

Wissensrepräsentation als Hypertexte

Beispiel und Problematik einer Verstehenstechnologie*

Burghard Rieger

Lehrstuhl für Computerlinguistik

Fachbereich II: Sprach- und Literaturwissenschaften

Universität Trier

Einleitung

Der inzwischen vollzogene Übergang von der *Daten-* zur *Informationsverarbeitung* ist verbunden mit der elektronischen Speicherung und weitgehenden Verfügbarkeit von (fachspezifischen) Teilbereichen des vorhandenen Wissens in Daten- und Informationsbanken, die z.T. weltweit abrufbar sind. Der weiterführende Schritt von der *Informations-* zur *Wissensverarbeitung* ist derzeit als Herausforderung erkannt und wird inzwischen als zukunftsweisende Aufgabenstellung in der Forschung angegangen. Zur Lösung der hierbei anstehenden Probleme im Schnittbereich von Informatik, kognitiver Psychologie, Computerlinguistik und den Semantiktheorien der Sprachwissenschaft und Philosophie liegen inzwischen Ansätze vor, denen—bei allen Unterschieden im Zusammenhang der Hypothesenentwicklung, der Theorienbildung, der Modellierung und der experimentellen Überprüfung oder praktischen Umsetzung in Anwendungen—der Einsatz von Computern gemeinsam ist. Dabei steht die interdisziplinäre Untersuchung solche Strukturen und Prozesse im Vordergrund, die natürliche (oder künstliche) informationsverarbeitende Systeme in die Lage versetzen, sprachliche Zeichen zu entwickeln, zu erlernen oder auch bloß zu verwenden, wenn sie miteinander kommunizieren. Die Diskussion unterschiedlicher Formate bei der Darstellung von Wissen und der Möglichkeiten seiner Verarbeitung hat inzwischen zur Spezifizierung von (symbolischen und distributionellen) Repräsentationen geführt, auf denen Operationen definierbar sind, die als Modellierungen von Verarbeitungsprozessen propositionaler Zusammenhänge des Wissens und konzeptueller Strukturen des Gedächtnisses gelten und in sog. Sprach-verstehenden Systemen der künstlichen Intelligenz Forschung (*KI*) ebenso wie in kognitionstheoretisch motivierten Analysen natürlichsprachlicher Bedeutung etwa der Situationssemantik (*SS*) unternommen werden.

Obwohl die auf diesem Gebiet tätigen Wissenschaftler deutlich machen, daß unsere Kenntnisse über die vielfältigen und höchst komplexen intellektuellen Aktivitäten, die heute unter der Bezeichnung *Kognition* zusammengefaßt werden, noch höchst lückenhaft sind, scheinen die gemeinsamen Bemühungen sich zunehmend auf Probleme des *Sprachverstehens*, seine Modellierung bzw. Simulation als *Wissensverarbeitung* und die selbstorganisierende Generierung der dazu benötigten dynamischen Strukturen als *Semiosis* zu konzentrieren.

Jüngste Fortschritte dieser interdisziplinären Forschungen in der maschinellen Verarbeitung natürlicher Sprache (*NLP*)¹ sowie die Umsetzung ihrer zunächst theoretischen Ergebnisse in praktische Anwendungszusammenhänge etwa in der Entwicklung von Arbeitsumgebungen, Expertensystemen, Ausbildungs- und Denkwerkzeugen, etc. sind bisher zwar im wesentlichen auf den Bereich der Ingenieur- und Naturwissenschaften beschränkt geblieben. Eine Übertragung auf die und eine Anwendung in den Geisteswissenschaften scheint inzwischen aber immerhin

*Dieses Skript basiert auf einem Vortrag "Wissensverarbeitung als Verstehenstechnologie: die computerlinguistische Perspektive" auf der 5.SAVE-Frühjahrstagung 1990 in Karlsruhe und führt einige dort entwickelte Überlegungen weiter.

Erschienen in: Jäger, L./ Switalla, B. (Hrsg.): Germanistik in der Mediengesellschaft. München (W. Fink) 1994, S. 373–403.

¹*NLP*: natural language processing

vorstellbar angesichts der *Hypertext*-Systeme im Bereich der Informations- und Wissensverarbeitung. Durch sie werden neue technologische Entwicklungen zunehmend auch für jene Disziplinen attraktiv, deren Grundlage das Verstehen natürlichsprachlicher Texte ist sowie das in ihnen überlieferte—durch unterschiedliche Interpretationen und wandelndes Verständnis—sich verändernde Wissen. Die *Germanistik 2000* wird sich dieser neuen Technologien hoffentlich schon bald ebenso selbstverständlich bedienen, wie deren *germanistische* Relevanz von heutigen Vertretern des Fachs noch übersehen wird.

1 Sprache und Systemsituation

1.1 Das gesamte Wissen unserer Zeit ist (oder könnte doch) in Texten aufbewahrt oder wiedergegeben werden. Seit ihrer schriftsprachlichen Fixierung und insbesondere seit der Entwicklung und Vervollkommnung der drucktechnischen Möglichkeiten hat sich die natürliche Sprache als das universelle Repräsentationsmedium erwiesen. Es ist den rein abbildenden (auch neueren) Formen bloßer klanglicher und/oder visueller Darstellung durch seine vergleichsweise hohe semiotische Abstraktheit überlegen: *Sprache* konnte und kann nicht nur den unterschiedlichsten Aufgaben entsprechen, die eine gleichermaßen auf Generalisierung wie Spezifizierung ausgerichtete, umfassende Informations- und Wissensvermittlung stellt; *Sprache* erwies und erweist sich darüber hinaus auch als ein überaus flexibles und anpassungsfähiges Medium sehr grundlegender, zudem veränderbarer Strukturierungen von Welt und Wirklichkeit (oder Ausschnitten daraus), dessen sich der Mensch zur Orientierung wie zum Überleben bisher sehr erfolgreich bediente.

Im Licht der Theorie informationsverarbeitender Systeme wird die Bedeutung des Sprachvermögens als Teil der kognitiven Leistungen des Menschen erkennbar. Der Beitrag, den die Sprachkompetenz eines (natürlichen wie künstlichen) kognitiven Systems bei der Verarbeitung seiner Umgebungsinformationen liefert, ist kaum zu überschätzen. Im Vergleich zu anderen kognitiven Systemen, mit denen der Mensch als Lebewesen konkurriert, zeigt sich, daß die menschliche Überlegenheit bei der Erfüllung von Ausgleichs- und Anpassungsfunktionen zur Systemstabilisierung und -erhaltung vornehmlich in solchen Bereichen liegt, die mit der Verarbeitung von Wissen, seinem beschleunigten Erwerb, seiner besonderen Repräsentation und seiner allgemeinen Verfügbarkeit durch *sprachliche* Vermittlung zusammenhängen.

So scheint das kognitive System *Mensch*, das gerade durch sein Sprachvermögen vor anderen informationsverarbeitenden Systemen ausgezeichnet scheint, durch eben diese besondere Fähigkeit auch dazu verleitet zu werden, die Welt, die es mit anderen Lebewesen als gemeinsame *Lebenswelt* teilt, als nahezu ausschließlich *seine Umwelt* mißzuverstehen. Die immer noch weitgehend anthropozentrische Nutzung ihrer Ressourcen hat die Welt ja nicht nur verschmutzt und anderen Systemen als mögliche Lebenswelt entzogen, sondern diese Nutzung hat—mit zunehmendem technologischen Wissen sogar sich noch beschleunigend—schon den globalen Systemzusammenhang destabilisiert, dessen Teil (nicht dessen Beherrscher oder Kontrollinstanz) der Mensch ist. Das ist zwar eine auch für die wissensbasierte, computerlinguistische Forschung scheinbar irrelevante und daher noch wenig geläufige ökologische Einsicht, sie hat dafür aber in der kognitiven und systemtheoretisch orientierten Sprachphilosophie prominente Proponenten ².

Die besondere Aktualität der sprachorientierten Verstehensforschung durch kognitive Modellierungen auf dem Computer hat in dieser Dimension des Sprachvermögens ihren Grund:

²So etwa BARWISE und PERRY, die ihrer *Situationssemantik* (1983) wünschen "that the book will, in some small way, contribute to a rethinking of the relation of people to the world around them, a world full of constraints and meaning, both for people and for the other beings with whom they share it." (S. xiv).

Bedingung der Möglichkeit des Erwerbs und der Verarbeitung von Wissen zu sein, dessen un-systematische Nutzung—als Interpretation von Welt—zunehmend System-gefährdend, dessen systematische Nutzung aber ohne die Hilfe künstlicher sprachverstehender Systeme inzwischen kaum mehr umfassend möglich scheint.

1.2 Wie alle lebenden Systeme in dieser Welt, die sich aufgrund der ihnen aus ihrer Umwelt zugänglichen Informationen (und deren Verarbeitung) an veränderte Bedingungen anzupassen vermögen, verfügt auch der Mensch über die Fähigkeit, aus solchen Erfahrungen zu lernen und sein Verhalten entsprechend zu ändern. Anders als die übrigen lebenden Systeme jedoch, die die Resultate dieser Verarbeitungsprozesse vornehmlich nur über genetische Veränderungen speichern und weitergeben können zur System-stabilisierenden Erhaltung ihrer Art, verfügt das kognitive System *Mensch* über die Möglichkeit, die quasi *vertikale* Vermittlung *endotopen* Wissens³ über Folgen sich reproduzierender und dabei anpassender Systemgenerationen durch eine sehr viel schnellere und effizientere Verhaltensanpassung zu ergänzen, welche aufgrund einer quasi *horizontalen* Vermittlung *exotopen* Wissens⁴ schon zwischen den Einzelsystemen derselben Generation wirksam werden kann.

Dies geschieht durch situative Repräsentation von *Welt* in Form von Zeichensystemen der Sprache und ihrer Realisierung in Texten. Texte funktionieren dabei wie potentielle *Situationen*, deren *informations-systematische* Besonderheit darin besteht, die Identität der Ort-Zeit-Paare für alle Koordinaten von System und Systemumgebung zu dispensieren. Erst durch diese Abstraktion kann ein *System* in relativer Unabhängigkeit unterschieden werden von seiner *Systemumgebung*, was deren weitere Differenzierung möglich und nötig macht. Danach sind unmittelbare, Ort-Zeit-identische Umgebungen eines Systems (in *Situationen*) von solchen *Quasi*-Umgebungen (in *Repräsentationen*) zu unterscheiden, die wegen ihrer Ort-Zeit-Abstraktion erst aufgrund von Aktualisierungen durch entsprechend angepaßte Systeme wieder zu deren Systemumgebungen werden können.

Für den hier betrachteten Zusammenhang wird dabei—informations-systematisch gesprochen—der Ausschnitt von *Welt*, der einem *System* zugänglich ist, beträchtlich erweitert und zwar um jene Orts-Zeit-versetzten, weil sprachlich repräsentierten und so vermittelten *Quasi*-Umgebungen, die seine textsprachliche Überlieferung als System-transzendierende (*kulturelle*) Einbettung ausmacht. Sie zu aktualisieren, d.h. zu verstehen ist einem sprachverarbeitenden System nur aufgrund seiner Sprachkompetenz möglich, die seine Anpassung (*attunement*) an diese Umgebung darstellt.

Die Aktualisierung solcher *Quasi*-Umgebungen bewirkt dabei nicht nur eine bloße Vermehrung von Wechselbeziehungen der Art, wie sie ein System zu seiner unmittelbaren Systemumgebung unterhält, vielmehr bildet diese Aktualisierung eine neue, *semiotische* Qualität von Systemerfahrung. Sie erlaubt es, Erfahrungen nicht nur zu machen sondern wie Hypothesen anzusetzen und deren Resultate auch wieder zu verwerfen. Diese gleichsam experimentelle Qualität semiotisch repräsentierter Umgebungen, die ein System durch Aktualisierung zu seiner Systemumgebung zu machen vermag, begründet dessen *kognitive* Fähigkeit. Sie läßt sich zusammenfassend als das Vermögen eines informationsverarbeitenden Systems kennzeichnen, seine Erfahrungsergebnisse von den ihnen zugrundeliegenden Erfahrungen zu unterscheiden und diese Resultate in Zeichen und deren Aggregationen als sein Wissen zu repräsentieren.

1.3 Die in Form von Zeichenaggregationen sprachlicher Texte repräsentierten *Quasi*-Umgebungen sind eine der Bedingungen dafür, daß es informationsverarbeitenden Systemen, die über die Fähigkeit der semiotischen Aktualisierung (Sprachkompetenz) verfügen, möglich ist,

³Speicherung und Weitergabe system-genetischer Eigenerfahrung

⁴Identifikation und Übernahme Symbol-repräsentierter Fremderfahrung

sich des Wissens und der Erfahrungen von Generationen ohne deren genetische Vermittlung zu bedienen. Angesichts dieser Möglichkeiten sprachkompetenter, kognitiver Systeme, eigenes und fremdes Wissen sowohl aus *endotopen* als auch *exotopen* Erfahrungszusammenhängen vor allem *horizontal* zu kommunizieren, verliert die auch weiterhin mögliche *vertikale* Vermittlung offenbar an Bedeutung.

Erst die *Sprachkompetenz* als herausragende kognitive Fähigkeit scheint der Spezies *Mensch* dabei ermöglicht zu haben, das *Wissen* als Menge der verfügbaren faktischen und hypothetischen Erfahrungsergebnisse in Form von (gesprochenen oder geschriebenen) Texten zum Gegenstand von faktischen und hypothetischen Erfahrungen zu machen, die dieses *Wissen* selbst verändern. Obwohl das System *Mensch*—wie jedes informationsverarbeitende System—in unmittelbarer Wechselbeziehung zu seiner Systemumgebung steht, enthält diese Umgebung selbst zum Teil auch wieder die (in Zeichenaggregaten symbolisch repräsentierten) Quasi-Umgebungen, deren Aktualisierung die unmittelbare Ort-Zeit-Identität von System und Systemumgebung quasi rekursiv und mittelbar beliebig zu erweitern vermag. Die Unmittelbarkeit dieser Wechselbeziehung beschränkt sich dabei—wie für jedes kognitive System—auf die Ort-Zeit-Identität, während die Vermitteltheit dieser Wechselbeziehung—wie für jedes sprachkompetente System—in ihrer der semiotischen Aktualisierung bedürftigen Zeichenrelation besteht. Sie erlaubt es, neben praktischen auch theoretische, neben faktischen auch hypothetische, neben realen auch fiktive Erfahrungen zu erproben und deren Resultate als Wissen zu vermitteln.

Gerade diese Fähigkeit sprachkompetenter Systeme zur Aktualisierung von Ort-Zeit-versetzten Quasi-Umgebungen, durch die das eigene Erfahrungswissen aus Ort-Zeit-identischen Systemumgebungen sich ständig verändert, kennzeichnet die Offenheit oder Dynamik Sprachverstehender, kognitiver Systeme, für deren pragmatische Eingebettetheit der Begriff der *Situation* kennzeichnend ist⁵.

2 Bedeutung und Situationssemantik

2.1 Trotz ihrer relativ langen Geschichte hat die Sprachphilosophie und Semantik für eine Reihe von Problemen, die sich im Zusammenhang der Analyse, Repräsentation und Verarbeitung natürlichsprachlicher Ausdrücke als "Bedeutung-vermittelndeSSstrukturen stellten, keine oder nur unbefriedigende Antworten zu geben vermocht:

- ▷ Angesichts wahrheitsfunktioneller Modellierung wurde unübersehbar, daß *Bedeutungen* von sprachlichen Ausdrücken uns korrekte Information aber auch Ungenauigkeiten und Irrtümer sowohl über die *externe* Wirklichkeit (als System-Umgebung) als auch über die *innere* Welt (als Systems-Struktur) vermitteln.
- ▷ Aufgrund abbildungstheoretischer Modelle ergab sich, daß *Ausdrücke* in systematischer, d.h. nicht-zufälliger Weise verbunden sein müssen mit Entitäten der externen Wirklichkeit (System-Umgebung) einerseits und mit Komponenten der inneren Welt (System-Struktur) andererseits. Wären sie das nicht, vermöchte die (verbale/schriftliche) Äußerung einer Ausdrucks keine Bedeutung zu vermitteln; wir (als beteiligte Systeme) würden sie nur als Geräusch hören bzw. als Kritzel sehen.
- ▷ Im Rahmen semiotischer Modellbildungen zeigte sich, daß die Fähigkeit, mittels *Sprache* sowohl unsere Gedanken als auch die Wirklichkeit zu strukturieren, auch auf die Sprache selbst müsse angewandt werden können: Ausdrücke lassen sich danach klassifizieren, wie sie einerseits Wirklichkeit (die System-Umgebung) und andererseits unsere Wissen über sie (das System selbst) strukturieren.

⁵vgl. hierzu etwa RIEGER (1989a), S.70–78 und RIEGER/THIOPOULOS (1989)

Aufgrund einer vergleichsweise radikalen Revision einiger Basisannahmen modelltheoretischer Semantiktheorien haben BARWISE und PERRY (1983) die Analyse natürlichsprachlicher Bedeutung als formale Analyse der kommunikativen Zusammenhänge konzipiert, in denen sprachliche Ausdrücke geäußert, interpretiert und verstanden werden. Anders als in den Mögliche-Welt-Semantiken, die ko- und kontextuellen Einbettungen nur als pragmatische Erweiterungen ergänzend einzubeziehen vermochten, ist die derzeit wohl aktuellste Modellbildung in der sprachphilosophischen und sprachwissenschaftlichen Semantik auf ein systemtheoretisch fundiertes Konzept der pragmatischen Kontexte aufgebaut, das *Situation* heißt.

Der Neuansatz läuft auf die Entwicklung einer abstrakten Theorie zur Klassifizierung pragmatischer Kontexte von sprachlichen Ausdrücken hinaus. Als Basis der *Situationsemantik* bildet sie die Grundlage dessen, was den Informations-verarbeitenden (natürlichen wie künstlichen) Systemen (*organisms*) von der sie umgebenden Wirklichkeit und externen Realität als ihre System-Umgebung jeweils zugänglich ist.

2.2 Was über Informations-verarbeitende Systeme und ihre Adaptation an die sie umgebende Wirklichkeit ausgeführt wurde, kann nun für Sprach-verarbeitende (natürliche) Systeme, die in eine zwischen einzelnen Systemen (Individuen) und Klassen von ihnen (Generationen) vermittelnde Sprachgemeinschaft (*linguistic community*) eingebettet sind, zur formal-theoretisch wie empirisch-praktisch zugänglichen Basis erklärt werden, welche die Analyse jener Prozesse erlaubt, aufgrund der sich diese Systeme beschleunigt anzupassen vermögen: anhand und mithilfe von in *Texten* realisierter, sprachlicher Strukturen.

Danach ist das System *Mensch* auf die von ihm beherrschte Sprache eingestellt, so daß ihre Strukturen ihm im Gebrauch die Orientierung liefern bei der Identifikation, Auswahl, Analyse, Klassifikation und Repräsentation von Situationen. Als solche vermag er die für ihn bedeutsamen Komponenten seiner System-Umgebung zu identifizieren und als jenen Ausschnitt von Realität zu erkennen, den er als seine *Um-Welt* erfährt und interpretieren kann. Wenn Realität derart als zunächst heterogene Überdeckung von Situationen gedacht wird, aus der informationsverarbeitende Systeme die ihnen zugänglichen Ausschnitte als ihre System-Umgebungen wahrnehmen, dann ist das Vermögen, einige Situationen als einander ähnlich zu erkennen und in Zusammenhang zu bringen, abhängig von der Fähigkeit, auf bestimmte ihrer Komponenten in gleicher oder ähnlicher Weise zu reagieren, was deren Invarianz oder Gleichförmigkeit ausmacht. Die Konstitution derartiger Zusammenhänge, die von der Struktur und der Einstellung des Informations-verarbeitenden Systems abhängen, bildet aber gleichzeitig auch die Grundlage dafür, aus dem Verständnis und der Kenntnis der Verhältnisse einer Situation auf die Verhältnisse in ähnlichen Situationen zu schließen, die mit diesen durch Invariante verbunden sind⁶.

2.3 Situationen bestehen—je nach Struktur und Einstellung des informationsverarbeitenden Systems, das seine System-Umgebung als situative *Um-Welt* aus der umgebenden Realität konstituiert—aus Komponenten. Diese erscheinen im wesentlichen als individuelle Entitäten, die bestimmte Eigenschaften haben, sich an unterschiedlichen Raum-Zeit-Punkten befinden und in systematischen Relationen zueinander stehen, welche durch die jeweiligen Invarianten bestimmt sind. *Wörter* stellen als Worttypen solche besonderen Invarianten dar, die zur gleichen Zeit an sehr verschiedenen Situationen teilhaben können, aber innerhalb verschiedener Äußerungen als Worttoken erscheinen, wodurch sie in *Texten* sowohl Zusammenhänge

⁶One situation s can contain information about another situation s' only if there is a systematic relation M that holds between situations sharing some configuration of uniformities with s and situations that share some other configuration of uniformities with s' . These uniformities may be physical objects, abstract objects like words, physical or abstract properties or relations, places, times, or other uniformities. But in any case, it is the relationship M that allows us to say that the first situation means the second. (BARWISE/PERRY, S.14)

von Äußerungs-Situationen als auch Zusammenhänge beschriebener Situationen konstituieren. Die aufgrund besonderer Gebrauchsregularitäten einer Sprache dabei ausgebildeten Strukturen sind—obwohl systematischer Natur—im Unterschied zu Naturgesetzen veränderlich und unterliegen—obwohl als bloße Konventionen der Verwendung etabliert—dennoch bestimmten Restriktionen.

Mit Hilfe dieser Invarianten lassen sich nun Situationen als abstrakte Typen (*abstract situations*) bzw. deren Aktualisierungen (*actual situations*) definieren, die mit tatsächlichen Situationen (*real situations*) nur in (den auch belegten) Teilbereichen übereinzustimmen brauchen. Als *Ereignisfolge* kann jede (partielle) Funktion von *Raum-Zeit-Punkten* auf *Situations-Typen* gelten, wobei eine *Ereignisfolge*, die nur für einen Punkt definiert ist, *Sachverhalt* heißt.

Durch Ausnutzung der im Kontext *abstrakter, aktueller* und *realer* Situationen enthaltenen Invarianten kann dabei der Übergang—trotz einer systematischen Unterbestimmtheit von Bedeutung und Interpretation—vollzogen werden: von der sprachlichen Bedeutung (*linguistic meaning*) eines Ausdrucks über die Interpretation seiner Gebrauchsweise (*context of use*) bis zu dem, was seine tatsächliche Äußerung (*on a particular occasion*) an Information vermittelt. Die Betonung des informatorischen Gehalts von Äußerungen bei BARWISE/PERRY führt dabei auf die direkte Entsprechung von externer und mentaler Bedeutsamkeit von Ausdrücken und deren wechselseitiger Interpretierbarkeit.

Die Konzeption des adaptiven (natürlichen oder künstlichen) Systems, das sich in bezug auf seine Umwelt *intelligent* verhält aufgrund und vermöge seiner (veränderbaren) Anpassung an die Invarianten, welche seine Umgebung in *Situationen* strukturieren, weil es selber diese Strukturen (als Kenntnis von Invarianten) besitzt, scheint im Begriff der relationalen Bedeutung deren *externe* und *interne* Modellierungen miteinander vermitteln zu können: die interne Struktur eines informationsverarbeitenden Systems bestimmt nicht nur, welche externen Strukturen seiner Umgebung es zu verarbeiten vermag, sondern System und Systemumgebung erscheinen Konstitutions-analytisch als ununterscheidbar.

Gleichwohl bleibt die Situationssemantik noch auf die Einheit *Satz* im Unterschied zum *Text* beschränkt, sie ist bisher ausschließlich *extern-semantisch* ausgearbeitet, und es fehlen ihr bisher noch alle Hinweise auf eine mögliche empirisch-simulative Rekonstruktion der Basisinheit *Situation*, welche die formale ergänzen könnte.

3 Verstehen und Informationstechnologie

3.1 Im Zuge einer allgemeinen Informatisierung unserer Gesellschaft und unter dem Druck wie den Verheißungen tiefgreifender kommunikations-technologischer Veränderungen werden derzeit weltweit große Anstrengungen unternommen und beträchtliche Mittel in die Forschung investiert zur Untersuchung und Analyse jener Strukturen und Prozesse, aufgrund deren die Menschen in der Lage sind, sprachlich-semiotische Systeme zu entwickeln, zu erlernen und zu verwenden, um miteinander zu kommunizieren. Das Vermögen der Menschen, sprachliche Texte produzieren und rezipieren zu können, die in ihnen vermittelten Bedeutungen zu verstehen und aus ihnen Kenntnisse zu erwerben, die sie als Wissen (ähnlich dem aus eigenen, praktischen Erfahrungen aufgebauten Erkenntnissen) verwenden, ist eine der wichtigsten Voraussetzungen bei der gemeinsamen Lösung neuer Probleme. Dieses Vermögen, nicht ausschließlich nur aus real-weltlichen Gegebenheiten und Erfahrungen *lernen* zu können, sondern auch aus den *semiotischen* Repräsentationen solcher (oder fiktiver) Gegebenheiten und Erfahrungen oder gar aus solchen *Lernresultaten* selbst, erweist sich mittlerweile jedoch als eher unzureichend ausgebildet. Intensivere Einweisung und verstärktes Training einerseits, technische und technologische Unterstützung andererseits scheinen hier dringend erforderlich angesichts sowohl der heute schon vorliegenden und in Zukunft eher noch zunehmenden Massen sprachlicher Texte,

als auch im Hinblick auf die derzeit bestehenden und zukünftig noch entstehenden Probleme sowie deren bisher kaum sichtbaren Lösungsmöglichkeiten.

Es muß betont werden, daß es bei diesem Defizit nicht um quantitativen *Mangel* von jener Art geht, welcher durch umfassendere Bereitstellung zusätzlicher Daten und Informationen sich beheben, und dessen Ausgleich daher bessere Problemlösungen automatisch finden oder anstehende Entscheidungen mechanisch und schneller treffen ließe; vielmehr gilt es zu erkennen, daß—wie bei allen industriellen Produktionsformen am Markt—auch bei den Produktions- und Verwertungszyklen von Daten und Informationen nicht nur ein *Überfluß* des Angebots entsteht, sondern auch massenhaft *Abfall*. Die technisch-informatische Ununterscheidbarkeit von verwertbarem Produkt und unverwertbarem Ausschuß potenziert dabei das Problem angesichts eines über alle Maßen anwachsenden Angebots an und der ständigen (elektronisch gewährleisteten) Verfügbarkeit von *Daten* und *Information*, das nach Kanalisierung verlangt und erst durch eine kognitiv-semiotische Unterscheidung von *relevanten* und *irrelevanten* Zusammenhängen verfügbares *Wissen* bereitstellt.

3.2 Eine sachgerechte Auswahl von jeweils wichtigem Material aus einem Gegenstandsreich im Hinblick auf bestimmte Zusammenhänge oder unter dem Aspekt bestimmter Interessen ist mittlerweile kaum mehr möglich. Sie überfordert in vielfacher Hinsicht den einzelnen menschlichen Problemlöser wie meistens auch ganze Expertenteams. Bei der Tragweite und den oft unabsehbaren Risiken nicht-optimaler⁷ Problemlösungen ergibt sich daher die Notwendigkeit, die spezifische Weise, in der das natürliche kognitive System *Mensch* das ihm verfügbare *Wissen* zu Problemlösungen nutzt, auch in künstlichen intelligenten Systemen nachzubilden. Dies erscheint umso dringlicher angesichts des Unvermögens sogenannter *intelligenter, wissensbasierter, sprachverstehender* Systeme der bisherigen kommunikationstechnologischen KI-Forschung⁸. Deren bisher eher behauptete als schon realisierte *kognitive* Fähigkeiten beruhen auf Funktionen, die *Problemlösungen* im wesentlichen als (algorithmische oder heuristische) Suche und Auswahl optimaler Wege durch die Menge antizipierter Lösungsmöglichkeiten im so definierten *Problemraum* modellieren. Derartige Systeme—oder ihre Weiterentwicklungen—werden aber kaum je in den Stand gesetzt werden können, das als unzulänglich erfahrene Vermögen des Menschen zu verbessern, sich in einer Umgebung zu orientieren, die ihm als *seine Welt* zu *verstehen* ja gerade deswegen so schwerfällt, weil er Menge und Art der sich ihm bietenden Möglichkeiten und ihm abverlangten Entscheidungen nicht entfernt mehr zu durchschauen, d.h. zu antizipieren vermag.

Das kognitive Unvermögen in ähnlich effizienter Weise auszugleichen, wie—auf physiologischer Ebene—eine die Linsenbrechung der Augen korrigierende Brille das defiziente Sehvermögen ihres Trägers seinen Erfordernissen anzupassen vermag, läßt inzwischen nach einer durchaus neuartigen funktionellen Operabilität der kommunikationstechnologischen Systeme und deren Handhabbarkeit suchen. Diese müßte—der Brille und ihrer wie selbstverständlich genutzten Leistung vergleichbar—die Benutzer quasi vergessen lassen, daß sie sich überhaupt eines Hilfsmittels bedienen.

3.3 Eine solche Auffassung von *Operabilität* hätte freilich ernst zu machen mit einem auch phänomenologischen Verständnis von Systemaktivität⁹, das weder die Repräsentation einer

⁷*nicht-optimal* werden hier Problemlösungen genannt, welche Handlungsweisen zur Folge haben, die die zum Zeitpunkt der Entscheidung für oder gegen eine oder mehrere Alternativen solchen Handelns verfügbaren Informationen und Kenntnisse über mögliche Konsequenzen nur deswegen nicht berücksichtigen, weil sie—wiewohl *vorhanden*—den Entscheidungsträgern nicht zugänglich waren oder aber als irrelevant erschienen.

⁸vgl. hierzu RIEGER (1989a), insbesondere *Kapitel 5* und *6*, S.103–160

⁹Den ersten aus dem Kreis der KI-Forschung selbst hervorgegangenen Ansatz in dieser Richtung machten WINOGRAD und FLORES (1986), die eine aus phänomenologischer Sicht vorgetragene Kritik der bisherigen KI-

System-umgebenden *Wirklichkeit* noch das Wirklichkeits-angemessene System-*Verhalten* in Form von *Programmen* zur Verarbeitung von *Daten* und *Informationen* schon voraussetzen kann. Eine solche *selbstorganisierende* Aktivität wäre vielmehr in Form von *Prozeduren* zu realisieren, deren zeitlicher Ablauf in *Prozessen* gerade jene Veränderungen zu simulieren gestattet, die sowohl das System selbst als auch die ihm zugänglichen Umgebungen in Abhängigkeit von den jeweils erreichten Zuständen modifizieren¹⁰.

Derartig *dynamische* Architekturen¹¹ *offener* Systeme können daher in einem neuen Sinne *wissensbasiert* insofern genannt werden, als sie *Verstehen* nicht mehr mit dem Prozeß des Interpretierens von Gegebenheiten oder Zusammenhängen aufgrund von vorgegebenem Wissen gleichsetzen, sondern als diejenige Aktivität des Systems selbst begreifen lassen, die vermöge der Zustands-abhängigen Interpretation von Gegebenheiten und Zusammenhängen aufgrund von System-eigenem Wissen eben dieses Wissen kontinuierlich verändert.

Im Modell vermöchte ein solches kognitives System danach aufgrund seiner *Wissensstrukturen* nur jene Elemente in seiner Umgebung zu erkennen, die sich aufgrund seiner eigenen *Wissensorganisation* als potentielle Zusammenhänge erschließen lassen. Als Menge möglicher Zusammenhangsstrukturen bilden sie die *Welt* des Systems, welche sich ihm durch veränderte Interpretationen der dadurch jeweils zugänglichen Aspekte erschließt. Der Prozeß des Erschließens wird dabei als zeitliche Abfolge von *Wissens-Zuständen* faßbar, deren Aufbau und Veränderung das *Verstehen* und *Lernen* des kognitiven Systems übergreifend als seine Weise der *Bedeutungskonstitution* modellieren.

Als kognitive Hilfssysteme des Menschen konzipiert, bestände die Leistung eines solchen informationsverarbeitenden *Tools* als einem *zeughaften* Modellsystem daher nicht mehr nur—wie in den sogenannten *wissensbasierten* Systemen der traditionellen KI-Forschung—in der Anwendung schon vorausgesetzter, jedenfalls unveränderlicher Strategien und Operationen auf vorgegebene Daten und Strukturen, sondern vielmehr darin, daß sein *kognitives* Vermögen—dem lebender Systeme vergleichbar—es innerhalb seiner strukturellen Beschränkungen und den durch seine jeweiligen Umgebungen vorgegebenen Restriktionen *lernfähig* macht, was sich in der kontinuierlichen Veränderung seines Wissens durch Anwendung dieses Wissens wie auch in seiner damit sich verändernden Brauchbarkeit ausweise.

4 Wortbedeutung und Weltwissen

4.1 Technologisch war der Übergang von der Daten- zur Informationsverarbeitung gekennzeichnet durch die elektronische Speicherung und weltweit abrufbare Verfügbarkeit von (fachspezifischen) Teilbereichen des vorhandenen Wissens in Daten- und Informationsbanken. Der weiterführende Schritt von der Informations- zur Wissensverarbeitung, der als zukunftsweisende Aufgabenstellung und Herausforderung erkannt ist, wird im Rahmen der auf die Verarbei-

Forschung und ihrer Systementwicklungen zur Verarbeitung natürlicher Sprache (*natural language processing*) lieferten. Die Verwirklichung ihrer Forderung aber, sprachverstehende *Tools*—durchaus im HEIDEGGERSchen Sinne—als kognitive System-*Zeuge* zu entwickeln, würde es notwendig machen, gerade auch die *semiotischen* Bedingungen solcher Systementwürfe phänomenologisch zu reflektieren, was bei WINOGRAD/FLORES jedoch unterbleibt (vgl. RIEGER 1989b).

¹⁰MATURANA und VARELA nennen solche Systeme *autopoietisch* und bestimmen sie als "organized (defined as a unity) as a network of processes of production, transformation, and destruction of components that produces the components which: (i) through their interactions and transformations regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it as a concrete unity in the space in which they exist by specifying the topological domain of its realization as such a network." MATURANA/VARELA (1980), S. 135

¹¹Hierzu zählen zum Teil die *konnektionistischen* Modellbildungen *neuronaler Netzwerke*, die im Rahmen *massiv paralleler* Informationsverarbeitung zu neuen (verteilten) Formen der Prozessoren-*Anordnung* in Rechnersystemen (vgl. KEMKE (1988)) wie der Informations-*Repräsentation* (vgl. RIEGER (1991)) geführt haben.

tung der natürlichen Sprache gerichteten Systementwicklungen von KI-Forschung und Computerlinguistik gleichermaßen angestrebt. Dabei stehen allerdings noch Forschungsbereiche im Vordergrund, die eher durch Begriffe wie *Grammatikformalismen* und *Parsingalgorithmen* für natürliche Sprachen, *maschinelle Übersetzung*, *Wissens- und Bedeutungsrepräsentation*, *Verstehenssimulation* abgedeckt werden als durch neue Schlüsselwörter wie *Selbstorganisation*, *Lernfähigkeit*, *Fehlertoleranz* und *Dynamik*, die den Forderungskatalog beherrschen, durch den wünschenswerte Systemeigenschaften in der Anwenderdiskussion umrissen werden. Sie richtet sich verstärkt auf *Expertensysteme*, *integrierte Arbeitsplätze*, *Kommunikations- und Denkwerkzeuge*, die keine bloßen Schlagworte mehr sind: vielfach bezeichnen sie schon Produktentwicklungen, die entweder a de Schwell merkantier Nutzung stehen oder aber tatsächlich umgesetzte, praktische Anwendungsforschungen bezeichnen, zu denen vielfältige, im weitesten Sinne kognitionswissenschaftliche Erkenntnisse beigetragen haben.

Insbesondere ist der Bereich des Wissenserwerbs aus Texten und die Modellierung der hierfür verantwortlichen Sprachverstehens- und Lernprozesse noch weitgehend ungeklärt. Der Umstand, daß es sich bei dieser Forschungsproblematik nicht mehr nur ausschließlich um eine ingenieurwissenschaftliche Umsetzung von theoretisch geklärten und formal expliziten Modellbildungen handelt, sondern um einen mit der sprachlichen Bedeutungskonstitution verbundenen, allgemeinen *semiotischen* Prozeß (nämlich ein Erkennen solcher Zusammenhängen, die durch dieses Erkennen vielfach erst gestiftet werden), macht deutlich, weshalb hier die Computerlinguistik in höherem Maße gefordert ist als die traditionelle KI-Forschung, deren (z.T. negative) Ergebnisse und (z.T. unzureichende) Systeme gleichwohl die anstehende Problematik zumindest zu erkennen halfen.

4.2 Die zunächst theoretischen Systementwicklungen zur Wissens-basierten Verarbeitung von Daten und Informationen in der KI-Forschung, sowie deren praktische Umsetzung in Anwendungszusammenhängen etwa von Expertensystemen, Arbeitsumgebungen, etc. sind bisher im wesentlichen dem Bereich der Ingenieur- und Naturwissenschaften zugute gekommen. Deren weitgehend auf formalisierbarem Wissen beruhende Problemlösungen sowie die dabei verwendeten logisch-deduktiven Verfahren ihrer Durchführung, Überprüfung und Evaluierung legten es nahe, gerade innerhalb dieser Gebiete nach übergreifenden Algorithmisierungen zu suchen. Trotz zahlreicher entweder schon als Software-Produkte auf dem Markt befindlicher oder aber noch in der Entwicklung stehender sog. "intelligenter", weil Wissens-basierter Systeme, kann von einer erfolgreichen Anwendung und Übertragung dieser Ansätze im Bereich der Geisteswissenschaften bisher nicht die Rede sein.

Anders als in Natur- und Ingenieurwissenschaften, deren Erkenntnisinteressen, Forschungsgegenstände und Untersuchungsverfahren in einer von der natürlichen Sprache unterschiedenen formalen Repräsentation höchster Intersubjektivität darstellbar sind, bildet die natürliche Sprache und die Masse der in natürlicher Sprache formulierten Zeugnisse und Dokumente die Basis der Geisteswissenschaften. Deren informatorische Grundlage besteht damit—von den vereinzelt Bildmaterialien einmal abgesehen—im wesentlichen aus sprachlich überliefertem Wissen, das in Form von Textmaterialien mehr oder weniger zugänglich ist. Charakteristisch für den spezifischen Zusammenhang und die Dynamik dieses Wissens ist aber, daß seine unterschiedlichen interpretatorischen Auslegungen sowie ein jeweils (historisch, sozial, edukativ, wissenschaftlich, etc.) sich veränderndes Verständnis dabei wiederum in Texten vermittelt wird, deren unterschiedliche Auslegungen sowie ihr jeweils verändertes Verständnis sich wiederum in Texten ausdrückt, und so fort—was unter anderem die Überlieferung von Geschichte ausmacht und lebendiges Bewußtsein von Geschichtlichkeit begründet.

Den weitgehend in logischen Ausdrücken formalisierbaren Wissensbeständen der exakten Wissenschaften stehen damit die in natürlich-sprachlichen Texten formulierten (oder doch for-

mulierbaren) Verstehenszusammenhänge der Geisteswissenschaften gegenüber, wobei letztere—nicht zuletzt durch das Medium der natürlichen Sprache—ihrem *rationalistischen* Mangel an methodischer Strenge und formalem Rigorismus die Universalität in Richtung und Skopus ihres *hermeneutischen* Verstehens- und Erklärungsanspruchs entgegenstellen kann. In dieser Offenheit und Flexibilität liegt begründet, daß es bisher noch keine—den Algorithmen der logisch-deduktiven Verarbeitungsprozesse vergleichbaren—Algorithmisierungen jener hermeneutischen Prozesse zu geben scheint, die mit *unscharfem Wissen* und *vagen Bedeutungen* in *analoger Weise* umzugehen vermögen, wie es das Lernverhalten kognitiver, informationsverarbeitender Systeme in solchen Umgebungen nahelegt, die diese als *Umwelten* zu interpretieren und als ihre *Welten* zu *verstehen* vermögen.

4.3 Die Repräsentation von *Wissen*, das Verstehen von *Bedeutung* und die Analyse von *Texten* sind durch die Disziplinen der kognitiven Wissenschaften zu einer zentralen Problemstellung der natürlichsprachlichen Semantik geworden, die im Hinblick auf eine Modellierung im Computer sowohl deskriptive und explikative als auch simulative Aspekte vereinigen¹². Dabei ist die systematische Verbindung von adäquater Analyse, Repräsentation und Verarbeitung auch nicht-präzise definierbarer Bedeutungen, welche durch natürlichsprachliche Zeichenaggregate vermittelt werden, zu einer Art Schlüsselproblem avanciert.

Für die meisten Anwendungszusammenhänge ist die *Linearität* der textuellen Darbietungsweise dessen, was wir als *Bedeutung* oder *Inhalt* von sprachlichen Ausdrücken zu bezeichnen uns angewöhnt haben, ein nicht nur ausreichendes Darstellungsprinzip, sondern in seiner linguistischen Funktion sogar eine Bedingung semiotischer Struktur- und Systembildung. Ohne den Zwang zur Linearisierung ist etwa die Konstitution *syntagmatischer* und *paradigmatischer* Relationen zwischen sprachlichen Elementen nicht denkbar, die schon F. DE SAUSSURE als eine strukturbildend-funktionelle Beschreibung sehr fundamentaler Ordnungsprinzipien der kommunikativen Verwendung und Aggregation von Zeichen unterschied. Darüber hinaus stellt Linearisierung aber auch ein gewissermaßen erzieherisches Prinzip dar bei der "allmählichen Verfertigung der Gedanken beim Sprechen" und/oder Schreiben zur Überführung vieldimensionaler, konzeptueller Zusammenhänge in Texte, wie ebenso auch umgekehrt beim allmählichen Aufbau und Vollzug gedanklicher und konzeptueller Zusammenhänge aus der Linearität sprachlicher Texte. Diese Prozesse, bei deren Analyse und Rekonstruktion sich das strukturbildende Moment der Linearisierung wie eine Art Leitlinie nutzen läßt zur Modellierung von Bedeutung und Wissen in Texten, können übergreifend—sowohl als *Verstehen* in der Anknüpfung an schon vorhandene wie auch als *Lernen* im Zustandekommen neuer Zusammenhänge—als *Bedeutungskonstitution*¹³. bezeichnet werden, deren (simulative/kreative) Nachbildung im Computer heute die zentrale Herausforderung der einschlägigen Forschungen bildet.

4.4 Insbesondere unter dem Aspekt des Erwerbs von *Wissen* und des Lernens von *Bedeutungen* ist deutlich geworden, daß die Versuche, die Wahrheitsfunktionalität von *Sätzen* als regelgeleitete Komposition der Bedeutungen von *Wörtern* zu rekonstruieren, sich nicht ohne weiteres auf die Aggregation größerer, textueller Einheiten übertragen lassen. Um besser verstehen zu können, wie sich aus der Verbindung von *Sätzen* in *Texten* etwa Strukturen ergeben, die—als sprachlich repräsentierter *Bedeutungszusammenhang*—jenes (aktuelle wie hypothetische) *Wissen* vermitteln, das zu aktualisieren gerade solchen Systemen möglich ist, die Textstrukturen als System-Umgebungen zu erkennen und zu interpretieren, und dabei aus ihnen wie aus *Situationen* zu lernen vermögen, bedarf es möglicherweise eines völlig anderen Ansatzes.

¹²NORVIG (1987)

¹³Der Begriff wird hier in dem schon früher entwickelten Verständnis von RIEGER (1977) verwendet

Denn hierbei geht es nicht mehr nur um eine analytische Beschreibung, sondern um die nachvollziehende Simulation eines *semiotischen* Prozesses, der die situative Verwendung von sprachlichen Zeichen und deren kommunikative Aggregation in *Texten* zum Gegenstand hat. Anders als in einer *Satzsemantik*, in der Äußerungen als von ihren pragmatischen Einbettungen losgelöst nur unter dem kompositorischen Aspekt der beteiligten Wörter betrachtet werden, läßt *Textbedeutung* sich als pragmatische Information nur dynamisch rekonstruieren aufgrund der Regularitäten, die jede performative Zeichenaggregation zum Zweck der Kommunikation befolgt und verändert, und anhand der Strukturen, die diese Regularitäten in sprachlichen Texten erkennen und wiedererkennen lassen. Für beide, Regularitäten und Strukturen, sind dabei die systematischen Korrelationen entscheidend: sie ermöglichen es, einerseits über die *lineare* Verknüpfung beliebiger Zeichenaggregate zu abstrahieren, was die variablen Strukturen *syntagmatischer* Verkettungs-Beschränkungen erkennen und wiedererkennen läßt, und andererseits über die *situative* Verknüpfung beliebiger Zeichenaggregate zu abstrahieren, was die variablen Strukturen *paradigmatischer* Ersetzungs-Beschränkungen zu interpretieren und zu verstehen erlaubt. Das hierzu benötigte strukturelle Wissen kann ein kognitives, informationsverarbeitendes System dabei selbst aufgrund seiner eigenen Struktur und Verarbeitungsorganisation erwerben und verändern, ohne daß es ihm—wie in herkömmlichen Verarbeitungssystemen—als weitgehend statische Wissensbasis vorgegeben werden müßte.

Man kann—was aus dem bisher Gesagten auch schon deutlich geworden sein dürfte—zusammenfassend sagen, daß es sich bei Entwurf und Konstruktion solcher neuartigen intelligenten, Wissens-basierten Systeme, die wir hier *kognitiv* genannt haben, nicht mehr nur um eine ingenieurwissenschaftliche Umsetzung von theoretisch geklärten und formal expliziten Modellbildungen handelt, sondern daß es um die Modellierung eines mit der sprachlichen Bedeutungskonstitution verbundenen, allgemeinen kognitiv-semiotischen Prozesses geht. Da er—quasi sich selbst organisierend—auf das Erkennen solcher (sprachlicher) Zusammenhänge zu zielen scheint, die durch dieses (verstehende) Erkennen vielfach erst gestiftet werden, wird deutlich, weshalb hier die Computerlinguistik in höherem Maß gefordert ist als die traditionelle KI-Forschung, deren (z.T. negativen) Ergebnisse und (z.T. unzureichenden) Systeme gleichwohl die anstehende Problematik zumindest zu erkennen halfen.

5 Hypertext

5.1 Anhand einer der vielversprechenden Systementwicklungen, die—als *Denkwerkzeug* apostrophiert¹⁴—seit Neuestem (etwa 1986) viel diskutiert wird, soll im folgenden versucht werden, wenigstens einige Aufgaben- und Problemstellungen beispielhaft zu illustrieren, die mit einer über die bloß *technologischen* Veränderungen der Textproduktion und -vermittlung hinausgehenden Veränderung der *kommunikativ-medialen* Funktionen wie auch der diese erst ermöglichenden *text-sprachlichen* Strukturen verbunden sind. Auf die inzwischen einsetzende Nutzung vieldimensional strukturierter Informationsräume mit Bild- und Ton-Erweiterungen durch *Video-Clip*-Techniken in Form von *multi-medialen* Hypertexturen kann hier nur hingewiesen werden¹⁵, zumal sie in ihren Auswirkungen bisher kaum abschätzbar sind.

¹⁴J. Conklin, der 1987 den bisher wohl besten, weil umfassenden, sachlich informativen und gut lesbaren Überblick über Entwicklung und derzeit verfügbare Produkte zum *Hypertext*-Konzept vorgelegt hat, schreibt u.a.: "People who think for a living—writers, scientists, artists, designers, etc.—must contend with the fact that the brain can create ideas faster than the hand can write them or the mouth can speak them. There is always a balance between refining the current idea, returning to a previous idea to refine it, and attending to any of the vague "proto-ideas" which are hovering at the edge of consciousness. Hypertext simply offers a sufficiently sophisticated "pencil" to begin to engage the richness, variety, and interrelatedness of creative thought." (S. 40)

¹⁵P.A. GLOOR/ N.A. STREITZ (1990)

Obwohl im Prinzip aus Komponenten aufgebaut, die traditionellerweise den Systementwicklungen der maschinellen Verarbeitung natürlicher Sprache (*NLP*) ohnehin zugerechnet werden, bilden die sog. *Hypertext (HT)*-Systeme eine neue Entwicklungsstufe im Übergang von der Informations- zur Wissensverarbeitung nicht nur sprachlich kodierter Information. Auf älteren Ansätzen und Überlegungen aufbauend¹⁶, die das menschliche Vermögen zur Konzeptualisierung durch maschinelle Hilfen zu verbessern suchten, wird diese neue Kommunikations- und Verstehens-Technologie aufgrund fortgeschrittener Hard- und Software-Lösungen heute zu preiswerten und damit weitverbreiteten Anwendungssystemen führen. Deren Prinzipien sollen hier kurz vorgestellt werden, um (mögliche) Veränderungen ein- und abschätzen zu können, die sich aus der allgemeinen Nutzung dieser neuen (nicht allein auf sprachliche Vermittlung von Information beschränkte) Technologie für den Umgang mit Sprache als dem hier thematisierten, kognitiv-semiotischen Repräsentationsmedium des Menschen, für das Welt und Wissen verändernde Sprachverstehen und für die sprachliche Kommunikation ergeben könnten.

Den meisten modernen *Systemen* der Daten- und Informationsverarbeitung ist gemeinsam, daß sie als *Verzeichnisse* aufgebaut sind, die *Dateien* enthalten, welche aus *Texten* bestehen und ihrerseits aus *Zeichen* zusammengesetzt sind. Diese lineare Ordnung ist für die Mehrzahl der Aufgabenstellungen heutiger Daten- und Informationsverarbeitungssysteme ausreichend. Mit der Verfügbarkeit von zunehmend komfortableren Entwicklungsumgebungen und Arbeitsplätzen (*Workstations*), die zahlreiche z.T. sehr unterschiedlicher Werkzeuge zur Bewältigung einer Vielzahl von Aufgaben etwa der Textbearbeitung, der Informations- und Dokumentationsarbeit, der graphischen Darstellung, der numerischen Kalkulation, der statistischen Auswertung, der formal-logischen Analyse und Überprüfung, etc. integrieren und für einzelne Teilbereiche mit den erforderlichen Daten und Wissensbasen ausstatten, wuchs allerdings das Bedürfnis danach, neben der ein-dimensionalen (linearen) Darbietung von Informationen im Text auch höhere Formen der mehr-dimensionalen Daten- und Informations-Organisation zur Verfügung zu haben.

So stehen heute Mechanismen bereit, die einen direkten Verweis von einem Teil eines Texts auf einen anderen im selben oder in einem (oder mehreren) anderen Texten nicht nur elektronisch herzustellen sondern als Zusammenhang auch derart zu repräsentieren erlauben, daß er jederzeit reproduzierbar ist. Verschiedene Einzelinformationen können auf diese Weise zu kleineren oder größeren Einheiten zusammengefaßt und durch Verweise auf ihre wechselseitigen Komponenten so dargestellt werden, daß unterschiedliche Zuordnungen auch unterschiedlichen Verbindungen entsprechen.

5.2 Dieses im Grunde einfache Schema von *HT*-Systemen hat nun durch die *Fenstertechnik* auf dem Bildschirm und direkte Unterstützung durch *Datenbasen* für eine neue Sicht von Information gesorgt, weil sprachlich vermittelte Bedeutung sich nicht mehr nur als *lineare Verkettung* von Zeichen sondern erstmals als *vieldimensionale Aggregation* von bedeutungstragenden Elementen sehr verschiedener semiotischer Ebenen darstellen, erkennen, und sichtbar machen ließ.

Jedem (benennbaren) Fenster auf dem Bildschirm sind (identifizierbare) Objekte in der Datenbasis zugeordnet, und es können Verbindungen hergestellt (und benannt) werden sowohl zwischen diesen Objekten innerhalb der Datenbasis (durch Zeigerstrukturen: *pointer*) als auch graphisch auf dem Bildschirm (durch kommentierbare Etiketten/Bildsymbole: *icons*). Die *Datenbasis* erscheint so als ein Netzwerk von Verweis-*Kanten* zwischen textuellen (oder graphischen) *Knoten*, das—als gerichteter (Hyper-)Graph selbst eine Art Hyperdokument—in beliebigen Teilausschnitten auf den Bildschirm gebracht werden kann. Da jedem *Bildschirm-*

¹⁶Engelbart (1963)

Fenster ein Knoten in der Datenbasis entspricht, kann dessen Inhalt—durch Öffnen des betreffenden Fensters—auf den Bildschirm gebracht werden, wobei freilich die Zahl der gleichzeitig offenen Knoten/Fenster beschränkt ist.

Maus-gesteuerten Standardfunktionen der Fenstertechnik, die das Öffnen, Verschieben, Verändern der Größen/Ausschnitte, Öffnen und Schließen von Fenstern auf dem Bildschirm umfassen, werden dabei ergänzt durch *HT*-spezifische Funktionen: die mnemotechnisch über Etiketle/Bildsymbole angezeigten Inhalte von Knoten und Kanten werden durch Anklicken als Fenster geöffnet, Schließen eines Fensters bewirkt Übertragung aller im Fenster vorgenommenen Veränderungen in die Datenbasis, Anklicken des Icons per Maus öffnet das Fenster wieder auf dem Bildschirm, etc.

Da Fenster eine beliebige Anzahl von etikettierten *Verbindungen* (Namen von Kanten zu anderen Fenstern) als Verweise enthalten können, die *Zeigern* zu anderen Knoten in der Datenbasis entsprechen, kann der Benutzer von jedem Knoten beliebig viele neue Verbindungen herstellen sowohl zu neuen Knoten (z.B. um Anmerkungen, Kommentare, Zusätze aufzunehmen) oder zu schon bestehenden Knoten (z.B. um eine noch nicht bestehende Verbindung neu zu definieren). Das rekursive Bauprinzip der *HT*-Architektur läßt dabei eine beliebige (nur durch die Rechnerleistung beschränkte) Tiefe vieldimensionaler Beziehungsgeflechte zu, deren spezifische Ausfüllung von der jeweiligen Ordnungsebene einer semiotischen und/oder konzeptuellen Strukturierung abhängt. Das macht *HT*-Systeme zu geeigneten Werkzeugen bei der Maschinen-unterstützten Sammlung, Darstellung, Veränderung und Verfügbarkeit vieldimensionaler konzeptueller Verbindungsstrukturen, deren Zusammenhänge sich sowohl aus den Mikrostrukturen der in den einzelnen Knoten enthaltenen Informationen als auch aufgrund von interpretativen Akten der die Makrostruktur des Hyperdokuments bearbeitenden Benutzer ergeben können.

Damit scheinen *HT*-Systeme sich in der Tat als Kandidaten anzubieten, deren Einsatz als *Denk-Zeuge* den kommunikativen Umgang mit Sprache und mit in sprachlichen Strukturen repräsentierten Inhalten und Bedeutungen revolutionieren könnte. Als ein dem menschlichen Gedächtnis (vermutlich) näheres, weil assoziativ-parallel (nicht deduktiv-seriell) strukturiertes Modell der Wissensverarbeitung befreit es seine Benutzer ganz offenbar vom Zwang der Linearisierung, dem sie bei der Umsetzung ihrer (vermutlich) vieldimensionalen, kognitiv-konzeptuellen Vorstellungen in sprachlich vermittelbare Zusammenhänge bisher unterworfen waren.

6 Das Strukturierungsproblem

6.1 Im Falle eines Korpus von Texten unterschiedlicher Autoren zu einem übergeifenden Thema könnte dieses Korpus das Hyperdokument bilden. Im Hyperdokument erschiene etwa jeder Autorentext als ein Knoten, dessen wechselseitigen Verbindungen (wegen der übergeifenden Thematik) durch ihr gleichartiges Vokabular, identische oder verwandte Inhalte, Sachbezüge, Ideen, aber auch durch direkte wechselseitige Verweise, etc. hergestellt würden. Dem an diesem Thema interessierten Neuling fällt eine Einarbeitung als Benutzer dieser *HT*-Struktur leicht: textmateriell bestehende Zusammenhänge brauchte er nicht mehr herauszufinden, da sie ihm optisch als Graph dargeboten werden; geläufige konzeptuelle Zusammenhänge erkennt er wieder, ihm unbekanntes kann er als solche erkennen und—falls gewünscht—die ihnen zugrunde liegenden Informationen als Inhalte der betreffenden Knoten bzw. Kanten sich ansehen; zugeordnete neue versuchsweise definieren, um schließlich einen Gesamtzusammenhang der bearbeiteten Thematik als "sein" Hyperdokument abzulegen. Eine solche *HT*-Struktur ist wie eine Moment-Aufnahme der gedanklichen, textualen Verstehens-Arbeit eines Bearbeiters über ein Thema, mit einer (für ihn und andere Bearbeiter) wieder abrufbaren Entwicklungsgeschich-

te bis zu diesem Zustand, eine Art Verlaufsmodell seines sprachlich-kondensierten Verstehens-Prozesses, wiederaufnehmbar, nicht abgeschlossen, jederzeit veränderbar in Breite (durch neue Knoten) und Tiefe (durch neue Verbindungen).

6.2 Im Verlaufe der (individuellen und/oder kollektiven Arbeit) an einem Hyperdokument (i.e. einer Datenbasis, einem Entwurf eines Textes, einem konzeptuellen Brainstorming bei der Entwicklungsarbeit, etc.) kann die Struktur der alten und neuen Knoten und Kanten schnell eine Komplexität erreichen, die im multi-dimensionalen Repräsentationsformat des *HT*-Systems angelegt ist.

Wegen dieser vieldimensionalen Repräsentationsstruktur von *HT*, die das Prinzip der linearen Ordnung sprachlicher Texte wesentlich ergänzt, werden aber—über die graphische Darstellung des inhaltlichen Verweissystems (oder Ausschnitten daraus) als Netzwerk von Knoten und Kanten hinaus—solche Ordnungsprinzipien wichtig, die als Grundlage unterschiedlicher Verfahren der Durchmusterung und Suche (*browsing*) von in *HT*-Strukturen organisierter Information dienen können. Conklin¹⁷ nennt drei Such-Prinzipien:

- ▷ entlang bestehender Verbindungen, wodurch die angetroffenen Knoten nacheinander als Fenster geöffnet werden und deren Inhalt sichtbar machen;
- ▷ aufgrund der Netzwerkstruktur (oder Teilen davon), wobei nach bestimmten Zeichenketten, Schlüsselwörtern oder Attribut-Werten gefragt wird;
- ▷ vermittelt einer strukturellen Steuerung, die es aufgrund graphischer Darstellung der jeweiligen Regionen erlaubt, durchs Netzwerk zu navigieren, wobei die Knoten als Icons und die Verbindungen als Kanten dargestellt werden, jeweils mit oder ohne ihre Namen bzw. Etiketten.

Da die Suchstrategie den gesamten Strukturzusammenhang des Hyperdokuments (oder Teile davon) als Graphen abbildet, bietet sie gleichzeitig Aufschluß über Nähe und Ferne von Knoten sowie die Anzahl und das Gewicht ihrer Verbindungen als wichtige Hinweise auf konzeptuelle Zusammenhänge und Nachbarschaften.

6.3 Die Möglichkeit zu struktureller Orientierung und zielgerichteter Bewegung in *HT*-Repräsentationen mittels eines geeigneten *Navigationsinstruments (HND)*¹⁸ bildet dabei eine entscheidende—bisher jedoch erst in Ansätzen realisierte—Voraussetzung der praktischen Nutzung und *Zeug*-haften Handhabbarkeit von *HT*-Systemen als einer neuartiger Wissenstechnologie.

Das leuchtet sofort ein, wenn man sich die Veränderungen vor Augen führt, die mit dem Übergang von der linearen Strukturierung herkömmlicher Textdokumente zur Vieldimensionalität von Hyperdokumenten verbunden ist:

- ▷ Die Unterscheidung, die wir in bezug auf lineare sprachliche Texte noch zwischen *Textautor* und *Textleser* sinnvoll machen können, wird angesichts viel-dimensionaler Hypertexte und ihrer *Benutzer*, die zumindest potentiell immer sowohl *HT-Leser* als auch *HT-Autoren* sind, bedeutungslos.
- ▷ Von jedem Benutzer werden in den unterschiedlichsten Weisen, je nach Erfahrungsstand und herangetragenem Verständnis durch Hinzufügung und Aktivierung von, wie durch Erweiterung und Kommentierung zu einzelnen Schlüsselwörtern, Relationen oder ganzen Konzeptzusammenhängen jene Veränderungen in ein Hyperdokument eingeführt, die—sowohl als erweiternde Ergänzungen wie auch als vereinfachende Zusammenfassungen—nicht mehr dem Zwang zur Linearisierung unterworfen sind und damit schnell die so

¹⁷Conklin (1987), S. 19

¹⁸für *Hypertext Navigation Device*

entstehenden Hyperstrukturen nach Gliederung und Tiefe unüberschaubar werden lassen.

- ▷ Selbst wenn nur ein Benutzer in einem Sachbereich an einem Hyperdokument arbeitet, wird dieses je nach Domaine und Kreativität dieses Autors schon bald ein Maß an Komplexität erreichen, daß es ihm schwer macht, sein *HT*-System ohne *Navigationshilfe* fruchtbar zu nutzen.

Vor diesem Hintergrund versagen Herkömmliche Strukturierungsprinzipien (prädikativer Zuordnungen) ersichtlich deswegen, weil sie als Wertzuweisungen auf propositionaler Basis eine Linearisierung schon voraussetzen, die aufgegeben zu haben gerade die (formale) Stärke jeder Hyperstruktur ist.

7 Modellierung und Modellanwendung

7.1 Anknüpfend an früher entwickelte Modellbildungen¹⁹ bietet die Theorie der unscharfen (*fuzzy*) Mengen in Verbindung mit statistischen Verfahren der quantitativen Analyse großer Textcorpora nicht nur die Möglichkeit, Vagheit natürlich-sprachlicher Wortbedeutungen als Stereotype zu repräsentieren und in ihrem systematischen Zusammenhang topologisch als vieldimensionale Raumstruktur darzustellen²⁰. Vielmehr erlauben die besonderen (formalen und inhaltlichen) Eigenschaften derart modellierter lexikalisch-semantischer Beziehungen sowohl eine eher Ergebnis-orientierte (statische) Interpretation in Form eines Netzwerks, als auch eine eher Prozeß-orientierte (dynamische) Auffassung sich verändernder, variabler Resultate der zugrundeliegenden Prozeduren. Diese können in ihrem systematischen Zusammenhang das *Bedeutungs-konstituierende* Vermögen insofern simulieren, als es möglich erscheint, die *syntagmatischen* und *paradigmatischen* Beschränkungen über formale Abstraktionen in einem geeigneten funktionalen Modell zu (re-)konstruieren und als einen Prozeß der *Emergenz* von mit Zeichenketten verbundenen Bedeutungen zu implementieren.

Die hierzu auf der Wortebene verwendeten textanalytischen Verfahren sind deskriptiv-statistischer Natur und beruhen im wesentlichen auf einer Korrelationsmessung α von Worttypen in einem Korpus *pragmatisch-homogener* Texte, sowie auf einer *verteilten* Repräsentation von Verwendungsregularitäten, deren Unterschiede über ein Distanzmaß δ numerisch präzisiert werden können.

Als Resultat eines daraus ableitbaren zweifachen Abstraktionsschritts (*Tab. 1*) läßt sich die Lage und Position jeden Bedeutungspunkts $z \in SR$ in einer vieldimensionalen metrischen Struktur, dem sog. *semantischen Raum* (*SR*), auch als Funktion aller Unterschiede (δ - oder Distanzwerte) aller Verwendungs-Regularitäten (α - oder Korrelationswerte) von Wörtern $x \in V$ deuten, die sich für jeden Worttyp des untersuchten Vokabulars zu sämtlichen anderen aus den analysierten Texten ergeben (*Abb. 1*). Als Zuordnung von Wörtern zu ihren *Bedeutungspunkten* stellt diese Funktion eine (mögliche) Operationalisierung der *syntagmatischen* und *paradigmatischen* Restriktionen dar, welche Folgen von Wörtern in Texten pragmatisch homogener Korpora erkennen lassen. Dabei läßt ihre numerische Spezifizierung diese Funktion nicht nur *formal* als Komposition der beiden restringierten Relationen $\delta \mid y$ und $\alpha \mid x$ charakterisieren sondern auch *empirisch* rekonstruieren.

Auf diese Weise modelliert der Analysealgorithmus des Systems—durch die Verarbeitung der natürlichsprachlichen Texte—die Fähigkeit, vermöge der dem System jeweils verfügbaren strukturellen Information sowohl seine (textuelle) Umgebung zu erkennen und aufgrund

¹⁹vgl. hierzu den Überblick der von Verf. vorgelegten Modellansätze und Systementwicklungen in RIEGER (1989a)

²⁰vgl. RIEGER (1990)

	$V \times V$	α -Abstraktion	\Downarrow	$C \times C$	δ -Abstraktion	\Downarrow	$S \times S$	
α	$x_1 \quad \dots \quad x_N$			δ	$y_1 \quad \dots \quad y_N$		∂	$z_1 \quad \dots \quad z_N$
x_1	$\alpha_{11} \quad \dots \quad \alpha_{1N}$			y_1	$\delta_{11} \quad \dots \quad \delta_{1N}$		z_1	$\partial_{11} \quad \dots \quad \partial_{1N}$
\vdots	$\vdots \quad \ddots \quad \vdots$	$\xrightarrow{\alpha x_i}$		\vdots	$\vdots \quad \ddots \quad \vdots$	$\xrightarrow{\delta y_j}$	\vdots	$\vdots \quad \ddots \quad \vdots$
x_N	$\alpha_{N1} \quad \dots \quad \alpha_{NN}$			y_N	$\delta_{N1} \quad \dots \quad \delta_{NN}$		y_N	$\partial_{N1} \quad \dots \quad \partial_{NN}$
		\Uparrow		\Uparrow				
		<i>Syntagmatische</i>		<i>Paradigmatische</i>				
		<i>R e s t r i k t i o n e n</i>		<i>R e s t r i k t i o n e n</i>				

Tabelle 1: Formalisierung (*syntagmatischer/paradigmatischer*) Restriktionen (*constraints*) durch eine zweistufige, konsekutive (α - und δ -) Abstraktion von Verwendungs-Regularitäten x_i und deren Ähnlichkeiten/Unterschieden y_j .

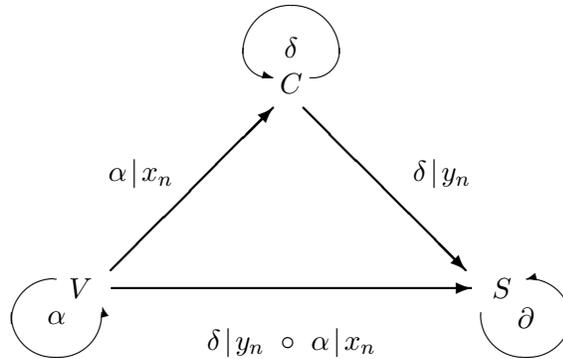


Abbildung 1: Abbildungsrelationen α und δ zwischen den strukturierten Mengen des Vokabulars $x_n \in V$, seiner Verwendungs-Regularitäten $y_n \in C$ und seiner Bedeutungspunkte $z_n \in S$.

seines (semantischen) Wissen zu interpretieren, als auch dieses eigene (semantische) Wissen zu repräsentieren und aufgrund veränderter (textueller) Umgebung zu modifizieren. Beides zusammen—obwohl bisher auf eine vor-prädikative, nicht-propositionale, assoziative Bedeutungsebene beschränkt—macht das *Verstehen* eines solchen kognitiven, informationsverarbeitenden Systems aus.

7.2 Damit läßt sich in bezug auf *selbstorganisierende* Systeme eine Frage beantworten, die für herkömmliche Modelle zur Wissens- und Bedeutungsrepräsentation auch nur zu stellen bisher gar nicht sinnvoll schien: woher nämlich die natürlichsprachlichen Kennzeichnungen (*labels*) kommen, welche die abstrakten Entitäten (*Knoten, Kanten*) in den graphentheoretischen Modellen allererst zu Darstellungen der durch sie repräsentierten Bedeutungen oder ihrer Denotate (*Objekte, Relationen, Sachverhalte, etc.*) machen. Während diese Identifikation von Modellelement und seiner Interpretation auf der vom Modellbenutzer schon verstandenen Bedeutung des natürlichsprachlichen Ausdrucks beruht, die damit vom Modell vorausgesetzt nicht aber modelliert wird, leistet das hier vorgestellte Modell einer strukturellen Bedeutungsrepräsentation genau dies: der Prozeß der *Bedeutungskonstitution* wird durch den Textanalyse-

Algorithmus des Systems modelliert, welcher diesem erlaubt, seine interne Repräsentation in Form von sprachlich gekennzeichneten Bedeutungspunkten und ihren Konfigurationen im *semantischen Raum* je nach textueller Umgebung, in der das System arbeitet, aufzubauen, anzuwenden und zu verändern. Das geschieht im wesentlichen durch die hier entwickelte zweistufige, konsekutive Abbildung des Vokabulars auf sich selbst. Sie läßt sich—ohne dabei auf das Wort- und Welt-Wissen (*die semantische Kompetenz*) von Systementwicklern oder Testpersonen zurückgreifen zu müssen—empirisch überprüfen anhand der Verwendungs-Regularitäten von Wörtern, wie sie von wirklichen Sprechern/Schreibern in natürlichsprachlichen Texten zum Zweck tatsächlicher (oder zumindest intendierter) Kommunikation (*kommunikative Performanz*) befolgt und/oder auch verändert werden.

Darauf baut beispielsweise ein Projekt auf, das eine den modernen Technologien entsprechende Antwort zu finden sucht auf den durch eben diese Technologien ausgelösten Informationsüberfluß sowie auf die mit industrieller Überflußproduktion verbundenen Probleme der *Entsorgung*. Ein solche sollte—wenn schon nicht auf eine Beseitigung so doch zumindest—auf eine Kanalisierung des anfallenden Informations-*Mülls* oder *Abfalls* hinauslaufen und zwar mit dem Ziel, durch Kanalisierung das Verhältnis von Informations-Nachfragern und Informations-Anbietern zu spezifizieren. Ein damit sich verengender Informations-Markt könnte bei *sauberer* Produktion auch zu einem Kosten-günstigeren Informations-Angebot führen.

Die Rede ist von der oben schon angesprochenen Flut sprachlicher Produkte, Texte und Neuerscheinungen. Sie wird in den meisten industriellen Anwendungsbereichen, wissenschaftlichen Disziplinen und selbst in einzelnen, sehr spezialisierten Teilgebieten—durch Textverarbeitungssysteme erleichtert und innerhalb Welt-umspannender Informationsnetze verbreitet—von der Menge der Forscher/Entwickler/Anwender weltweit gespeist, gleichzeitig aber kann sie von den einzelnen betroffenen Forschern/Entwicklern/Anwendern kaum mehr bewältigt, d.h. nicht mehr rezipiert werden im Sinne einer umfassenden und intensiven Lektüre.

Ein automatisches System zum orientierenden *diagonalen Lesen* (*skimming*) von Texten²¹, das lernfähig sein *Wissen* verändert je nach *gelesenen* Texten und aufgrund frei wählbarer Stichwörter inhaltlich *relevante* Zusammenhänge dieses Wissens unter wiederum frei wählbaren *Aspekten* dem Benutzer darbietet, kann in dieser Situation helfen, die Menge des in sprachlichen Texten verfügbaren Wissens zu sichten, zu organisieren, und auf jene Zusammenhangsstrukturen einzuschränken, deren sprachliche und/oder graphische Darbietung—unter dem gewünschten Aspekt der betreffenden und das in den Texten vermittelte Wissen charakterisierenden Stichwörter—erkennen läßt, welche der vom System verarbeiteten Texte noch am ehesten eine zeitaufwendige Lektüre mit Aussicht auf Kenntniszuwachs und Verstehensgewinn rechtfertigen.

7.3 Auf der Basis des bisher grundlagentheoretisch erarbeiteten und empirisch überprüften *kognitiv-semiotischen* Modells zur (algorithmischen) Analyse und Repräsentation (vor-prädikativer) Bedeutungen in Texten bieten sich die besonderen (formalen und inhaltlichen) Modelleigenschaften an, die lexikalisch-semantischen Beziehungen als Teil des in natürlichsprachlichen Texten vermittelten, konzeptuellen Wissens nicht mehr auf ein (statisches) Netzwerk vorgegebener Konzeptstrukturen abzubilden, sondern als eine Sammlung von Prozeduren darzustellen, deren Operationen auf einer Menge von strukturierten Basisdaten nicht nur (dynamisch) sich ändernde, variable Verarbeitungsergebnisse liefern, sondern die Menge und Struktur dieser Basisdaten zu verändern vermag.

Anders als in den (propositionalen) Formaten zur (prädikativen) Bedeutungs- und Wissensrepräsentation der bisherigen KI-Forschung werden die skizzierten Prozeduren als *zeitkritische*

²¹Rieger 1988b

Algorithmen unterschiedlicher (z.T. noch zu testender) Aufgabenstellungen und Operationscharakteristiken weiterzuentwickeln sein,

- ▷ welche semantische Beziehungen zwischen Konzepten nicht voraussetzen müssen, sondern diese induktiv aus den Strukturen der analysierten Corpora als Funktion des Gebrauchs von Wörtern in Texten zu berechnen gestatten;
- ▷ welche—durch die Trennung von Basisstruktur und den auf dieser Basis operierenden Prozeduren—es erlauben, semantische Beziehungen zwischen den *stereotypischen* Repräsentationen (*Bedeutungspunkten* im *semantischen Hyperraum*) von deren—je nach Aspekt, Perspektive, Kontext—variablen konzeptuellen Abhängigkeiten untereinander zu unterscheiden;
- ▷ welche—auf der Grundlage dieser konzeptuellen Hierarchien und der sie aktivierenden Prozeduren— *assoziativ-analoges* im Unterschied zu *deduktiv-logischem* Schließen als *semantische Inferenzen* modellieren;
- ▷ welche schließlich—durch ihre teils rekursiven, teils rückbezüglichen Strukturen—die Resultate solcher Verarbeitungsprozesse eben diesen Verarbeitungsprozessen zu unterwerfen vermögen.

Diese Prozeduren bilden den Kern bilden für das *TESKI*-System²² zum "orientierenden Verständnis" von natürlich-sprachlichen Texten eines Sachgebiets, das sich derzeit in der Entwicklung befindet. Es soll den Systembenutzern ermöglichen, sich aufgrund von Anfragen zu frei wählbaren Stichwörtern, Begriffen, etc. über konzeptuelle Zusammenhänge zu informieren, die vom System unter diesen (oder anderen semantisch relevanten) Stichwörtern durch die Verarbeitung der Texte dieses Gegenstandsbereichs (als seinen *semantischen Hyperraum*) aufgebaut und—Aspekt-abhängig nach Kriterien inhaltlicher Zusammengehörigkeit geordnet—in Form dynamischer, dispositioneller Dependenzstrukturen (*DDS*)²³ ausgegeben werden.

Eine sich anpassende Veränderung der Basisstruktur ist durch kontinuierliche Hinzunahme jeweils neuer Primärtexte ebenso möglich, wie durch Rückkopplung (*Updating*), was über die Verarbeitung der Abfrage-Antwort-Dialoge geschehen soll. Für deren Analyse kann das System die gleichen textanalytischen Verfahren einsetzen, wie beim automatischen Aufbau des *semantischen Hyperraums*, was eine System-immanente, Benutzer-abhängige Relevanz-Steuerung des Lernens/Vergessens von in der sprachlichen Umgebung des Systems erkennbaren Konzepten erlaubt. Mit seiner stereotypischen Repräsentation von vagen Bedeutungen lexikalischer Einheiten und den aus diesen durch konzeptuelle Abhängigkeiten aufgebauten *semantischen Dispositionen* werden darüber hinaus Beziehungen zwischen Bedeutungsrepräsentationen generiert, welche die in den vom System verarbeiteten Texten enthaltenen *relevanten* Informationen — je nach semantischer Perspektive und inhaltlichem Aspekt der Benutzer-Anfrage—inhaltlich re-organisiert und variabel darbieten.

Die Eigenschaften der *dispositionellen Dependenzstrukturen (DDS)* legen es nahe, sie auch als das dynamische, inhalts-gesteuerte Organisationsprinzip zu erproben, das als Kern eines leistungsfähigen *Hypertext-Navigation-Device (HND)* gefordert wurde. Erste, gewiß noch vorläufige Resultate aus Versuchen mit Teilimplementationen sind inzwischen so ermutigend, daß die Entwicklungsarbeit—durch industrielle Anwender unterstützt—verstärkt auch im Rahmen jener *konnektionistischen* Neuansätze fortgesetzt werden, die das kognitive Verhalten *natürlicher* sprachverarbeitender Systeme in bezug auf fehlerhafte, unsichere bzw. unscharfe

²²RIEGER (1988b)

²³Eine eingehende Darstellung mit formaler Einführung und zahlreichen empirischen Beispielen solcher *DDS* werden in RIEGER (1989a), *Kapitel 9*, S.237–289 gegeben.

Bedeutungen (und deren Repräsentationen) sowie auf Vorgänge des Erkennens, Identifizierens, Lernens und Veränderns von solchen Bedeutungen (und deren Repräsentationen) anhand *künstlicher* dynamischer, selbst-regulierender und reflexiver Systeme studieren.

8 Die möglichen Folgen

8.1 Unbestritten sind die mit *HT*-Strukturen verbundenen Vorteile, die es erlauben, eine beliebig komplexe Verweisstruktur beliebiger Tiefe nicht nur aufzubauen und als statische Struktur von bestehenden Zusammenhängen darzustellen, sondern als ein veränderbares System wechselseitig aufeinander beziehbarer Konzepte, Inhalte und Bedeutungen in nahezu beliebigen Informationsträgern (Sprache, Graphik, Farbbild, Film, Bildanimation, Ton, Klangsynthese, etc.) praktisch verfügbar zu haben. Was als Fußnoten-Arsenal und Lesarten-Apparat jedem Historiker und Philologen geläufig und als (unvollkommenes) Abbild der (im einzelnen vielfältigeren) Bezüge innerhalb eines Sachgebiets in seiner (linearisierten) Darstellungsweise jedem Benutzer (mehr oder weniger) schmerzlich bewußt ist, wird in der dynamischen und auf Abruf bereitstehenden *HT*-Strukturierung (vieldimensionaler) Zusammenhänge eines Gegenstandsbereichs praktisch überwunden. Diese Aspekte, welche den im Zwang zur Linearisierung seiner Vorstellungen, Ideen und Gedanken Geübten als durchaus *positiv* zu bewertende Möglichkeiten von *HT*-Strukturen erscheinen werden, sollen uns hier jedoch nicht weiter beschäftigen.

Vielmehr soll hingewiesen werden auf einen bisher übersehenen Zusammenhang, der sich im Umgang mit *HT*-Strukturen dann ergibt, wenn deren (absehbare) Einführung als *edukatives* Medium etwa in der Primar- und Sekundarstufen-Ausbildung nicht mehr Ausnahme sondern die *multi-mediale* Regel sein wird. Hier stellen sich zahlreiche Fragen, deren Beantwortung umso schwerer fällt, je tiefgreifender die möglichen Veränderungen sein werden, die sich durch den Welt-erschließenden Umgang mit *HT*-Strukturen im sprachlichen wie kommunikativen Verhalten und Vermögen der Betroffenen ergeben.

8.2 Man stelle sich eine *Kindergeschichte* oder die *Geschichte der Französischen Revolution* vor: als *HT*-Struktur wären sie beide ein vieldimensionales Gebilde. Dessen Aktivierung auf einem Bildschirm bringt dabei nicht eine Geschichte (mit Anfang und Ende) hervor, sondern stellt ein *modulares Wissenssystem* in einer Art *Baukasten* bereit. Als strukturierte Menge der durch vielfältige Beziehungen untereinander verbundenen Informationsblöcke enthält er quasi die *P o t e n z m e n g e* aller möglichen Geschichten, aus denen ein potentieller Benutzer (Leser?) erst durch seine *Mensch-Maus-Manipulationen* eine zeitliche Abfolge als "seine" Geschichte (Autor?) aufbaut²⁴.

Schon heute ist eine Generation vorstellbar, die—anstatt mit einer zweidimensionalen Ebene (Tafel, Papier, etc.) zur Aufnahme linearisierter Symbolketten (d.h. *sprachlicher Texte*) während ihrer Alphabetisierungsphase konfrontiert worden zu sein—mit dem Bildschirm (groß, farbig, flimmerfrei) als primärem Ausbildungsmedium aufwächst. Den heute noch im Leseerlebnis von Texten gemachten (emotionalen, intellektuellen, sozialen) Erfahrungen von *interpretierter Welt*—auch als Stimulans und Motivation zum Erlernen der sie vermittelnden Schriftzeichen—entsprechen bei jener zukünftigen Generation die (emotionalen intellektuellen, sozialen) Sensationen der via Bildschirm erlebbaren *Mensch-Maus-Manipulationen*, über deren *Welt-erschließende* Kraft sich bestenfalls spekulieren läßt.

Es wurde oben wiederholt auf die beliebige Manipulierbarkeit von *HT*-Strukturen in Computern hingewiesen und darauf, daß damit die für alles sprachliche Mitteilen bisher geltende

²⁴Da es sich bei den Informationsblöcken um Einheiten aus sprachlichen bildlichen akustischen, filmischen, etc. Komponenten handeln kann, würde jede Realisation daraus im Prinzip auch eine *multi-mediale* Geschichte sein.

Notwendigkeit, "Gemeintes" in ein (zeitliches und/oder räumliches) Nacheinander *sprachlicher Linearität* transformieren zu müssen, wenn nicht vollends aufgehoben so doch entscheidend gelockert werde. Im Hinblick auf einen allgemeinen *edukativen* und weltweiten *kommunikativen* Einsatz von *multi-medial* angereicherten *HT*-Systemen in nahezu allen Bereichen des privaten und öffentlichen Lebens einer Gesellschaft wäre selbst eine Veränderung der *natürlichen Sprache* als Medium der Mensch-Mensch-Kommunikation in Richtung auf einen durch Hyperstrukturen vermittelten *Code* denkbar, der ohne eine technologische Unterstützung durch die *Welt-strukturierende* Automaten nicht mehr auskommt. So unvorstellbar eine solche Entwicklung heute auch erscheinen mag, so einsichtig scheint das Argument, wonach das grundlegende Ordnungsprinzip sprachlicher Strukturierung in *syntagmatische* und *paradigmatische* Beziehungen von Zeichenaggregaten abgelöst und ersetzt werden könnte durch die *prozessuale* Verfügbarkeit vieldimensionaler Zusammenhangsstrukturen, welche die *HT*-Systeme in optischer, topologischer und manipulativer Hinsicht gewähren.

Wenn die Grundlage unseres wechselseitigen Kommunizierens nicht mehr *natürlich-sprachliche Texte* (oder wenigstens die durch syntagmatische und paradigmatische Regularitäten eingeschränkte Zeichenaggregate) sind, sondern aus *Icons*, *Symbolen* und *Zeichencodes* an Knoten und Kanten komplexer Netzwerke und Graphenstrukturen bestände, was—so ist zu fragen—würde verhindern, daß sich die Bedeutung von beispielsweise "Demokratie" in einer nur am Bildschirm erfahrbaren *Hypertext-Struktur* eines komplexen Relationengeflechts erschöpft, dessen "Wirklichkeit" darin bestünde, an einem Terminal—Maus-gesteuert und durch vielfältig Stereo-Ton- und Video-Clip-Information ergänzt—*explorativ* durchmessen zu werden ?

Die Möglichkeit jedenfalls sei angedeutet, daß die in nahezu beliebiger Tiefe nahezu beliebig manipulierbaren Zusammenhangsstrukturen, die durch *HT*-Systeme auf dem Bildschirm zugänglich werden, nicht nur dessen zweidimensionale Oberfläche vergessen lassen könnten sondern auch die Notwendigkeit der kognitiv-semiotischen Beherrschung (*Verstehen*) dieser Oberfläche. Ihre als Zwang zur *Linearisierung* erfahrenen Bedeutung-konstituierenden Restriktionen, welche durch die Regelmäßigkeiten *syntagmatischer* Verknüpfungen von *paradigmatischen* Selektionen erst (intersubjektive) Interpretation von Zeichenaggregationen ermöglichen, könnten sich als eine Bedingung der Interpretierbarkeit auch jener vieldimensionalen Strukturen erweisen, deren bloße Manipulation auf dem Bildschirm nicht schon ihr *Verstehen* ausmacht, jedenfalls *Bedeutungskonstitution* weder ersetzen noch befördern wird.

Literatur

- BARWISE, J. / PERRY, J. (1983): Situations and Attitudes. Cambridge, MA (MIT Press)
- CONKLIN, J. (1987): "Hypertext: An Introduction and Survey", *Computer* 20,9: 17–41
- ENGELBART, D.C. (1963): "A Conceptual Framework for the Augmentation of Man's Intellect" in: Howerton, P.W. / Weeks, D.C. (Hrsg): Vistas in Information Handling, Vol. I. London (Spartan)
- GLOOR, P.A. / STREITZ, N.A. (Hrsg.)(1990): Hypertext und Hypermedia. Von theoretischen Konzepten zur praktischen Anwendung [Informatik Fachberichte 249]. Berlin/Heidelberg (Springer)
- HEIDEGGER, M. (1927): Sein und Zeit. Tübingen (M. Niemeyer) ¹⁰1963
- KAEHR, R. / VON GOLDAMMER, E. (1988): "Again Computers and the Brain" *Journal of Molecular Electronics* 4,12: 232–238
- KEMKE, C. (1988): "Der Neuere Konnektionismus. Ein Überblick", *Informatik-Spektrum* 11: 143–162
- MATURANA, H.R. / VARELA, F.J. (1980) Autopoiesis and Cognition. The Reality of the Living. Dordrecht/Boston/London (Reidel Publishing)

- NORVIG, P. (1987): Unified Theory of Inference for Text Understanding. (EECS-Report UCB/CSD 87/339) University of California, Berkeley
- RIEGER, B. (1977): "Bedeutungskonstitution. Einige Bemerkungen zur semiotischen Problematik eines linguistischen Problems", *Zeitschrift für Literaturwissenschaft und Linguistik* 27/28: 55–68
- RIEGER, B.(1988b): TESKI – A natural language TExt-SKImmer for shallow understanding and conceptual structuring of textually conveyed knowledge. (LDV/CL-Report 10/88) FB II: Linguistische Datenverarbeitung, Universität Trier
- RIEGER, B.(1989a): Unschärfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten. Bern / Frankfurt / New York / Paris (P. Lang)
- RIEGER, B.(1989b): "Computerlinguistik und Verstehenstechnologie. Zur Abschätzung ihrer Aufgaben und möglichen Folgen" in: Gatzemeier, M. (Hrsg.): Verantwortung in Wissenschaft und Technik. Mannheim/Wien/Zürich (B.I. Wissenschaftsverlag), S.256–276
- RIEGER, B.(1990): "Unschärfe Semantik: numerische Modellierung von Wortbedeutungen als *Fuzzy-Mengen*" in: Friemel, H.-J. / Müller-Schönberger, G. / Schütt, A. (Hrsg.): Forum '90 – Wissenschaft und Technik. Neue Anwendungen mit Hilfe aktueller Computer-Technologien. [Informatik Fachberichte 259] Berlin / Heidelberg / New York (Springer), S.80–104
- RIEGER, B.(1991): "On Distributed Representation in Word Semantics", International Computer Science Institute, [ICSI-Report TR-91-012], University of California, Berkeley
- RIEGER, B. / THIOPOULOS, C. (1989): "Situations, Topoi, and Dispositions. On the phenomenological modelling of meaning", in: Retti, J. / Leidlmaier, K. (Hrsg.): Proceedings: 5. Österreichische Artificial-Intelligence-Tagung, [Informatik-Fachberichte 208] Berlin / Heidelberg / New York (Springer), S.365–375
- RIEGER, B. / THIOPOULOS, C. (1992): "Process Dynamics and Semiotics: A self-organizing lexical system in hypertext" in: Rieger, B./ Köhler, R. (Hrsg.): QUALICO-91 – Proceedings of the First International Quantitative Linguistics Conference, Trier, Germany Amsterdam (Elsevier Science) [im Erscheinen]
- SAUSSURE, F. DE (1916): Cours de Linguistique Générale. Kritische Edition von R. Engler, Wiesbaden (Backhaus) 1967
- WINOGRAD, T. / FLORES, F. (1986): Understanding Computers and Cognition: A New Foundation for Design. Norwood, NJ (Ablex Publishing)