase in diesem Fall generisch, d.h. wie *aller modernen Hotels*, zu interpretieren ist. In der sprachorientierten KI-Forschung wurde

(62) *Ein anspruchsvolles Restaurant ist das Aushängeschild des modernen Hotels.*

eine Reihe von Heuristiken untersucht, mit denen generische und nicht-generische Nominalphrasen unterschieden werden können (vgl. z.B. Sidner 1979, v. Hahn et al. 1980). Beispielsweise kann ein NSS leicht feststellen, daß in (58) eine generische Interpretation für *die berühmte Forschungsstelle am Mittelweg* nicht sinnvoll ist, da sie als räumlicher Bezugspunkt für ein anderes Referenzobjekt (das Hotel in (58)) dienen soll.

5. REPRÄSENTATION UND AUSWERTUNG DER SEMANTIK NATÜRLICHSPRACHLICHER AUSDRÜCKE

Der Kern des Verarbeitungsabschnittes 'Auswertung' eines NSS ist stark von der jeweiligen Repräsentationssprache abhängig, die zur Darstellung der semantischen Struktur der natürlichsprachlichen Eingabe verwendet wird. Da im Rahmen der vorliegenden Einführung in die sprachorientierte KI-Forschung nicht zusätzlich die Ergebnisse und Methoden des umfangreichen KI -Teilgebietes *Repräsentation von Wissen* (vgl. z.B. Bobrow/Collins 1975, Brachman/Smith 1980) vermittelt werden können, muß sich der folgende Abschnitt darauf beschränken, auf einige für NSS typische Darstellungsprobleme, Repräsentations- und Auswertungstechniken hinzuweisen.

Zunächst ist festzustellen, daß nicht alle in der KI verwendeten Wissensrepräsentationssprachen zur Darstellung der Semantik von natürlichsprachlichen Eingaben geeignet sind. Beispielsweise sind *Produktionensysteme* und *Actor-Systeme* für diese spezielle Aufgabe ungeeignet, was nicht ausschließt, daß sie sich in anderen Komponenten eines NSS zur Codierung von Wissensquellen durchaus bewähren¹.

¹ Beispielsweise haben sich Produktionensysteme u.a. für das Syntagmen-Lexikon (vgl. Abschnitt 3.1.) und Actor-Systeme u.a. zur Darstellung von lexikalisch-pragmatischem Wissen (z.B. im Wortexperten-Parser, vgl. Abschnitt 4.4) bewährt.

Außerdem wirft die Darstellung der Semantik natürlichsprachlicher Eingaben einige Probleme auf, die sich bei der Wissensrepräsentation für bildverstehende Systeme, Expertensysteme oder Theorembeweiser nicht in dieser Weise stellen. Zu den schwierigsten und in den bisherigen NSS stets nur ansatzweise oder punktuell gelösten Darstellungsproblemen gehört die *formale Repräsentation der Semantik* von:

- Vagen Quantoren (z.B. *mehrere, viele, einige*)
- Frageoperatoren (z.B. *Warum, Wie*)
- Konjunktionen von Termen (z.B. *Peter und Maria heiraten* vs. *Peter und Maria gehen zur Uni*)
- natürlichsprachlichen Junktoren (z.B. *nicht X oder Y, nicht nur X sondern auch Y*)
- komplexen Quantifikationen (z.B. *ungefähr zehn von den rund 100 bis 200 Bestellungen*)
- Hecken (z.B. *fast alle, einigermaßen groß, außerordentlich teuer*)
- Modalpartikel (z.B. *Leider, Selbstverständlich, Hoffentlich*)
- Komparativkonstruktionen (z.B. *Der Sieg beruhte mehr auf Kraft als auf Können*)
- deiktischen Ausdrücken (z.B. *hier, jetzt, später*)
- Kontinua (z.B. *Wasser, Verkehr, Musik*)

Die in NSS zum Aufbau semantischer Repräsentationskonstruktionen verwendeten Darstellungstechniken lassen sich in folgende vier Gruppen einteilen:

- Prozedurale Repräsentation mithilfe von KI-Programmiersprachen (z.B. PLANNER (vgl. Hewitt 1971); FUZZY (vgl. LeFaivre 1977))
- Netz-basierte Repräsentation (z.B. Semantische Netze (vgl. Findler 1979), CD-Graphen, vgl. Abschnitt 5.2.)
- Frame-basierte Repräsentation (z.B. FRL (vgl. Roberts/Goldstein 1977), KRL (vgl. Bobrow/Winograd 1977), KL-ONE (vgl. Brachman 1978), AIMOS (vgl. Sridharan 1978))
- Logik-orientierte Repräsentation (z.B. die für deutschsprachige NSS entwickelten Sprachen KS (vgl. Kolvenbach et al. 1979), MSRL/SRL (vgl. Habel et al. 1980), SURF/DEEP (vgl. Jameson et al. 1980))

Trotz der unterschiedlichen Syntax und Semantik der genannten Klassen von Repräsentationssprachen treten bei der Darstellung der Semantik natürlichsprachlicher Eingaben in allen angeführten Sprachen bestimmte Grundkonstruktionen in entsprechend angepaßter Form immer wieder auf. Zu diesen Grundkonstruktionen gehört die *Kasusrahmen-Darstellung* für Verben (vgl. auch Abschnitt 4.3.). Da sich die Anwendung von *Frames* als semantische Repräsentationssprache fast ausschließlich auf die *Kasusrahmen-Darstellung* erstreckt, braucht im folgenden nicht näher auf diese Darstellungstechnik eingegangen zu werden.

5.1. DIE VERWENDUNG VON DATENBANKANFRAGESPRACHEN UND KI-PROGRAMMIERSPRACHEN IM VERARBEITUNGSABSCHNITT 'ANALYSE'

Bevor anhand von drei Beispielen auf wichtige Repräsentations- und Auswertungstechniken eingegangen wird, sei noch auf die Möglichkeit hingewiesen, die Semantik der natürlichsprachlichen Eingabe in einer konventionellen *Datenbankanfragesprache* oder einer KI-Programmiersprache zu erfassen.

Bei der Entwicklung von *natürlichsprachlichen Schnittstellen* (z.B. USL, vgl. Ott 1979) besteht häufig das Hauptziel des NSS darin, die Eingabe möglichst direkt und effizient in eine gegebene Datenbankanfragesprache zu übersetzen (z.B. ISBL und SQL in USL, vgl. (63) - (66)¹). Da diese Sprachen im Gegensatz zu einigen in der KI entworfenen Repräsentationssprachen nicht speziell als Zielsprachen für das Parsing natürlichsprachlicher Eingaben entwickelt wurden, ist eine direkte Überführung in eine Datenbankanfragesprache meist mit einer drastischen Beschränkung der möglichen semantischen Konstruktionen

(63) BEN: *Welcher Mitarbeiter des Managers Schmidt wohnt in Heidelberg?*

ISBL-Ausdruck:

((WOHN; C2 = 'HEIDELBERG')*(MITARBEITER * (MANAGER; C1 = 'SCHMIDT'; C2= C3); C1 = C3)%C1

(64) SYS: *Braun.*

(65) BEN: *Wer verdient mehr als sein Manager?*

ISBL-Ausdruck:

((VERDIEN * MANAGER; C1 = C3); C2 > C4); C1 = C6)%C1

(66) SYS: *Niemand.*

verbunden². Der Verarbeitungsabschnitt 'Auswertung' wird in solchen Systemen vollständig vom Datenbankverwaltungssystem übernommen, so daß auf die Semantik des natürlichsprachlichen Ausdrucks abgestimmte Auswertungstechniken (z.B. für vage Quantoren) von vornherein ausgeschlossen sind.

Am meisten Ähnlichkeit mit der genannten Verwendung von Datenbankanfragesprachen hat die u.a. in Winograds System SHRDLU (vgl. Winograd 1972) erfolgreich erprobte Technik *Ausdrücke einer KI-Programmiersprache* wie PLANNER als semantische Repräsentationssprache zu verwenden (vgl. Fig. 17³). Auch in diesem Fall ist der Auswertungsprozeß durch den vorgegebenen Interpreter der Programmiersprache bestimmt. Allerdings können im Gegensatz zu konventionellen Datenbankanfragesprachen alle für die Auswertung relevanten Inferenzregeln und Heuristiken ebenfalls in der KI-Programmiersprache formuliert werden. Außerdem sind KI-Programmiersprachen so flexibel, daß in ihnen ohne weiteres spezielle Auswertungsverfahren für natürlichsprachliche Konstrukte formuliert werden können.

¹ Die ISBL-Ausdrücke werden über einer relationalen Datenbank ausgewertet. Mit C_n wird die n-te Spalte einer Relation angesprochen und '*' entspricht dem Join-Operator, ';' dem Selektionsoperator und '%' dem Projektionsoperator.

² Beispielsweise lassen sich die oben angeführten Repräsentationsprobleme kaum in konventionellen Datenbankanfragesprachen lösen.

³ Die Syntax des PLANNER-Ausdrucks wurde etwas vereinfacht.

```

(FIND ALL ?X1 (X1)
 (GOAL (IS ?X1 PYRAMIDE))
 (FIND 3 ?X2 (X2)
 (GOAL (IS ?X2 BLOCK))
 (GOAL (STUETZT ?X2 ?X1))))

```

Fig. 17: PLANNER-Ausdruck für *Welche Pyramiden werden von drei Blöcken gestützt?*

Der wesentliche Nachteil dieser Form der *prozeduralen Repräsentation*, der dazu führte, daß diese Technik in neueren Systementwürfen kaum noch verwendet wird, besteht darin, daß die Repräsentationskonstruktionen aufgrund ihrer prozeduralen Anteile (z.B. FIND ALL, FIND 3, GOAL in Fig. 17) selbst nicht Gegenstand der im übrigen System ablaufenden Inferenzprozesse sein können. Dies hat sich aber in der neueren Forschung als notwendig erwiesen, da aus der Tatsache, daß z.B. eine bestimmte Frage vom Benutzer in einer bestimmten Form gestellt wurde, mithilfe von Inferenzregeln wichtige Konsequenzen¹ für das Partnermodell des Systems (vgl. Abschnitt 2.1.) und die Beantwortungsstrategie abgeleitet werden können. Im Gegensatz zu Datenbankanfragesprachen und KI-Programmiersprachen wurden die meisten netz-basierten Repräsentationssprachen, die im folgenden am Beispiel der CD-Graphen eingeführt werden, speziell in Hinblick auf die semantische Repräsentation natürlichsprachlicher Ausdrücke entworfen und mit einem auf diese Aufgabe abgestimmten Inventar von Sprachkonstrukten ausgestattet.

5.2. CD-GRAPHEN ALS BEISPIEL FÜR NETZ-BASIERTE REPRÄSENTATIONSSPRACHEN

In der von R. Schank innerhalb der letzten dreizehn Jahre entwickelten Theorie der *konzeptuellen Dependenz* (Abk. CD-Theorie) werden als semantische Repräsentationssprache für natürlichsprachliche Ausdrücke sog. *CD-Graphen*² verwendet. Charakteristisch für die CD-Theorie ist der Anspruch, daß CD-Graphen weniger die reine Semantik der natürlichsprachlichen Eingabe als die aus dem Verstehen der Eingabe resultierende Repräsentation des beschriebenen Ereignisses, Zustands oder Handlungsablaufs erfassen. Weitere Entwurfsziele für CD-Graphen sind die Unabhängigkeit der Elemente der Repräsentationssprache von Ausdrucksformen der natürlichen Sprachen und die daraus resultierende Universalität der Repräsentation sowie die Eindeutigkeit der Interpretationskonstruktionen.

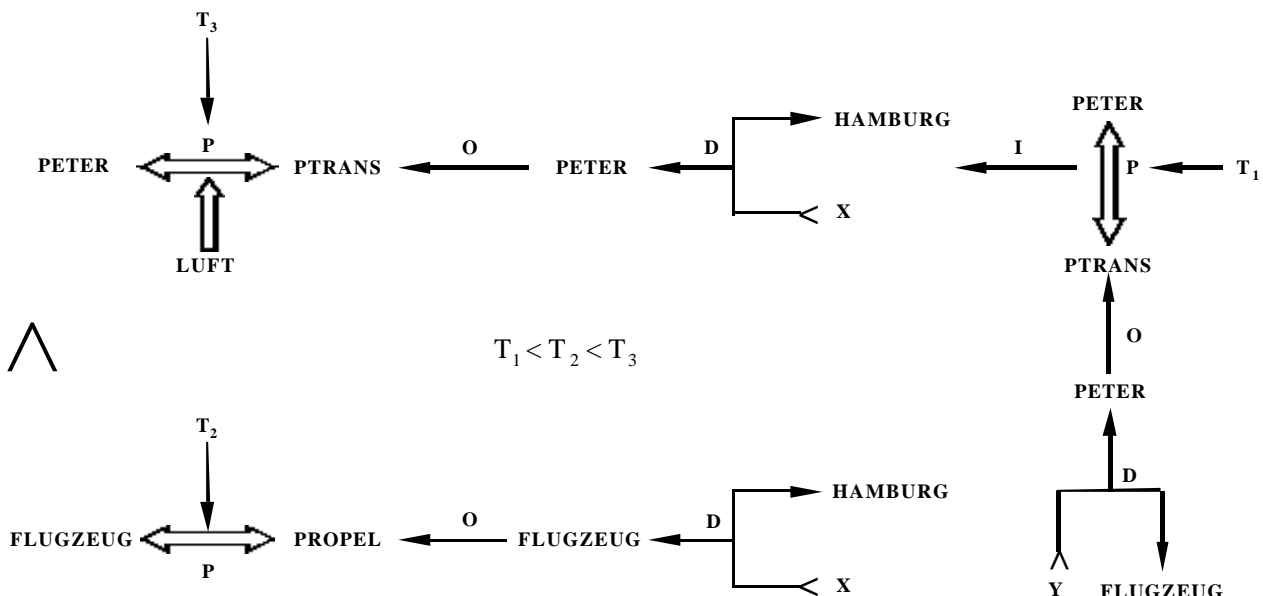


Fig. 18: CD-Repräsentation für *Peter flog nach Hamburg*

¹ Beispielsweise kann ein Hotelier aus der Frage des Gastes *Ist vielleicht ein Zimmertelephon vorhanden?* schließen, daß der Gast kein Luxushotel erwartet, in dem ein Zimmertelephon selbstverständlich vorausgesetzt werden kann.

² Durch die für die KI typische experimentelle Vorgehensweise haben sich Einzelheiten der CD-Darstellung durch die Erfahrungen bei der Konstruktion von NSS vielfach geändert und auch der z.Zt. verwendete Repräsentationsformalismus wird von Schank nicht als endgültige Version betrachtet (vgl. Schank/Riesbeck 1981, S. 11).

Jeder Knoten eines CD-Graphs kann einer der folgenden, in der CD-Theorie unterschiedenen *konzeptuellen Kategorien* zugeordnet werden (in eckigen Klammern werden Beispiele aus den Figuren 18 und 19 genannt): PP (Picture Producer) für Objekte [PETER, HAMBURG, FLUGZEUG], PA (Picture Aider) für Eigenschaften von Objekten [PART(MARIA)], ACT (Action) für eine von 11 primitiven Aktionen [PTRANS, PROPEL, GRASP, INGEST], AA (Action Aider) zur Modifikation von ACTs [ENG], LOC (Location) für Orte [LUFT], T (TIME) für Zeitpunkte [T₁, T₂, T₃].

In Abhängigkeit von der konzeptuellen Kategorie der beteiligten Knoten werden unterschiedliche Typen von Kanten zur Verbindung von Knoten verwendet. Beispielsweise werden Knoten vom Typ PP und PA durch einen dreigliedrigen Doppelpfeil verbunden (vgl. die Verbindung zwischen HANS und PART(PETER) in Fig. 19), während zwischen Knoten vom Typ AA und ACT ein eingliedriger Pfeil als Kante dient (vgl. die Verbindung zwischen ENG und GRASP in Fig. 19). Die Kanten eines CD-Graphs werden zusätzlich durch Angaben über Tiefenkasus (z.B. O = Object), Kausalitätstypen (z.B. h = 'Wird verhindert durch'), Zeiten (z.B. p = Vergangenheit) und Modalitäten (z.B. c = konditional) bewertet.

Teilstrukturen von CD-Graphen, in denen ein ACT durch einen zweigliedrigen Doppelpfeil mit einem als Actor fungierenden PP verbunden ist, werden zusammen mit den PPs für die Tiefenkasusrollen Objekt und dem aus Source und Goal bestehenden Directive (Abk.: D) als *Konzeptualisierung* bezeichnet. Durch logische Junktoren und Kausalrelationen sowie durch die Tiefenkasusrolle Instrumental (Abk.: I) können Konzeptualisierungen zueinander in Beziehung gesetzt werden (vgl. Fig. 18 und 19).

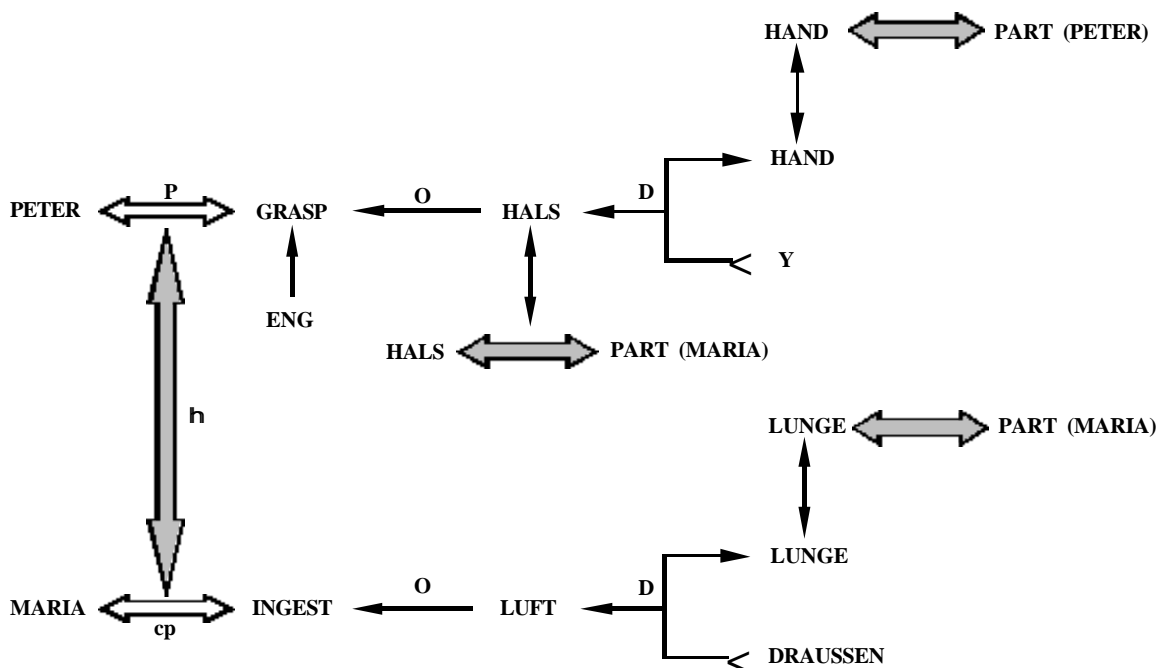


Fig. 19: CD-Repräsentation für *Peter würgte Maria*

Beispiele für *primitive Aktionen* in der CD-Theorie sind PTRANS (den Ort von etwas verändern), PROPEL (eine Kraft auf ein Objekt ausüben), GRASP (etwas greifen) und INGEST (einem Lebewesen etwas einverleiben), die alle in Fig. 18 und 19 verwendet werden. Mit jeder primitiven Aktion ist eine Menge von Inferenzregeln assoziiert. Beispielsweise (67) eine der mit INGEST assoziierten Inferenzregeln. Aus Gründen der Beschreibungsökonomie

- (67) Falls das OBJEKT von INGEST eine FLÜSSIGKEIT ist, dann negiere
'ACTOR ist DURSTIG' und
falls das OBJECT von INGEST nicht GASFÖRMIG ist, dann negiere
'ACTOR ist HUNGRIG'

wird in der CD-Theorie versucht, das Inventar von primitiven Aktionen möglichst klein zu halten. Neue ACTs werden nur eingeführt, wenn die Menge der damit assoziierten Inferenzregeln nicht durch bereits eingeführte ACTs abgedeckt werden kann.

Natürlichsprachliche Ausdrücke, die komplexe Sachverhalte, Handlungen oder Ereignisse beschreiben, werden in der CD-Theorie nach dem Prinzip der *lexikalischen Dekomposition* auf eine relationale Struktur von Konzeptualisierungen abgebildet und damit in eine Reihe primitiver Aktionen zerlegt. So wird dem Satz *Peter flog nach Hamburg* in der CD-Theorie eine aus drei Konzeptualisierungen bestehende Struktur zugeordnet (vgl. auch Fig. 18):

- (C1) Peter verändert den Ort von Peter in Richtung von Y zum Flugzeug
(C2) Das Flugzeug übt eine Kraft auf das Flugzeug aus in Richtung von X nach Hamburg
(C3) Peter verändert in der Luft den Ort von Peter in Richtung von X nach Hamburg

Dabei wird die Relation zwischen den Konzeptualisierungen (C1) und (C3) durch den Tiefenkasus Instrumental (Abk.: I) charakterisiert und für die durch (C1), (C2) und (C3) beschriebenen Teilereignisse wird eine zeitliche Reihenfolge spezifiziert.

Das in der CD-Theorie realisierte Dekompositionsprinzip wirft die Frage auf, welcher Detaillierungs- und Ausführlichkeitsgrad bei einer CD-Darstellung jeweils zu wählen ist, d.h. ob beispielsweise noch Konzeptualisierungen für *Peter kaufte ein Flugticket* und *Peter saß im Flugzeug* in dem durch Fig. 18 dargestellten CD-Graph zu ergänzen sind. Wie in Abschnitt 5.4. gezeigt wird, sieht die CD-Theorie den in Fig. 18 dargestellten Graph nur als Teil einer umfassenderen Struktur, in der die für einen Flug typischen Szenarios und Handlungsschemata als sog. *Scripts* erfaßt werden.

In Fig. 19 wird der Satz *Peter würgte Maria* auf zwei durch die Relation 'wird verhindert durch' verbundene Konzeptualisierung abgebildet:

- (C4) Maria verleiht Maria Luft ein von draussen in Marias Lunge
(C5) Peter greift mit Peters Hand eng um den Hals von Maria.

Dieses Beispiel zeigt, daß bei der Überführung einer natürlichsprachlichen Eingabe in eine CD-Darstellung auch auf Weltwissen des Systems beruhende Standardannahmen einfließen. In (C5) ist nämlich als Annahme die in der Eingabe nicht enthaltene Information angeführt, daß Peter Maria mit der Hand und nicht etwa mit einem Strick würgte.

CD-Graphen¹ haben sich als semantische Repräsentationssprache bei der Konstruktion einer ganzen Serie von experimentellen NSS (u.a. MARGIE, SAM, QUALM, PAM, TALE-SPIN, POLITICS, FRUMP, IPP, BORIS) und vereinzelt auch in Anwendungsprojekten (z.B. CADHELP, vgl. Cullingford et al. 1982) bewährt.

Allerdings wurden in anwendungsorientierten Systemen CD-Graphen bisher kaum zur Repräsentation der Semantik natürlichsprachlicher Anfragen an Datenbanksysteme oder Expertensysteme eingesetzt, weil die

¹ Für die maschinelle Speicherung und Verarbeitung werden CD-Graphen in eine Listendarstellung überführt. In Abschnitt 4.4. wurde bereits ein Beispiel für die linearisierte Darstellungsform gegeben.

für diesen Bereich typischen Probleme der *Quantifikation* (z.B. Bestimmung des Skopus von natürlichsprachlichen Quantoren) in der bisherigen Entwicklung der CD-Theorie bewußt ausgeklammert wurden (vgl. Schank/ Riesbeck 1981, S.25/26). Allgemein kann man feststellen, daß verglichen mit logik-orientierten Repräsentationskonstruktionen, deren Auswertung im folgenden an einem Beispiel eingeführt wird, alle in der KI bekannten netz-basierten Repräsentationssprachen im Bereich der Quantifikation entweder nur einfachste Strukturen zulassen oder zu relativ aufwendigen Repräsentationskonstruktionen (vgl. z.B. Hendrix 1976) führen.

5.3. EIN VERFAHREN ZUR AUSWERTUNG LOGIK-ORIENTIERTER REPRÄSENTATIONS-KONSTRUKTIONEN

Das im folgenden zu erläuternde Beispiel für die Auswertung einer logik-orientierten Repräsentationskonstruktion geht zunächst wieder von der bereits in Abschnitt 4.1. als Zielsprache für einen semantisch orientierten Parser verwendeten Repräsentationssprache SURF aus. Wie für alle in der sprachorientierten KI-Forschung verwendeten logik-orientierten Repräsentationssprachen ist die Syntax und in einigen Teilen auch die Semantik von SURF speziell auf die Repräsentationsaufgabe abgestimmt, so daß sich gegenüber der üblichen Form der Prädikatenlogik erster Stufe z.T. erhebliche Unterschiede ergeben.

Dem Beispielsatz (68), dessen Auswertung in Fig. 20 auszugsweise wiedergegeben wird, können aufgrund des nicht explizit markierten Skopus der darin auftretenden natürlichsprachlichen Quantoren, zwei Lesarten zugeordnet werden, die durch (69) und (70) paraphrasiert werden können.

(68) *Befindet sich eine Lampe neben jedem Bett?*

(69) Gibt es eine einzige Lampe, für die gilt, daß sie neben jedem Bett steht?

(70) Gibt es mehrere Lampen, so daß für jedes Bett gilt, daß eine Lampe neben ihm steht?

In dem vom Parser erzeugten SURF-Ausdruck (vgl. Fig. 20¹) wir zunächst der Skopus der Quantoren 'E' (für 'Eine') und 'Jed' nicht explizit dargestellt. Dies entspricht der Funktion von SURF als oberflächennaher Darstellung, die im System HAM-ANS Zielsprache für den Parser als auch Quellsprache für die Generierungskomponente ist (vgl. Jameson et al. 1980). Da eine semantische Auswertung aber nur für eindeutige Strukturen erfolgen kann, muß der SURF ein Ausdruck zunächst in eine Darstellung mit eindeutigen Skopusmarkierungen überführt werden.

Die logik-orientierte Repräsentationssprache DEEP, die speziell als Quellsprache für effiziente Auswertungs- und Inferenzprozesse und als Zielsprache für die Komposition der Auswertungsergebnisse entworfen wurde, unterscheidet sich von SURF im wesentlichen dadurch, daß der Skopus von Quantoren, Negations- und Modalpartikeln markiert ist und aus Lambda-Ausdrücken gebildeten Deskriptionen aufgelöst sind.

Die Überführung von SURF-Ausdrücken in DEEP-Ausdrücke wird durch die *Normalisierungskomponente* (vgl. auch Fig. 7) vorgenommen, in der u.a. aufgrund von Information über *Quantorenreänge*, die der intrinsischen Stärke natürlichsprachlicher Quantoren wie *alle, jede, beide, einige, eine, mehrere, etliche* entsprechen (vgl. Hendrix 1976) und der relativen Position der im Satz auftretenden Quantoren in einem rekursiv und top-down arbeitenden Verfahren Skopusmarkierungen und ggf. Quantorenumordnungen vorgenommen werden. In Fig.20 entsteht nach der Normalisierung eine DEEP-Formel, die der plausibleren Lesart entspricht.

¹ Das erste Element der in einem SURF-Ausdruck enthaltenen Listen markiert jeweils den syntaktischen Typ des folgenden Teilausdrucks (z.B. af-d für atomare Formel vom Typ Deskription) und dient zur effizienten Selektion der Komponenten des SURF-Ausdrucks).

SURF-AUSDRUCK

```
(af-d: IS
  (t-s: (q-qt: E-)(lambda:x1 (af-a: ISA x104 LAMPE)))
  (lambda: x2
    (af-a: RAEUMLICH NEBEN x2
      (t-s: (q-w: JED) (lambda: x3 (af-a: ISA x3 BETT))))))
```



NORMALISIERUNG

DEEP-AUSDRUCK

```
(f-d: (t-q: (for: (q-w: JED) x3) (af-a: ISA x3 BETT))
  (f-d: (t-q: (for: (q-qt : E-) x1)(af-a: ISA x1 LAMPE))
    (af-a: (RAEUMLICH NEBEN x1 x3))))
```

AUSWERTUNG

TEST: <Vollständiger DEEP-Ausdruck>

GENERATE: x3: (ISA x3 BETT)

GENERATE = (ISA BETT1 BETT)

TEST: (f-d: (t-q: (for: (q-qt: E-) x1)(af-a:ISA x1 LAMPE))

(af-a: (RAEUMLICH NEBEN x1 BETT1)))

GENERATE: x1: (ISA x1 LAMPE)

.

.

.

GENERATE: x1: (RAEUMLICH NEBEN x1 BETT1)

GENERATE = ((UNGEFAEHR (RAEUMLICH NEBEN LAMPE1 BETT1)) 0.6)

.

.

.

GENERATE: x3: (ISA x3 BETT)

.

.

.

Erfolge = (BETT1 BETT3 BETT4) *Misserfolge* =(BETT2)

```
TEST = ((f-d:(t-q: (for: (q-w: (FAST JED) x3)(af-a: ISA x3 BETT))
  (f-d: (t-q: (for: (q-qt: E-) x1)(af-a: ISA x1 LAMPE))
    (af-a: (RAEUMLICH NEBEN x1 x3)))) 0.6)
```

Fig. 20: Beispiel für die Auswertung eines SURF-Ausdruckes

Die Auswertung erfolgt in einem rekursiven Verfahren mit zwei Gruppen durch *patterngesteuerten Prozeduraufruf* aktivierter Prozeduren, den sog. TEST- und GENERATE-Prozeduren. Die TEST-Prozeduren überprüfen unter Verwendung aller dem System zur Verfügung stehenden Wissensquellen und Inferenzkapazitäten den Wahrheitswert voll instantiiertener DEEP-Formeln und ordnen dem Auswertungsergebnis eine Zahl aus dem reellen Intervall $[0,1]$ zu, die vom System als Sicherheitsfaktor für die Zuordnung des Wahrheitswertes interpretiert wird (vgl. Fig. 20). Die GENERATE-Prozeduren erzeugen nacheinander Konstanten, durch deren Substitution für eine bestimmte Variable einer DEEP-Formel der Wahrheitswert 'wahr' zugeordnet werden kann. TEST- und GENERATE-Prozeduren rufen sich bei der Top-down-Auswertung einer DEEP-Formel gegenseitig rekursiv auf, wobei die Auswertung stets mit einer TEST-Prozedur begonnen wird (vgl. Fig. 20). Während der Auswertung wird eine eingegebene DEEP-Formel immer weiter dekomponiert, um dann anschließend durch einen Kompositionsprozeß ggf. in veränderter Form als Auswertungsergebnis aufgebaut zu werden.

Für das Beispiel in Fig. 20 wird zunächst im referentiellen Wissen des Systems durch eine GENERATE-Prozedur nach einem Bett gesucht und dann für das gefundene BETT1 mithilfe einer TEST-Prozedur geprüft, ob eine Lampe neben ihm steht. Mit LAMPE1 wird eine Lampe gefunden, die ungefähr neben BETT1 steht. Die Flexibilität der Auswertungskomponente zeigt sich u.a. darin, daß dabei auch Modifikationen von Prädikationen durch Hecken wie *ungefähr* berücksichtigt werden können, die unsichere Ergebnisse von *approximativen Inferenzprozessen* oder Unsicherheiten bei der Verbalisierung der Ergebnisse visueller Suchprozesse signalisieren. Nachdem mit BETT1 und LAMPE1 eine erfolgreiche Instantiierung der eingebetteten Formel gefunden ist, wird durch die bereits einmal durchlaufene GENERATE-Prozedur eine weitere Individuenkonstante der Sorte BETT erzeugt. Nachdem für die vier im referentiellen Wissen gefundenen Betten nur drei erfolgreiche und eine nicht erfolgreiche Instantiierungen der eingebetteten Formel gefunden wurde, überführt ein Kompositionsprozeß schließlich den eingegebenen Quantor 'Jed' in den durch eine Hecke modifizierten Quantor-Ausdruck 'Fast Jed'.

Im Gegensatz zu der üblichen prädikatenlogischen Auswertung eines Allquantors braucht das System die Frage *Befindet sich eine Lampe neben jedem Bett* trotz eines Gegenbeispiels (vgl. BETT2 in Fig. 20) also nicht mit *Nein* zu beantworten, sondern kann kommunikativ adäquat mit der Formulierung *Ja, neben fast jedem* reagieren.

Das charakteristische Pattern von TEST- und GENERATE-Prozeduren bestimmt eine mehr oder weniger große Klasse von DEEP-Ausdrücken, auf deren Auswertung die Prozeduren jeweils spezialisiert sind. Dadurch, daß mehrere alternative oder neben sehr speziellen auch sehr allgemeine Prozeduren für eine Klasse von DEEP-Ausdrücken vorhanden sind, können kontextabhängig verschiedene *Auswertungsstrategien* gewählt werden. Beispielsweise gibt es zusätzlich zu der üblichen Interpretation der Konjunktion in der Prädikatenlogik bei dem hier vorgestellten Auswertungsverfahren die Möglichkeit, für die Konjunktion (AND (REF x1 BEQUEM) (ISA x1 STUHL)) zunächst durch eine spezielle GENERATE-Prozedur Konstanten der Sorte Sessel durch (ISA x1 SESSEL) zu erzeugen.

Auf diese Weise wird es möglich, Wissen in den Auswertungsprozeß einzubringen und somit auch die Evaluation komplexer Repräsentationskonstruktionen für natürlichsprachliche Eingaben effizient zu gestalten.

5.4. SCRIPTS ZUR UNTERSTÜTZUNG DER AUSWERTUNG VON SATZFOLGEN

Es gehört auch zu den Aufgaben des Verarbeitungsabschnitts 'Auswertung', jede einzelne analysierte Eingabe in den größeren Zusammenhang des bereits gelesenen Textes bzw. des vorausgegangenen Dialoges einzubetten und daraus Erwartungen bezüglich weiterer Eingaben abzuleiten. Eine Möglichkeit solche größeren Zusammenhänge zu erkennen, die besonders in textverstehenden Systemen wie SAM (vgl. Schank/Abelson 1977) verwendet wurde, besteht darin zu versuchen, jede einzelne Eingabe auf eine als allgemeines Hintergrundwissen gespeicherte Beschreibung einer *stereotypen Ereignisfolge* zu beziehen.

Solche stereotypen Ereignisfolgen können in NSS durch sog. *Scripts* erfaßt werden, die ähnlich wie Frames (vgl. Minsky 1975) speziell zur Unterstützung erwartungsgesteuerter Verarbeitungsprozesse entwickelt wurden. Fig. 21 zeigt ein Script¹ für Restaurantbesuche, dessen Hauptbestandteil eine Folge von Ereignisbeschreibungen ist.

Bei der Verarbeitung einer Satzfolge wie (71) - (72) wird zunächst die für Satz (71) erzeugte semantische Repräsentationskonstruktion mit den Kessätzen aller verfügbaren Scripts

(71) *Fritz ging in die Pizzeria Casa Mia und bestellte eine große Portion Tortellini.*

(72) *Verärgert durch die überhöhte Rechnung verließ er nach einer halben Stunde das Restaurant.*

verglichen. Im vorliegenden Beispiel instantiiert das erste Konjunkt von Satz (71) den Kessatz des Restaurant-Scripts, wodurch dieser aktiviert wird. Danach sind einige Script-Variable mit Werten belegt (Gast: = Fritz, Restaurant: = (Pizzeria Casa Mia), Mahlzeit: = (eine große Portion Tortellini)) und als weitere Eingaben werden Ereignisbeschreibungen 5-10 des Restaurant-Scripts erwartet. Durch die Auswertung von (72) wird der Script-Variablen 'Rechnung' die Deskription 'eine überhöhte Rechnung' zugewiesen und der Script nach Instantiierung der Ereignisbeschreibung 10 ('Fritz verläßt die Pizzeria Casa Mia') deaktiviert. Einige Vorteile einer script-gesteuerten Auswertung sind:

- das System kann Fragen (z.B. *Was hat Fritz gegessen?* nach (71) (72)), für die eine Antwort nicht explizit im Eingabetext enthalten ist, aufgrund eines partiell instantiierten Scripts beantworten (*eine große Portion Tortellini*)
- das System kann Schlußfolgerungen ziehen, die nur durch eine script-basierte Ergänzung von in der verbalen Beschreibung fehlenden Handlungsschritten ausgelöst werden können (z.B. *Fritz aß eine große Portion Tortellini* → *Fritz ist danach nicht sehr hungrig*)
- das System kann aufgrund der im Kontext eines Scripts vorausgesetzten Personen und Objekte einige Nominalphrasen mit bestimmtem Artikel anaphorisch deuten, obwohl die entsprechenden Referenzobjekte im Text nicht explizit vorerwähnt sind (z.B. *die überhöhte Rechnung* in (72))

¹ Für die maschinelle Verarbeitung müssen entgegen der vereinfachten Darstellung in Fig. 21 die Ereignisbeschreibungen in einer formalen Repräsentationssprache (z.B. CD-Graphen) vorliegen.

Beteiligte Personen:	<i>Gast, Kellner</i>
Beteiligte Objekte:	<i>Restaurant, Tisch, Speisekarte, Mahlzeit, Rechnung, Trinkgeld</i>
Ereignisse:	<ol style="list-style-type: none"> 1. <i>Gast geht ins Restaurant</i> 2. <i>Gast geht an Tisch</i> 3. <i>Kellner bringt Speisekarte</i> 4. <i>Gast bestellt Mahlzeit</i> 5. <i>Kellner bringt Mahlzeit</i> 6. <i>Gast ißt Mahlzeit</i> 7. <i>Kellner bringt Rechnung</i> 8. <i>Gast zahlt Rechnung</i> 9. <i>Gast hinterläßt Trinkgeld für Kellner</i> 10. <i>Gast verläßt Restaurant</i>
Kennsatz:	<i>Ereignis 1</i>
Hauptkonzept:	<i>Ereignis 6</i>

Fig. 21. Beispiel für ein Script

Da für eine Vielzahl von weniger standardisierten Szenarios keine Scripts gespeichert sind, kann die Auswertung solcher Ereignisbeschreibungen nicht script-basiert erfolgen, sondern muß über das Erkennen von *Handlungsmotiven*, *Handlungszielen* und *Plänen* der beteiligten Personen laufen (vgl. Schank/Abelson 1977, Schank/Riesbeck 1981).

Außerdem hat sich inzwischen gezeigt, daß Scripts Weltwissen oft nicht auf der angemessenen Abstraktionsstufe erfassen. Beispielsweise könnte man einerseits spezielle Scripts für Luxusrestaurants und Selbstbedienungsrestaurants postulieren, andererseits könnte man Teile des Restaurant-Scripts in Fig. 21 als Ausprägung eines allgemeinen Dienstleistungs-Scripts auffassen. Mit dem Konzept der *MOPs* (Memory Organization Packages) erprobt Schank zur Zeit eine Organisation von Wissen, in der die Probleme der Vermischung, Überlappung Interferenz und Vererbung aus Erfahrung beruhender Wissensbereiche gelöst werden sollen (vgl. Schank 1982).

6. VERFAHREN ZUR GENERIERUNG NATÜRLICHER SPRACHE

Der Verarbeitungsabschnitt 'Generierung' (vgl. Kapitel 2) ist in den meisten NSS weniger ausgeprägt und wesentlich leistungsschwächer als die Verarbeitungsabschnitte 'Analyse' und 'Auswertung'¹. Die Tatsache, daß der menschliche Hörer dazu in der Lage ist, auch schlecht formulierte, ungrammatische und kaum verständliche Äußerungen (z.B. Äußerungen eines Kindes, das Sprechen lernt) mithilfe von Weltwissen und Kontextinformation zu interpretieren, legte es zunächst nahe, in anwendungsorientierten NSS die Sprachgenerierungskomponente nur so weit auszubauen, daß die Qualität der Äußerungen des NSS vom Benutzer gerade noch toleriert werden könnte. Als Standardverfahren zur Sprachausgabe, die aus dem Blickwinkel der KI zu undifferenziert sind und daher im folgenden nicht behandelt werden, verwenden viele NSS folgende Techniken:

- Ausgabe vorgefertigter Textteile
- Formatierte Ausgabe von Auszügen aus der Wissensbasis des Systems
- Kontextsensitive Instantiierung von vorgefertigten Textschemata

Inzwischen stellte man fest, daß die fehlende Kohärenz, zu starke Stilisierung und Monotonie, zu hohe Redundanz und schlechte Lesbarkeit der ausgegebenen Äußerungen die Akzeptanz von natürlichsprachlichen KI-Systemen erheblich beeinträchtigen kann. Man hat daher begonnen, Komponenten für den Verarbeitungsabschnitt 'Generierung' zu entwickeln, die neben dem allgemeinen Hintergrundwissen besonders dialog- bzw. textbezogene Wissensquellen ausnutzen, um *kommunikativ-adäquate*, verständliche und kohärente Äußerungen zu erzeugen (vgl. Mann et al. 1981).

6.1. WORTWAHL UND NP-GENERIERUNG ALS AUFGABEN DER VERBALISIERUNGSKOMPONENTE

Um eine im Verarbeitungsabschnitt 'Auswertung' erzeugte semantische Repräsentationskonstruktion (vgl. Fig. 6 in Kapitel 2), deren Elementen nur kontextabhängig ein Lemma oder ein natürlichsprachlicher Ausdruck zugeordnet werden kann (z.B. semantische Primitive, vgl. Abschnitt 5.2), verbalisieren zu können, muß die Generierungskomponente eines NSS über Prozeduren zur Realisierung der Wortwahl verfügen.

In dem von Goldman entwickelten System BABEL, das als Generierungskomponente von MARGIE verwendet wurde (vgl. Schank 1975), werden primitive Aktionen der CD-Theorie wie INGEST (vgl. Fig. 19) mithilfe von *Diskriminationsnetzen* kontextabhängig in ein Verb überführt (vgl. Fig. 22). Dabei werden die Eigenschaften von 'Picture Producers' (vgl. Abschnitt 5.2), die eine bestimmte Tiefenkasusrolle in der mit der primitiven Aktion verbundenen Konzeptualisierung übernehmen, sukzessive überprüft, bis eine eindeutige Wortwahl möglich ist. Beispielsweise wird die primitive Aktion INGEST genau dann als *Einatmen* verbalisiert, wenn in der entsprechenden Konzeptualisierung das Objekt nicht flüssig und gasförmig ist (vgl. Fig. 22). Nach der Wortwahl für eine primitive Aktion wird in BABEL der mit dem Verb assoziierte Kasusrahmen instantiiert und anschließend mithilfe einer *ATN-Generierungsgrammatik* in einen natürlichsprachlichen Satz überführt.

¹ Beispielsweise bestehen in USL die Antworten des Systems neben Partikeln wie *Ja, Nein, Doch* nur aus formatierten Ausgaben von Datenbankinhalten (vgl. Ott 1979, S. 129).

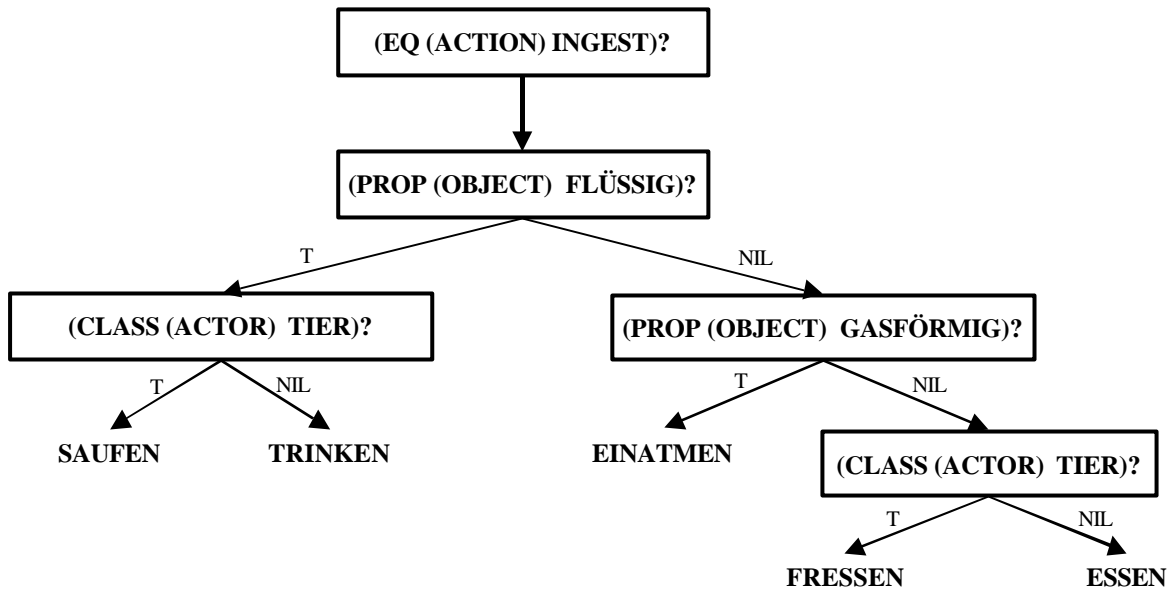


Fig. 22: Diskriminationsnetz zur Steuerung der Wortwahl

Falls die zu verbalisierende Repräsentationskonstruktion einen systeminternen Objektbezeichner wie BETT1 als Individuenkonstante erzeugt, muß dieser durch die Verbalisierungskomponente in eine Nominalphase überführt werden. Solche Individuenkonstanten treten nur dann in den Repräsentationskonstruktionen für die natürlichsprachliche Ausgabe auf, wenn wie in (73) und (74) für die zu generierende NP eine *extensionale Lesart* und

(73) *Das Fernsehgerät steht direkt neben dem Bett.*

(74) *Auf der ersten Etage befindet sich eine Dusche.*

nicht wie in (75) und (76) eine *intensionale Lesart* intendiert ist. Aufgabe der NP-Generie-

(75) *Der Farbfernseher gehört zur Standardausstattung eines Luxushotels.*

(76) *Eine Dusche ist leider nicht vorhanden.*

rungskomponente ist es, einen systeminternen Objektbezeichner durch die semantische Struktur einer NP zu ersetzen, die dem Hörer die Identifikation des intendierten Referenzobjektes ermöglicht (vgl. Fig. 23 und Jameson/Wahlster 1982). NP-Generierungskomponenten greifen zur Bewältigung dieser Aufgabe typischerweise auf folgende Wissensquellen zu:

- Referentielles Wissen
- Begriffliches Wissen
- Inferenzregeln
- Partnermodell

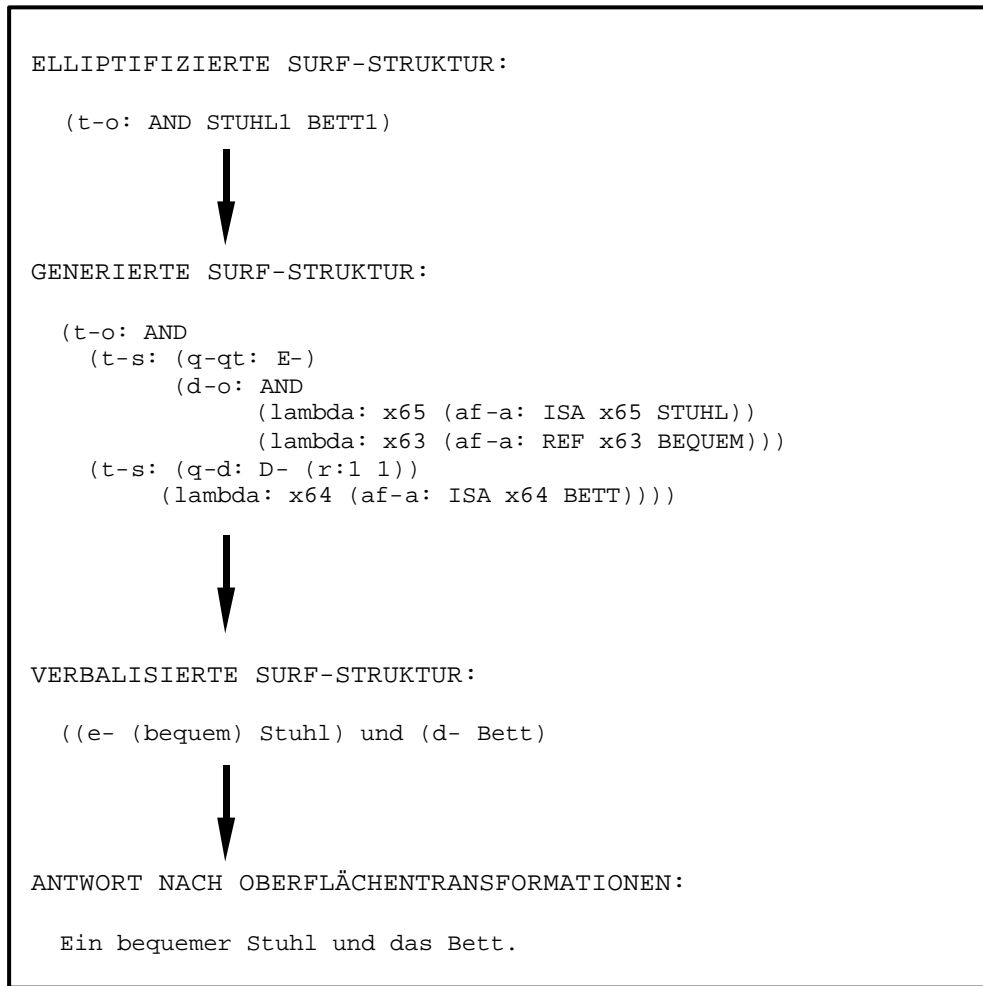


Fig. 23: Beispiel für einen NP-Generierungsprozeß

In Dialogsystemen hängt die Komplexität der NP-Generierungskomponente von der gewählten *Dialogsituation* ('Beide Dialogpartner kennen bzw. sehen Details der Diskurswelt' vs. 'Dem Hörer ist die konkrete Diskurswelt unbekannt') und den Konversationszielen des Systems ab ('System hat keine speziellen Ziele, sondern soll nur möglichst 'objektiv' und kooperativ antworten' vs. 'System soll eine Entscheidung des Benutzers beeinflussen').

Wenn die NP-Generierung aufgrund vorgegebener Konversationsziele zum *interessen-geleiteten kognitiven Entscheidungsprozeß* wird (z.B. bei der Beschreibung eines angebotenen Hotelzimmers, das durch das System vermietet werden soll) sind folgende Komponenten des Partnermodells von besonderer Bedeutung:

- Existenzvoraussetzungen und Partnererwartungen (a priori Partnermodell)
- Spezielle Wünsche und Einstellungen des Partners (dialogbasiertes Partnermodell)
- Referentielles Wissen des Partners (dialogbasiertes Partnermodell)
- Dimensionspräferenzen

Falls das Referenzobjekt beim Partner als unbekannt vorausgesetzt werden muß, wird es zunächst meist durch die NP-Generierungskomponente als indefinite NP eingeführt, um dann im weiteren Dialog bzw. Text durch referenzidentische Pronomen oder anaphorisch

gebrauchte NPs wieder aufgegriffen zu werden (z.B. '*neben einem antiken Sessel...Der Sessel...*').

Bei der Generierung einer indefiniten NP muß das System unter Verwendung des Partnermodells eine angemessene Anfangscharakterisierung auswählen. Wenn z.B. das intern als STUHL1 bezeichnete Objekt mithilfe der NP *ein beschädigter Stuhl* oder *ein bequemer Stuhl* eingeführt werden kann, muß das System aufgrund des Partnermodells die zweite Verbalisierungsmöglichkeit wählen, da nur diese mit dem Handlungsziel 'Vermietung eines Hotelzimmers' vereinbar ist. Mit einer solchen Erzeugung einer indefiniten NP muß stets eine entsprechende Ergänzung des referentiellen Wissens im *dialogbasierten Partnermodell* verbunden sein. Wenn die Existenz bestimmter Objekte aufgrund des *a priori Partnermodell* vorausgesetzt werden muß (z.B. ein Bett, eine Tür und ein Fenster in einem Hotelzimmer) sollte das System ohne vorherige Verwendung einer indefiniten NP gleich definite NPs wie *das Bett* als sog. *pragmatische Anaphern* (d.h. Anaphern ohne sprachliches Antezedens) erzeugen. Sonst werden erst nach der Verwendung indefiniter NPs anaphorisch gebrauchte definite NPs verwendet.

Auf ein und dasselbe Objekt kann in verschiedenen Dialogabschnitten durch unterschiedliche definite NPs verwiesen werden. Dabei wird nach dem Prinzip der Verwechslungsvermeidung das zu benennende Objekt jeweils von der Menge aller Objekte abgegrenzt, von denen das System annimmt, daß der Benutzer sie in der gegebenen Dialogsituation mit dem Referenzobjekt verwechseln könnte. Wenn das Referenzobjekt eines von mehreren Objekten ist, die als referentielles Wissen des Partners gespeichert sind und zur selben konzeptuellen Klasse gehören, muß das System eine Teilmenge der für das Referenzobjekt gespeicherten Merkmale auswählen, die es von den anderen Objekten dieser Klasse unterscheidet (z.B. der rote, gepolsterte Sessel). Dabei müssen die in der Eingabe des Benutzers präsupponierten Eigenschaften des Referenzobjektes blockiert werden z.B. (Welches von den drei Betten ist hart? Nicht: das harte Bett, sondern z.B. das mittlere) .

Im Fall der multiplen Benennbarkeit, d.h. wenn dem System bei der Generierung einer definiten NP alternativ mehrere Merkmale zur Verfügung stehen, muß das System für eine nicht redundante Codierung nacheinander folgende Kriterien auswerten bis ein Merkmal gewählt worden ist:

- Wähle das Merkmal, welches gemäß Partnermodell am besten den speziellen Wünschen und Einstellungen des Benutzers entspricht.
- Wähle das Merkmal mit dem höchsten Ausprägungsgrad, welches die beste Diskrimination des Referenzobjektes ermöglicht.
- Wähle das Merkmal mit der höchsten Dimensionspräferenz (z.B. Farbe vor Größe).

Eines der schwierigsten Probleme der NP-Generierung bei visuell präsenten Diskurswelten ist die eindeutige Abgrenzung eines Objektes, in dessen unmittelbarer Umgebung sich eine Gruppe äußerlich mit dem Referenzobjekt identischer Objekte befindet. In diesem Fall müssen in der definiten NP komplexe räumliche Relationen zur Abgrenzung des Referenzobjektes in der Objektgruppe und zur Identifikation der entsprechenden Objektgruppe verwendet werden (z.B. *der hintere braune Stuhl, welcher sich rechts neben dem mittleren roten Bild befindet*, vgl. auch v. Hahn et al. 1980).

6.2. VERFAHREN ZUR VERBREITERUNG DER KOMMUNIKATIVEN LEISTUNGEN DER GENERIERUNGSKOMPONENTE

Die im letzten Abschnitt eingeführten Module zur Wortwahl und NP-Generierung sind notwendige Grundbausteine des Verarbeitungsabschnitts 'Generierung', die alleine allerdings noch kein kommunikativ-adäquates, kohärentes und kooperatives Ausgabeverhalten des Systems garantieren. Dazu muß der Verarbeitungsabschnitt 'Generierung' zusätzlich z.B. noch Komponenten zur Überbeantwortung, Elliptifizierung und zur Verbalisierung von Teilen der Wissensbasis enthalten, wie sie im folgenden kurz charakterisiert werden sollen.

Die Systemreaktion in (79) ist ein typisches Beispiel für *kooperatives Dialogverhalten*: Es wird nicht nur die eigentlich erfragte Abfahrzeit sondern zusätzlich auch noch das Abfahrtsgleis genannt, womit eine sog. *Überbeantwortung* vorliegt. Das gezeigte Systemverhalten mag gerade wegen seiner hohen Plausibilität und Natürlichkeit zunächst

(78) BEN: *Wann fährt der nächste Zug nach Frankfurt ab?*

(79) SYS: *16.34, Gleis 13.*

trivial erscheinen. Doch wie die Arbeit von Cohen 1978 überzeugend nachweist, erfordert die formale Rekonstruktion dieses Verhaltens in einem NSS aufwendige *Planerkennungs-* und *Planerzeugungsprozesse*. Um zu der in (79) angegebenen Überbeantwortung zu gelangen, muß das System ein allgemeines Verfahren mit folgender Grobstruktur anwenden:

- (S1) Beobachte das Benutzerverhalten
- (S2) Erkenne den Plan des Benutzers
- (S3) Suche nach evtl. Hindernissen für den inferierten Plan, indem du die geplanten Schritte in einer gedachten Welt des Benutzers ausführst
- (S4) Mache dir die entdeckten Hindernisse zum eigenen Ziel
- (S5) Erzeuge einen Plan zum Erreichen dieses Ziels
- (S6) Führe den Plan aus und gehe nach (S1)

Bei Anwendung des skizzierten Verfahrens auf die Benutzerfrage (78) inferiert das System zunächst als Plan des Benutzers, daß dieser mit dem Zug nach Frankfurt fahren will. Das System stellt dann fest, daß die fehlende Information über das Abfahrtsgleis ein Hindernis für den Plan des Benutzers darstellen könnte, und beseitigt dieses durch die Zusatzinformation. Das Beispiel verdeutlicht die Abhängigkeit kommunikativer Leistungen wie der Überbeantwortung von kognitiven Fähigkeiten der Planerkennung und -erzeugung.

Neuere Arbeiten zeigen, daß zur Generierung kommunikativ adäquater Äußerungen in allen Phasen des Verarbeitungsabschnitts 'Generierung' Planungsprozesse benötigt werden. Beispielsweise benutzt das System KAMP (vgl. Appelt 1981) nicht nur auf der Ebene der Sprechaktplanung sondern auch zur Wortwahl und syntaktischen Strukturierung der Äußerung eine Planungskomponente. Im System TEXT (vgl. McKeown 1982) werden bei der Planung eines Textes verschiedene Wissensquellen wie semantisches Dialoggedächtnis, Fokus und Regeln zur Textstrukturierung ausgewertet.

Wenn NSS überhaupt natürlichsprachliche Ausgaben erzeugen, so sind diese meist elliptisch. Im Gegensatz zum Verarbeitungsabschnitt 'Analyse', für den spezielle Komponenten zur Ellipsenverarbeitung entworfen wurden (vgl. Abschnitt 4.5.1), sind bei der Generierung Ellipsen in den meisten NSS nicht das Ergebnis eines expliziten Elliptifizierungsprozesses über der semantischen Repräsentationskonstruktion eines nicht elliptischen Auswertungsergebnisses, sondern entstehen als Seiteneffekt dadurch, daß von der Auswertung nur einzelne Objektnamen oder Prädikate an die Generierungskomponente übergeben werden. Man kann sich am Beispiel der kotextuellen Ellipsen (vgl. Abschnitt 4.5.1) leicht klar machen, daß kooperatives und kommunikativ-adäquates Antwortverhalten aber einen expliziten Elliptifizierungsprozeß voraussetzt, in dem u.a. die Struktur der Bezugsäußerung berücksichtigt wird (vgl. Jameson/Wahlster 1982). Wenn das System z.B. auf die Frage (80) mit (81) als Überbeantwortung reagiert, weil es tatsächlich drei

(80) BEN: *Gibt es drei Betten und mindestens einen Schreibtisch im Zimmer?*

(81) SYS: *Ja, drei.*

Betten und drei Schreibtische gibt, so ist dieses Antwortverhalten inadäquat, weil der Benutzer *drei* entweder auf die Zahl der Betten bezieht oder unsicher ist welche Interpretation vorliegt. Andererseits ist die elliptische Antwort *Ja, zwei genau* dann adäquat, wenn tatsächlich drei Betten und zwei Schreibtische im Hotelzimmer vorhanden sind. Dies zeigt, daß es nicht ausreicht, die während der Auswertung vorgenommene Änderung der Quantifikation (in (80) - (81) mindestens *ein } drei*) einfach zu verbalisieren. Von einer Elliptifizierungskomponente muß jede mögliche Form der Elidierung daraufhin überprüft werden, ob die erzeugte Ellipse für den Benutzer verständlich ist, d.h. ob sie durch einen Vergleich mit der Bezugsäußerung eindeutig rekonstruierbar ist.

So gehört es zu den Aufgaben der Elliptifizierungskomponente festzustellen, daß (81) als Antwort auf (80) mehrdeutig und damit 'zu elliptisch' ist, während (82) eine

(82) SYS: *Ja, drei Betten und drei Schreibtische.*

korrekte Elliptifizierung des Auswertungsergebnisses 'Ja, es gibt drei Betten und drei Schreibtische im Zimmer' darstellt.

Durch die von Jameson entwickelte Elliptifizierungskomponente von HAM-ANS (vgl. Jameson/Wahlster 1982) wird das Auswertungsergebnis top-down analysiert und mit der semantischen Struktur der Eingabe verglichen, um nacheinander alle Teilketten in der Repräsentationskonstruktion für die Antwort zu tilgen, die mit der Repräsentationskonstruktion für die Eingabe übereinstimmen. Ausgehend von der kürzesten möglichen Ellipse wird dann die für den Verarbeitungsabschnitt 'Analyse' entwickelte Komponente zur Ellipsenrekonstruktion angewendet, um zu überprüfen, ob die geplante Äußerung verständlich ist.

Hier handelt es sich um eine *lokale Rückkopplungsschleife*, in der ein Teil des Verstehensprozesses auf der Hörerseite antizipiert wird. Dieses Beispiel zeigt auch deutlich, daß die bis vor wenigen Jahren weit verbreitete Annahme, die automatische Generierung natürlicher Sprache sei einfacher als die Analyse und unabhängig von anderen Systemkomponenten, unbegründet ist, da eine leistungsfähige Generierungskomponente aus den vielen Verbalisierungsalternativen eine in Bezug auf das jeweilige Konversationsziel optimale Formulierung nur dadurch auswählen kann, daß die Analysekomponenten des

Systems zur *Antizipation des nachfolgenden Verstehensprozesses* auf der Hörerseite verwendet werden.

Häufig ist es erforderlich, daß ein NSS auch Auszüge aus den von ihm verwendeten Wissensquellen verbalisieren kann (z.B. zur Beantwortung metakommunikativer Fragen des Benutzers). Besonders wichtig ist diese Systemfähigkeit für den Aufbau einer *Erklärungs-*

```
(PROC DEMON: MDEMON ZVAL: 0.8
(REF PAUTO BILLIG)
(OR* (GOAL (REF !AUTO TUEV-FAELLIG))
(G.OAL (REF !AUTO REPARATURBEDUERFTIG)))
(END?))
```

Fig. 24: Als DEDUCE-Prozedur codierte Inferenzregel

komponente, durch die inferenzbasierte Systemantworten aufgrund von 'Warum'-Fragen des Benutzers erklärt werden sollen (vgl. Wahlster 1981b). Aufgabe des Verarbeitungsabschnittes 'Generierung' ist es dabei, eine z.B. in der KI-Programmiersprache FUZZY als sog. DEDUCE-Prozedur codierte Inferenzregel (vgl. Fig. 24) in eine natürlichsprachliche Formulierung wie (83) zu überführen.

(83) *Ein Auto ist meist billig, wenn es TÜV-fällig oder reparaturbedürftig ist.*

Fig. 25 zeigt die Grobstruktur des Teils der Sprachgenerierungskomponente, der für die *Verbalisierung von Inferenzregeln* verantwortlich ist. Der Quelltext der Inferenzregel wird zunächst umcodiert und lexikalisch analysiert. Voraussetzung für dieses Verfahren ist die Program/Daten-Äquivalenz, wie sie in LISP und LISP-basierten KI-Programmiersprachen verwirklicht ist. Durch Zugriff auf das Wort-Lexikon wird eine präterminale Kette erstellt, die dem ATN-Generator übergeben wird. Die mit den Zustandsübergängen in der ATN-Generierungsgrammatik assoziierten Aktionen (vgl. auch Abschnitt 4.1.) lösen semantisch-pragmatische Prozesse wie die Generierung von definiten und indefiniten NPs, Pronomen und linguistische Hecken aus. Nach einem erfolgreichen Durchlauf durch die Generierungsgrammatik wird eine vorläufige Formulierung der Inferenzregel als präterminale Kette (vgl. (84) für die in Fig. 24 dargestellte Inferenzregel) an ein Produktionensystem übergeben, das durch *Oberflächentransformationen* einen wohlgeformten deutschsprachigen Satz wie (83) erzeugt.

(84) ((DET (E-)) (NOM (AUTO)) (VRB (SEIN)) (ADV (MEIST)) (ADJ (BILLIG)) (KON (WENN))
(PRN (ES)) (ADJ (TUEV-FAELLIG)) (VRB (SEIN)) (KON (ODER)) (PRN (ES)) (ADJ
(REPARATURBEDUERFTIG)) (VRB (SEIN)))

Mit Hilfe des Produktionensystem-Interpreters werden solange Tilgungs-, Insertions- und Permutationstransformationen über der präterminalen Kette ausgeführt, bis keine der in der Wissensbasis gespeicherten Produktionen (vgl. Fig. 25) mehr angewendet werden kann. Auf die in (84) angegebene präterminale Kette kann z.B. die in (85) dargestellte Produktion

(85) ((PRN (?PRN)) ??ADJOBJ (VRB (?VRB)) (KON ((*ANY ?KON '(UND ODER))))
(PRN,!PRN)) (ADJ (?ADJ)) (VRB (!VRB))
((PRN (!PRN)) !!ADJOBJ (KON !KON)) (ADJ (!ADJ)) (VRB (!VRB)))

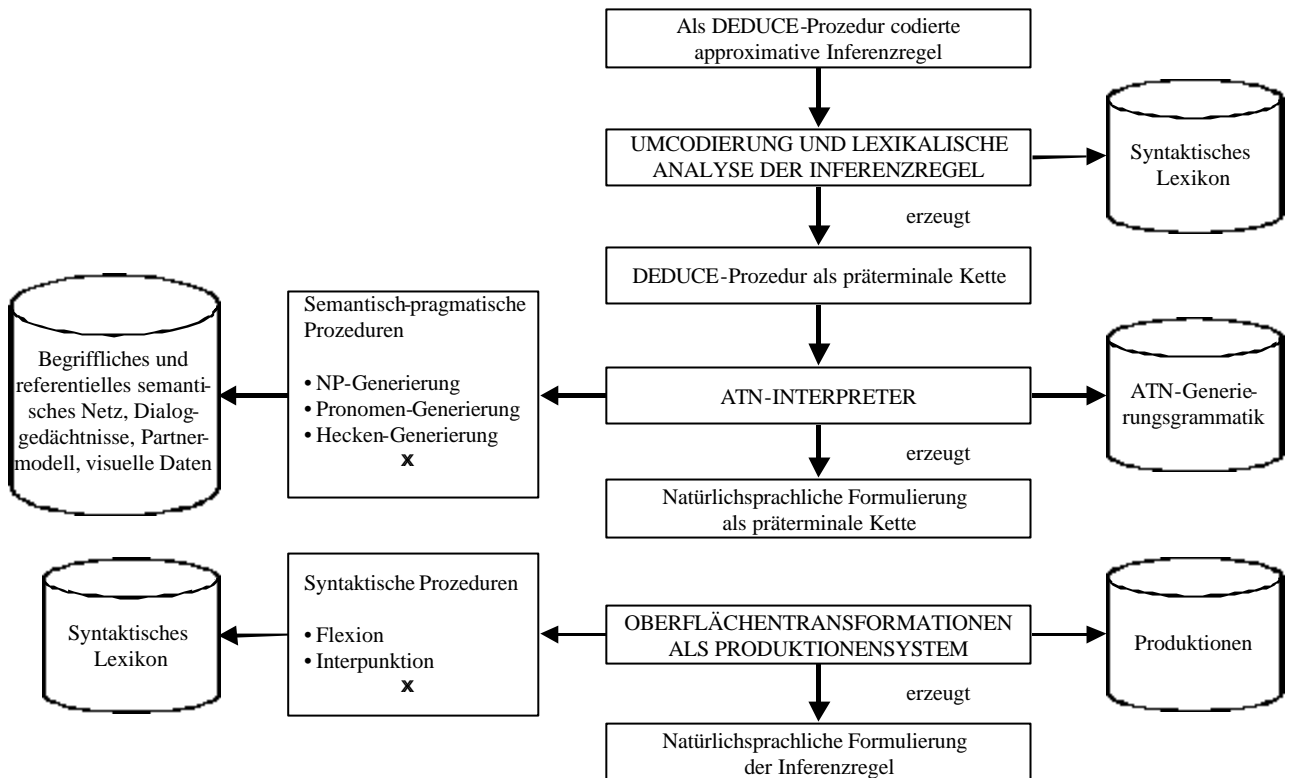


Fig. 25: Architektur einer Komponente zur Verbalisierung von Inferenzregeln

angewendet werden, die eine *Tilgung identischer Pronomen* bewirkt. Beim Vergleich der linken Seite dieser Regel mit der in (84) angegebenen Struktur, wird an die Variable ?PRN der Wert ES und an die Variable ?VRB der Wert SEIN gebunden. Durch die wertabgebenden Variablen !PRN und !VRB wird dabei gesichert, daß die in (85) formulierte Tilgungstransformation nur auf koordinierte Teilstrukturen mit identischem Pronomen und Verb angewendet wird. Nach Ausführung dieser Tilgungstransformation enthält der Arbeitsspeicher des Produktionensystems die in (86) angegebene Struktur. Auf diese Struktur

- (86) ((DET (E-)) (NOM (AUTO)) (VRB (SEIN)) (ADV (MEIST)) (ADJ (BILLIG))
 (KON (WENN)) (PRN (ES)) (ADJ (TUEV-FAELLIG)) (KON (ODER)) ADJ (REPA
 RATURBEDUERFTIG)) (VRB (SEIN)))

werden dann weitere Produktionen zur Kommainsertion und Flexion angewendet, um schließlich als Verbalisierung der Inferenzregel die Formulierung in (83) zu erhalten.

Zu den Oberflächentransformationen gehören neben der *Flexion* (vgl. Busemann 1982) und *Interpunktion* auch die Erzeugung feststehender Redewendungen und stilistischer Variationen der Wortstellung (z.B. Topikalisierungstransformationen, wie sie in der Generierungskomponente MUMBLE realisiert sind, vgl. McDonald 1980). Mit den Oberflächentransformationen wird der Sprachgenerierungsprozeß in natürlichsprachlichen Systemen und damit auch das letzte Kapitel dieser Einführung abgeschlossen.

AUSBLICK

Im vorliegenden Aufsatz wurde versucht, einen einführenden Überblick zu den wichtigsten Fragestellungen, Methoden und Resultaten des Forschungsgebietes 'Natürlichsprachliche Systeme' zu geben. An dieser Stelle sei der Anfänger davor gewarnt, natürlichsprachliche Systeme als 'Allheilmittel' für sämtliche Probleme der Mensch-Maschine-Kommunikation anzusehen. Selbstverständlich gibt es neben der natürlichen Sprache eine Reihe anderer Kommunikationsformen, deren Vor- und Nachteile gegenüber natürlichsprachlicher Kommunikation nicht prinzipiell festgestellt werden können, sondern nur durch die Evaluierung der jeweiligen Kommunikationserfordernisse in einer konkreten Anwendungssituation herausgearbeitet werden können.

Andererseits hoffe ich, den Leser durch diese Einführung dazu in die Lage versetzt zu haben, die in den letzten Jahren erfreulicherweise seltener zu hörenden Pauschalargumente gegen jegliche Form natürlichsprachlicher Systeme leicht entkräften zu können. Im folgenden werden einige der am häufigsten geäußerten Thesen und Antithesen zu natürlichsprachlichen Systemen, wie sie in Wahlster/v. Hahn 1981 zusammengestellt sind, nochmals wiedergegeben:

- These:* Die natürliche Sprache ist nicht den spezifischen Bedingungen der Mensch-Maschine-Kommunikation angepaßt.
- Antithese:* Die natürlichen Sprachen haben bei erweiterten Möglichkeiten zeitlicher und räumlicher Dilation (z.B. Brief, Telephon) je spezifische Register entwickelt, die sich schnell konventionalisiert haben. Ähnlich wird sich auch für die neuen Informationstechnologien eine spezielle 'Rhetorik' entwickeln. Bisher hat sich die natürliche Sprache noch jedem Medium angepaßt.
- These:* Natürlichsprachliche Ein- und Ausgaben sind gegenüber formalsprachlichen zu lang.
- Antithese:* Die breite Beherrschung von implizitem Wissen über Situation, Dialogstadium, Partneereigenschaften etc. führt in NSS zu stark elliptischem und anaphorischem Sprechen. Ausdrücke bisheriger formaler Anfragesprachen müssen dagegen bei Fehlern oder sachlicher Unrichtigkeit meist vollständig wiederholt werden. Außerdem können in der natürlichsprachlichen Kommunikation jederzeit Abkürzungen im Dialog definiert werden. Durch die Möglichkeiten zur Elliptifizierung und Anaphorisierung sind natürlichsprachliche Formulierungen gegenüber formalsprachlichen 'konzeptuell kürzer'.
- These:* Die Vagheit der natürlichen Sprache ist ein Defekt, der die natürliche Sprache für die Mensch-Maschine Kommunikation ungeeignet macht.
- Antithese:* Die Vagheit vieler natürlichsprachlicher Ausdrücke ist kein Defekt, sondern trägt entscheidend zur Nützlichkeit des Kommunikationswerkzeuges natürliche Sprache bei. Oft können die Griceschen Konversationsmaximen (z.B. 'Mache Deinen Beitrag nicht informativer als erforderlich', 'Sei relevant', vgl. Grice 1975) nur durch die Verwendung von Vagheit befolgt werden.
- These:* Das Verhalten natürlichsprachlicher KI-Systeme ist u.a. durch die Verwendung komplexer Inferenzverfahren für den Benutzer nicht transparent und kontrollierbar (vgl. z.B. Shneiderman 1980).
- Antithese:* Gerade die natürlichsprachliche Kommunikation ermöglicht argumentative Dialoge, in denen alle Äußerungen und Verhaltensweisen der Kommunikations-

partner kritisch hinterfragt und argumentativ abgesichert werden können. Erklärungskomponenten, deren Aufgabe es ist, auf Anfrage eine für den Benutzer verständliche und im jeweiligen Dialogzustand angemessene Erklärung für inferenz-basiertes Systemverhalten zu erzeugen, erhöhen die Transparenz und Kontrollierbarkeit von natürlich sprachlichen KI-Systemen gegenüber konventionellen Dialogsystemen ganz erheblich.

These: Die Menü-Technik ist effizienter als natürlichsprachliche Kommunikation.

Antithese: Stark standardisierte Kommunikation ist ein Sonderfall von natürlicher Kommunikation, z.B. Telefongesprächseröffnungen, Begrüßungen, Geschäftsbrief. Insofern hängt der Grad der Ritualisierung und damit der Einsatz von Menü-Techniken davon ab, wie tief strukturiert und wie verzweigt die Handlungsmuster in der speziellen Anwendungssituation sind.

Obwohl für die Hauptkomponenten eines natürlichsprachlichen Systems heute eine Auswahl von bewährten Methoden und Techniken zur Verfügung stehen, von denen einige in dieser Einführung vorgestellt werden konnten, sind noch viele Einzelprobleme, auch in klassischen Gebieten wie der lexikalischen Analyse und des Parsing ungeklärt und erfordern Grundlagenforschung von Informatikern mit guten KI-Kenntnissen in enger Zusammenarbeit mit Linguisten und Psychologen. Wie bereits in Abschnitt 1.1. festgestellt wurde, ist man beim derzeitigen Forschungsstand von einem 'natürlichsprachlichen Universalsystem' noch weit entfernt. Andererseits sind für viele industrielle Anwendungen, die eine geringe Komplexität der Sprachverstehens- bzw. Generierungsprozesse erfordern, heute produktreife NSS realisierbar.

Nach Hart 1982 läßt sich der derzeitige Entwicklungsstand durch die in Fig. 26 dargestellte Graphik charakterisieren. Es gibt einerseits effiziente und mit relativ geringem Betriebsmittel-

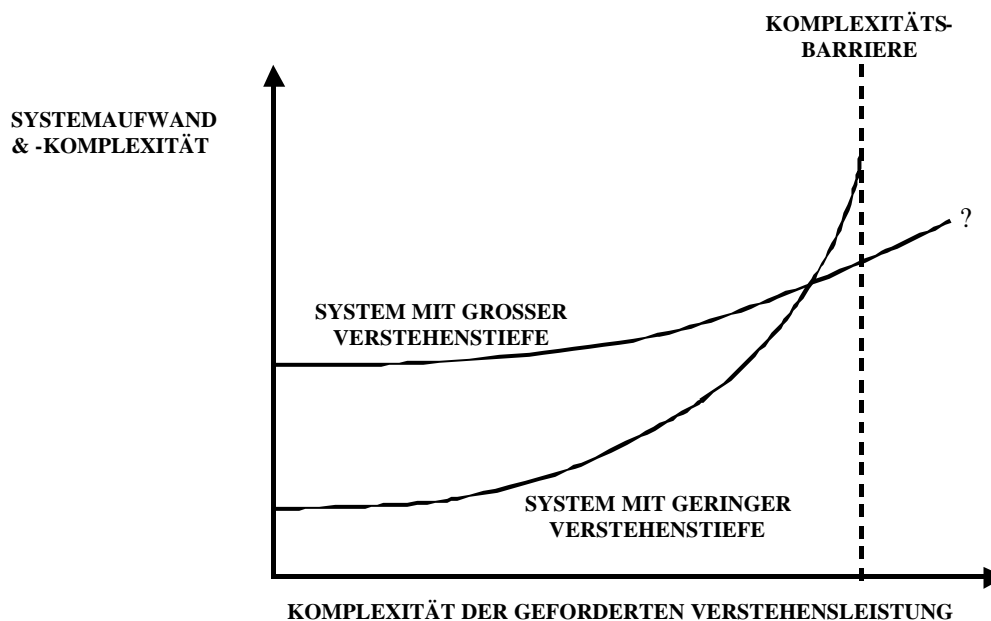


Fig. 26: Zusammenhang zwischen Systemkomplexität und geforderter Verstehensleistung

aufwand realisierbare Verfahren für Aufgaben, die Verstehensleistungen von geringerer Komplexität erfordern, und andererseits Verfahren mit sehr hohen Betriebsmittelanforderungen, durch die auch schwierige Aspekte des menschlichen Sprachverhaltens rekonstruiert werden können. Während man mit dem erstgenannten Typ von Verfahren bei steigender Komplexität der geforderten Verstehensleistung recht bald an eine *Komplexitätsbarriere* stößt, gelingt es mit dem zweiten Typ von Verfahren, der für einfache Aufgaben unnötig aufwendig ist, diese Komplexitätsbarriere zu durchbrechen. Als mittelfristige Forschungsstrategie schlägt Hart daher vor, in anwendungsorientierten Systemen bewährte Techniken aus Systemen mit geringerer Verstehentiefe und Systemen mit großer Verstehentiefe in mehrstufigen Systemen zu kombinieren.

Ich hoffe, daß sich einige der Leser dieser Einführung den faszinierenden Fragestellungen der sprachorientierten KI-Forschung zuwenden und mithelfen, die angesprochenen mehrstufigen Systeme zu realisieren, die anwendungsorientierten Verfahren weiter zu verbessern und die zahlreichen offenen Probleme der sprachorientierten KI-Forschung zu lösen.

DANKSAGUNG

Für zahlreiche Anregungen und Hinweise möchte ich allen Mitarbeitern der Forschungsstelle für Informationswissenschaft und Künstliche Intelligenz an der Universität Hamburg danken. Mein besonderer Dank gilt Frau A. Carstensen, Frau A. Scherff und Frau M. Mühlenberg für die sorgfältige Ausführung der Schreibarbeiten sowie Herrn J. Gottschalk und Frau S. Vogt für die gewissenhafte Ausführung von graphischen Arbeiten.

Die vorliegende Einführung möchte ich meiner Frau Doris Wahlster widmen, die meiner Schreibtischarbeit auch an Wochenenden und Abenden dieses Sommers großes Verständnis entgegengebracht hat.

LITERATUR

1) WICHTIGE SAMMELBÄNDE UND MONOGRAPHIEN

Obwohl Berichte, Mitteilungen und Memos der KI-Forschungszentren und als Institutsberichte veröffentlichte Dissertationen die aktuellste und detaillierteste Information beinhalten, wird zunächst auf allgemein in Buchform zugängliche Sammelbände und Monographien hingewiesen, da Institutsberichte oft schwer zugänglich, schnell vergriffen und teilweise sogar geheim (Industrie, Militär) sind.

Für eine erste Einführung sind die im folgenden genannten Bücher ausreichend; für eine eigenständige Forschungstätigkeit ist das Studium der einschlägigen Institutsberichte, Tagungsbände und Fachzeitschriften allerdings unerlässlich. Bei der notwendigen Selektion wurden neuere Publikationen sowie deutschsprachige Werke bevorzugt.

Barr, A., Feigenbaum, E.A. (eds.) (1981): The handbook of artificial intelligence. Vol. I, Kap. IV: Understanding natural language. S. 223-321; Kap. V: Understanding spoken language, S. 323-361, Los Altos: Kaufmann.

- Barr, A., Feigenbaum, E.A. (eds.) (1982): The handbook of artificial intelligence. Vol. II, S. 164-170, S. 238-239, S. 247-253, S. 370-374, Los Altos: Kaufmann.
- Bobrow, D., Collins, A.M. (eds.) (1975): Representation and understanding: Studies in cognitive science. New York: Academic.
- Bolc, L. (ed.) (1978). Natural language communication with computers. Berlin, Heidelberg, New York: Springer.
- Bolc, L. (ed.) (1980): Natural language question answering systems. München/London: Hanser/Macmillan.
- Bolc, L. (ed.) (1980): Natural language based computer systems. München/London: Hanser/Macmillan.
- Bolc, L. (ed.) (1980). Representation and processing of natural language. München/London: Hanser/Macmillan.
- Brady, M. (ed.) (1982): Computational theories of discourse. Cambridge: MIT Press.
- Charniak, E., Wilks, Y. (eds.) (1976): Computational semantics. An introduction to artificial intelligence and natural language comprehension. Amsterdam: North-Holland.
- Christaller, T., Metzger, D. (eds.) (1979, 1980) : Augmented transition network Grammatiken, Teil I und II, Berlin: Einhorn.
- Cohen, P.R., Feigenbaum, E.A. (eds.) (1982): The handbook of artificial intelligence. Vol. III, S. 65-74, Los Altos: Kaufmann
- Colby, K.M. (1975): Artificial paranoia: a computer simulation of paranoid processes. New York: Pergamon.
- Davey, A. (1978): Discourse production. A computer model of some aspects of a speaker. Edinburgh: Univ. Press.
- Eisenberg, P. (ed.) (1976): Maschinelle Sprachanalyse. Beiträge zur automatischen Sprachverarbeitung 1. Berlin, New York: de Gruyter.
- Eisenberg, P. (ed.) (1977): Semantik und Künstliche Intelligenz. Beiträge zur automatischen Sprachverarbeitung 2. Berlin, New York: de Gruyter.
- Findler, N.V. (ed.) (1979): Associative networks. Representation and use of knowledge by computers. New York: Academic.
- Freedle, R.O. (ed.) (1977): Discourse production and comprehension. Norwood: Ablex.
- Gallaire, H., Minker, J. (eds.) (1978): Logic and data bases. New York: Plenum.
- Gregg, L.W. (ed.) (1974): Knowledge and cognition. New York: Wiley.
- Hirst, G. (1981): Anaphora in natural language understanding: a survey. Berlin, Heidelberg, New York: Springer.

- Joshi, A., Webber, B.L., Sag, I. (eds.) (1981): Elements of discourse understanding. Cambridge: Cambridge Univ. Press.
- Kolvenbach, M., Lötscher, A., Lutz, H.D. (eds.) (1979): Künstliche Intelligenz und natürliche Sprache. Sprachverstehen und Problemlösen mit dem Computer. Tübingen: Narr.
- Krause, J. (1982): Mensch-Maschine-Interaktion in natürlicher Sprache. Evaluierungsstudien zu praxisorientierten Frage-Antwort-Systemen und ihre Methodik. Tübingen: Niemeyer.
- Lea, W. (ed.) (1980): Trends in speech recognition systems. Englewood Cliffs: Prentice Hall.
- Lehnert, W.C. (1978): The process of question answering: a computer simulation of cognition. Hillsdale: Lawrence Erlbaum.
- Lehnert, W.C., Ringle, M. (eds.) (1981): Knowledge representation for natural language processing. Hillsdale: Lawrence Erlbaum.
- Marcus, M.P. (1980): A theory of syntactic recognition for natural language. Cambridge: MIT Press.
- Metzing, D. (ed.) (1980): Frame conceptions and text understanding. Berlin, New York: de Gruyter.
- Metzing, D. (ed.) (1981): Dialogmuster und Dialogprozesse. Hamburg: Buske.
- Minsky, M. (ed.) (1968): Semantic information processing. Cambridge: MIT Press.
- Morik, K. (1982): Überzeugungssysteme der Künstlichen Intelligenz. Validierung vor dem Hintergrund linguistischer Theorien über implizite Äußerungen. Tübingen: Niemeyer 1982.
- Normen, D. A., Rumelhart, D.E. and the LNR Research Group (eds.) (1975): Explorations in cognition. San Francisco: Freeman.
- Rieger, B. (ed.) (1982): Empirical semantics. Bochum: Brockmeyer.
- Ritchie, G.D. (1980): Computational grammar. An artificial intelligence approach to linguistic description. Sussex: Harvester.
- Rollinger, C.-R., Schneider, H.-J. (eds.) (1980): Inferenzen in natürlichsprachlichen Systemen der Künstlichen Intelligenz. Berlin: Einhorn.
- Rustin, R. (ed.) (1973): Natural language processing. Englewood Cliffs: Prentice Hall.
- Sager, N. (1981): Natural language information processing: a computer grammar for English and its application. Reading: Addison-Wesley.
- Schank, R.C. (ed.) (1975): Conceptual information processing. Amsterdam: North-Holland.

- Schank, R.C., Abelson, R.P. (1977): Goals, plans, scripts and understanding: an enquiry into human knowledge structures. Hillsdale: Lawrence Erlbaum.
- Schank, R.C., Colby, K.M. (eds.) (1973): Computer models of thought and language. San Francisco: Freeman.
- Schank, R.C., Riesbeck, C.K. (1981). Inside computer understanding. Hillsdale: Lawrence Erlbaum.
- Schank, R.C. (1982): Dynamic memory. New York: Cambridge Univ. Press.
- Spiro, R., Bruce, B., Brewer, W. (eds.) (1980): Theoretical issues in reading comprehension. Hillsdale: Lawrence Erlbaum.
- Tennant, H. (1981): Natural language processing. An introduction to an emerging technology. New York: Petrocelli.
- Wahlster, W. (1981b): Natürlichsprachliche Argumentation in Dialogsystemen. KI-Verfahren zur Rekonstruktion und Erklärung approximativer Inferenzprozesse. Berlin, Heidelberg, New York: Springer.
- Walker, D.E. (ed.) (1978): Understanding spoken language. New York: North-Holland.
- Waterman, D., Hayes-Roth, F. (eds.) (1978): Pattern-directed inference Systems. New York: Academic.
- Wettler, M. (1980): Sprache, Gedächtnis, Verstehen. Berlin, New York: de Gruyter.
- Winograd, T. (1972): Understanding natural language. New York: Academic.
- Winograd, T. (1983): Language as a cognitive process. Vol. I: Syntax, Reading: AddisonWesley.
- Woods, W.A. (1979): Semantics for a question answering system. New York: Garland. (Buchveröffentlichung der Dissertation von 1967).
- Zampoli, A. (ed.) (1977): Linguistic structures processing. Amsterdam: North-Holland.

2) NEUERE ÜBERSICHTEN

Die folgende Auswahl enthält Übersichten zur sprachorientierten KI-Forschung, die seit 1979 erschienen sind:

- Charniak, E. (1981): Six topics in search of a parser: An overview of AI language research. In: Proc. of the IJCAI-81, Vancouver, S. 1079-1087.
- Eisenberg, P. (1980): Computerlinguistik. In: Althaus, H.P., Henne, H., Wiegand, H.E. (eds.): Lexikon der germanistischen Linguistik. Tübingen: Niemeyer, S. 809-815.

- Fausser, A., Roesner, D. (1979) : Computational Linguistics in West-Germany - a selected bibliography. Univ. Stuttgart, Institut für Informatik.
- Fausser, A., Rathke, Chr. (1981): Studie zum Stand der Forschung über natürlichsprachliche Frage/Antwort-Systeme. Univ. Stuttgart, Institut für Informatik.
BMFT-Forschungsbericht ID-81-006.
- Grosz, B.J. (1979): Utterance and objective: Issues in natural language communication. In: Proc. of the IJCAI-79, Tokio, S. 1067-1076.
- Hendler, J., Kehler, T.P., Michaelis, P.R., Philips, B., Ross, K.M., Tennant, H.R. (1981) : Issues in the development of natural-language front-ends . In: Proc. of the National Computer Conference, Chicago, Arlington: AFIPS Press, S. 643-648.
- Hendrix, G.G. (1980): Future prospects for computational linguistics. In: Proc. of the Annual Conference of the ACL, Philadelphia.
- Kaplan, S.J. (ed.) (1982): Special Section-Natural Language. In: SIGART Newsletter, Nr. 79, Januar 1982, S. 27-109
- Lenders, W. (1980): Linguistische Datenverarbeitung Stand der Forschung: In: Deutsche Sprache, 3, S. 213-264.
- Mann, W.C., Bates, M., Grosz, B., McDonald, D.O., McKeown, K.R., Swartout, W.R. (1981): Text generation: The state of the art and the literature. Univ. of Southern California, Information Sciences Institute, Report ISI/RR-81-101.
- Schank, R.C., Lehnert, W. (1979): Review of natural language processing. In: Wegner, P. (ed.): Research Directions in Software Technology. Cambridge: MIT Press, S. 750-766.
- Wahlster, W. (1981a): Natürlichsprachliche KI-Systeme: Entwicklungsstand und Forschungsperspektive. In: Siekmann, J.H. (ed.): GWAI-81. German Workshop on Artificial Intelligence, Bad Honnef; Berlin, Heidelberg, New York: Springer, S. 50-68.
- Waltz, D.L. (1981) : The state of the art in natural language understanding. Coordinated Science Laboratory, University of Illinois, Working Paper 27, Januar 1981.

3) ZITIERTE AUFSÄTZE

- Appelt, D.E. (1981): Planning natural language utterances to satisfy multiple goals. Ph. D. Dissertation, Stanford University, Stanford, Ca.
- Bates, M. (1980): Theorie und Praxis der ATN-Grammatiken. In: Christaller, Th., Metzger, D. (eds.): Augmented Transition Network Grammatiken. Bd. 2, Berlin: Einhorn.
- Birnbaum, L. , Selfridge, M. (1979) : Problems in conceptual analysis of natural language. Yale University, Computer Science Department, Research Report Nr. 168.

- Bobrow, D.G., Kaplan, R.M., Kay, M., Norman, D.A., Thompson, H., Winograd, T. (1977): GUS - a frame-driven dialog system. In: *Artificial Intelligence*, 8, 2, S. 155-173
- Bobrow, D.G., Winograd, T. (1977): An overview of KRL, a knowledge representation language. In: *Cognitive Science*, 1, S. 3-46.
- Bobrow, R.J. (1978). The RUS system. In: Webber, B.L., Bobrow, R.J. (eds.): *Research in natural language understanding. Quarterly progress report Nr. 3*, Bolt Beranek and Newman, Cambridge, Ma., BBN Report Nr. 3878.
- Bobrow, R.J. , Webber, B.L. (1980): Knowledge representation for syntactic/semantic processing. In: *Proc. of the 1st national conference on Artificial Intelligence*, Stanford, Ca., S. 316-323
- Brachman, R.J. (1978): A structural paradigm for representing knowledge. Solt Beranek and Newman, Cambridge, Ma., BBN Report Nr. 3605.
- Brachman, R.J., Bobrow, R.J., Cohen, P.R., Klovstadt, J.W., Webber, B.L., Woods, W.A. (1979): *Research in natural language understanding. Annual report*. Bolt Beranek and Newman, Cambridge, Ma., BBN Report Nr. 4274.
- Brachman, R.J., Smith, B.C. (1980): Special issue on knowledge representation. In: *SIGART Newsletter Nr. 70*.
- Burton, R.R. (1976): *Semantic grammar: An engineering technique for constructing natural language understanding systems*. Bolt Beranek and Newman, Cambridge, Ma., BBN Report Nr. 3453.
- Busemann, S. (1982): *Probleme der automatischen Generierung deutscher Sprache*. Universität Hamburg, Forschungsstelle für Informationswissenschaft und Künstliche Intelligenz, Memo ANS-8.
- Christaller, T., Metzger, D. (1982): Parsing interactions and a multi-level parser formalism based on cascaded ATNs. In: Sparck Jones, K., Wilks, Y. (ed.): *Automatic natural language parsing*. Cognitive Studies Center, Essex.
- Church, K. (1980). *On memory limitations in natural language processing*. MIT, LCS, Cambridge, Mass., Technical Report Nr. 245.
- Cohen, P. (1978): *On knowing what to say: Planning speech acts*. University of Toronto, Toronto, Technical Report Nr. 118.
- Collins, A., Warnock, E.H., Aiello, N.E., Miller, M.C. (1975): Reasoning from incomplete knowledge. In: Bobrow, D.G. , Collins, A. (eds.) : *Representation and understanding: Studies in cognitive science*. N.Y.: Academic.
- Cullingford, R.E., Krueger, M.W., Selfridge, M., Bienkowski, M.A. (1982): Automated explanations as a component of a computer-aided design system. In: *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-12, Nr. 2, S. 168-181.

- Eisenstadt, M. (1979): Alternative Parser für konzeptuelle Abhängigkeitsgrammatiken. In: Christaller, Th., Metzling, D. (eds.): Augmented Transition Network Grammatiken, Bd. 1, Berlin: Einhorn.
- Fausser, A. (1979): Überlegungen zu einer Theorie des Parsing. In: Christaller, Th, Metzling, D. (eds.): Augmented Transition Network Grammatiken, Bd. 1 , Berlin: Einhorn.
- Fausser, A. (1980): Inferenz und Kohärenz in KIPUS. In: Rollinger, C.-R., Schneider, H.-J. (eds.): Inferenzen in natürlichsprachlichen Systemen der Künstlichen Intelligenz. Berlin: Einhorn.
- Fillmore, C. (1968): The case for case. In: Bach, E., Harms, R. (eds.): Universals in linguistic theory. N.Y.: Holt, Rinehart and Winston.
- Görz, G. (1979): Kontrollstrukturen und ATN. In: Christaller, Th., Metzling, D. (eds.): Augmented Transition Network Grammatiken. Bd. 1, Berlin: Einhorn.
- Görz, G., Beckstein, C. (1980): GLP: Ein linguistischer Prozessor. Universität Erlangen Nürnberg, Rechenzentrum, Interner Arbeitsbericht Nr. 125.
- Grice, H.P. (1975): Logic and conversation. In: Cole, P., Morgan, J. L. (eds.): Syntax and semantics. Vol. 3, Speech acts, NY.: Academic.
- Grosz, B. (1977): The representation and use of focus in dialog understanding SRI International, Menlo Park, Ca., Technical Note Nr. 151.
- Haas, N., Hendrix, G.G. (1980): An approach to acquiring and applying knowledge. In: Proc. of the 1st national conference on Artificial Intelligence, Stanford, Ca., S. 235-239
- Habel , C.U., Rollinger, C.-R., Schmidt, A., Schneider, H.-J . (1980): A logic-oriented approach to automatic text understanding. In: Bolc, L. (ed.): Natural language based computer systems. München, London: Hanser/Macmillan.
- Hahn, U., Kuhlen, R. , Reimer, U. (1982): Konzeption und Aufbau des automatischen Textkondensierungssystems TOPIC. Universität Konstanz, Informationswissenschaft, Bericht TOPIC-1/82.
- v. Hahn, W. (1978): Überlegungen zum kommunikativen Status und der Testbarkeit von natürlichsprachlichen Artificial-Intelligence-Systemen. In: Sprache und Datenverarbeitung 1, S. 145-169.
- v. Hahn, W. (1979): Über Dialogkohärenz in natürlichsprachlichen AI-Systemen. Universität Hamburg, Germanisches Seminar, HAM-RPM Bericht Nr. 8.
- v. Hahn, W. , Hoepfner, W. , Jameson, A. , Wahlster, W. (1980): The anatomy of the natural language dialogue system HAM-RPM. In: Bolc, L. (ed.) : Natural language based computer systems. München, London: Hanser/Macmillan.
- Hart, P.E. (1982): Directions for AI in the eighties. In: SIGART Newsletter, Nr. 79, S. 11-16.

- Hayes, P.J., Carbonell, J.J. (1981): Multi-strategy construction-specific parsing for flexible data base query and update. In: Proc. of the 7th International Joint Conference on Artificial Intelligence, Vancouver, S. 432-439.
- Hein, H.W. (1981): A system for understanding continuous German speech. In: Siekmann, J. (ed.): GWAI-81. German workshop on Artificial Intelligence, Bad Honnef, Berlin, Heidelberg, N.Y.: Springer.
- Hendrix, G.G. (1976): Semantic aspects of translation. In: Walker, D.E. (ed.): Speech understanding research. SRI International, Menlo Park, Ca., final technical report.
- Hendrix, G.G. (1977): The LIFER manual: A guide to building practical natural language interfaces. SRI International, Menlo Park, Ca., Technical Note 138.
- Hewitt, C. (1971): Procedural embedding of knowledge in PLANNER. In: Proc. of the 2nd International Joint Conference on Artificial Intelligence, S. 169-182.
- Hobbs, J.R. (1976): Pronoun resolution. City University of New York, Dept. of Computer Science, Research Report Nr. 76-1.
- Hoepfner, W. (1982a): A multilayered approach to the handling of word formation. In: Horecky, J. (ed.): COLING-82. The ninth International Conference on Computational Linguistics, Prague. Amsterdam: North-Holland.
- Hoepfner, W. (1982b): ATN-Steuerung durch Kasusrahmen. In: Wahlster, W. (ed.): GWAI-82. German workshop on Artificial Intelligence, Bad Honnef, Berlin, Heidelberg, N.Y.: Springer.
- Hussmann, M., Genzmann, H. (1982): Performanz-orientiertes Parsing - Ansätze zur robusten Analyse natürlicher Sprache. Universität Hamburg, Forschungsstelle für Informationswissenschaft und Künstliche Intelligenz, Memo GEN-5.
- Jameson, A., Hoepfner, W., Wahlster, W. (1980): The natural language System HAM-RPM as a hotel manager: Some representational prerequisites. In: Wilhelm, R. (ed.): GI-10. Jahrestagung, Saarbrücken, Berlin, Heidelberg, N.Y.: Springer, S. 459-473.
- Jameson, A., Wahlster, W. (1982): User modelling in anaphora generation: Ellipsis and definite description. In: Proc. of the 1st European Conference on Artificial Intelligence, Orsay, S. 222-227.
- Kaplan, R.M. (1973): A general syntactic processor. In: Rustin, R. (ed.): Natural language processing. NY.: Algorithmics.
- Kay, M. (1973): The MIND system. In: Rustin, R. (ed.): Natural language processing. N.Y.: Algorithmics.
- Kwasny, S.C. (1980): Treatment of ungrammatical and extra-grammatical phenomena in natural language understanding. Ph.D. Dissertation, Ohio State University.

- LeFavre, R.A. (1977): FUZZY reference manual. Rutgers University, New Brunswick, Dept. of Computer Science.
- Liebisch, G. (1980): Das Konzept der natürlichsprachlichen Datenbankschnittstelle NATAN. In: Krallmann, D. (ed.): Dialogsysteme und Textverarbeitung. Essen.
- McDonald, D.D. (1980): Natural language production as a process of decision making under constraint. Ph.D. Dissertation, MIT, Cambridge, Mass.
- McKeown, K.R. (1982): Generating natural language text in response to questions about database structure. Ph.D. Dissertation, University of Pennsylvania, Philadelphia.
- Metzing, D. (1981): Tools for a procedural dialog model and some problems of application. In: Hägglund, S., Hein, U. (eds.): Proc. of the International Workshop on Models of Dialogue: Theory and Application. Linköping University.
- Minsky, M. (1975): A framework for representing knowledge. In: Winston, P. (ed.): The psychology of computer vision. NY.: McGraw-Hill.
- Nebel, B., Marburger, H. (1982): Das natürlichsprachliche System HAM-ANS: Intelligenter Zugriff auf heterogene Wissens- und Datenbasen. In: Nehmer, J. (ed.): GI-12. Jahrestagung, Kaiserslautern. Berlin, Heidelberg, NY.: Springer.
- Ott, N. (1979): Das experimentelle auf natürlicher Sprache basierende Informationssystem USL. In: Nachrichten für Dokumentation, 30, 3, S. 129-139.
- Parkison, R.C., Colby, K.M., Faught, W.S. (1977): Conversational language comprehension using integrated pattern-matching and parsing. In: Artificial Intelligence, 9, S. 111-134.
- Reddy, R., Erman, L., Fennell, R., Neely, R. (1976): The HEARSAY speech understanding system: An example of the recognition process. In: IEEE Transactions on Computers, C-25, S. 427-431.
- Rieger, C., Small, S. (1979): Word expert parsing. In: Proc. of the 6th International Joint Conference on Artificial Intelligence, S. 723-728.
- Riesbeck, C. (1975): Conceptual analysis. In: Schank, R.C. (ed.): Conceptual information processing. Amsterdam: North-Holland.
- Roberts, R.B., Goldstein, I.P. (1977): The FRL manual. MIT, AI Laboratory, Memo Nr. 409.
- Shneiderman, B. (1980): Natural vs. precise concise languages for human operation of computers: Research issues and experimental approaches. In: Proc. of the 18th annual meeting of the ACL, Philadelphia, S. 139-141.
- Sidner, C. (1979): Towards a computational theory of definite anaphora comprehension in English discourse. MIT, AI Laboratory, Mass., Report Nr. 537.
- Small, S. (1981): Viewing word expert parsing as linguistic theory. In: Proc. of the 7th International Joint Conference on Artificial Intelligence, Vancouver, S. 70-76.

- Sridharan, N.S. (1978): AIMDS user manual. Version 2, Rutgers University, New Brunswick, Dept. of Computer Science, Report Nr. CBM-TR-89.
- Wahlster, W. (1979): ATNs und die semantisch-pragmatische Steuerung der Analyse und Generierung natürlicher Sprache. In: Christaller, Th., Metzging, D. (ed.): Augmented Transition Network Grammatiken, Bd. 1, Berlin: Einhorn.
- Wahlster, W., v. Hahn, W. (1981): Mensch-Maschine-Kommunikation auf der Basis natürlicher Sprache. Universität Hamburg, Forschungsstelle für Informationswissenschaft und Künstliche Intelligenz, Memo GEN-2.
- Wahlster, W. (1982): Aufgaben, Standards und Perspektiven sprachorientierter KI-Forschung. Einige Überlegungen aus informatischer Sicht. In: Batori, I., Krause, J., Lutz, H.D. (eds.): Linguistische Datenverarbeitung. Versuch einer Standortbestimmung im Umfeld von Informationslinguistik und Künstlicher Intelligenz. Tübingen: Niemeyer.
- Waltz, D.L. (1978): An English question answering system for a large relational database. In: Communications of the ACM, 21, 7, S. 526-539.
- Webber, B.L. (1978): A formal approach to discourse anaphora. Bolt Beranek and Newman, Cambridge, Mass., BBN Report 3761.
- Wilensky, R. (1978): Understanding goal-based stories. Ph.D. Dissertation, Yale University, Computer Science Department, Report Nr. 140.
- Winograd, T. (1974): Five lectures on artificial intelligence. Stanford University, AI Laboratory, Memo 246.
- Witschas, W., Zänker, F., Helbig, H. (1980): FAS-80 - A natural language information system. In: Proc. of the 1st International workshop on natural communication with computers, Warschau, S. 120-124.
- Woods, W.A. (1970): Transition network grammars for natural language analysis. In: Communications of the ACM, 13, S. 591-606.
- Woods, W.A., Kaplan, R., Nash-Webber, B.L. (1972): The lunar sciences natural language information system. Bolt Beranek and Newman, Cambridge, Mass., BBN Report 2265.
- Woods, W.A. et al. (1976): Speech understanding systems: Final report. Bolt Beranek and Nefflan, Cambridge, Mass., Report Nr. 3438.
- Woods, W.A. (1980): Cascaded ATN grammars. In: AJCL, 6, 1, January-March 1980, S. 1-12.