

Computerlinguistische Ansätze zur Repräsentation und Verarbeitung von Wortbedeutungen

Eine Theorie der lexikalischen Semantik muß den Einfluß kontextueller und konzeptueller Faktoren auf die Interpretation der Lexeme mit berücksichtigen, um darlegen zu können, wie und warum die Bedeutung einzelner lexikalischer Einheiten variiert.¹ Dies zeigt sich auch in den Modellen der wortsemantisch orientierten sprachverarbeitenden Künstlichen Intelligenz (KI) und Computerlinguistik, die in diesem Beitrag skizziert werden sollen. Anhand dreier in der Literatur vorgeschlagenen Methoden zur Repräsentation und Verarbeitung von Wortbedeutungen soll dargestellt werden, welche Probleme in der wortsemantisch orientierten maschinellen Sprachverarbeitung schon behandelt werden können, aber auch, wo die Grenzen zur Zeit noch liegen. Zuerst sollen die Probleme erläutert werden, die bei der Analyse von Wortbedeutungen auftreten. Die Kontextabhängigkeit der Wortbedeutung wird sich als derjenige Faktor herausstellen, der das Hauptproblem bei der Verarbeitung darstellt. Danach wird auf den Begriff der Wissensrepräsentation, wie er in der Künstlichen Intelligenz gebraucht wird, kurz eingegangen und darauf aufbauend eine spezielle Klasse von semantischen Netzen dargestellt, nämlich KL-ONE-artige Repräsentationen. Die Verarbeitungsstrategien für Wortbedeutungen, die dann diskutiert werden sollen, sind Marker Passing, gewichtete Abduktion und schließlich die Angabe von „Qualiastrukturen“. Die zentrale Frage für diesen Beitrag ist: Wieweit kann man mit diesen genannten Methoden Bedeutungsunterschiede in der Wortsemantik erfassen? Ziel dieses Beitrags soll sein, einen Eindruck von der Komplexität der Wortsemantik zu vermitteln und zudem zu zeigen, daß die lexikalische Semantik im Grunde genommen ein Feld ist, in dem noch sehr viele Probleme bestehen, die für eine elaborierte Theorie der Sprachverarbeitung gelöst werden müssen.

1. Probleme in der lexikalischen Semantik

Wenn man ein Computersystem entwickeln will, das in der Lage sein soll, natürlichsprachliche Kommandos oder Anfragen zu verstehen und

¹ Dies gilt natürlich auch für die Satz- und Diskurssemantik, nur die lexikalische Semantik muß sich per se mit konzeptuellen Strukturen als der Extension dieser Einheiten und dem kontextuellen Faktor bei der Festlegung der aktuellen Bedeutung befassen und kann hierbei nicht, soweit bekannt, auf kompositionale Mechanismen zurückgreifen.

nicht nur Befehle, die in einem dem Computer angepaßten Code angegeben werden, müssen aufbauend auf morphologischen und syntaktischen Analysen Verfahren zur Bestimmung der Bedeutung der einzelnen Sätze und des Textes bereitgestellt werden.

Die Festlegung der Bedeutung muß bei der kleinsten Einheit des Satzes oder Textes anfangen, der lexikalischen Einheit. Von der Bedeutung einzelner lexikalischer Einheiten ausgehend werden komplexere Bedeutungen berechnet, bis die Satz- und schließlich die Textbedeutung bestimmt worden ist. Das der Satzbedeutung zugrundeliegende Prinzip ist die Kompositionalität: die Bedeutung eines Satzes ergibt sich aus den Bedeutungen seiner Elemente und der Art ihrer Zusammenfügung zu komplexeren Einheiten. Dieses Prinzip geht auf Frege zurück und bestimmt die Semantiktheorien der logisch orientierten Semantiker. Kompositionalität ist aber nicht das alleinige Prinzip der Semantik, wie schon daran zu sehen ist, daß sie in der Regel die Wortbedeutung als unanalysierte kleinste Einheit, aus der die komplexeren Repräsentationen aufgebaut werden, betrachtet.

Während man zum jetzigen Zeitpunkt recht gut weiß, wie die Satzbedeutung kompositional aufgebaut wird, gibt es noch keine einheitliche Annahme darüber, wie die nicht-kompositionalen Aspekte der Bedeutung zustandekommen. Nicht-kompositionale Bedeutungsaspekte betreffen alle Ebenen; sowohl die Wortsemantik als auch komplexere Einheiten wie Phrasen, Sätze und insbesondere ganze Diskurse. Oberhalb der Wortsemantik zeigt sie sich u.a. in Form von Skopusambiguität, anaphorischer Koreferenz und präferierten Interpretationen bei Ambiguität ganz allgemein.

Auf die Wortsemantik bezogen bedeutet Nichtkompositionalität: die Bedeutungen der einzelnen Lexeme werden durch die Art ihrer Kombination mit anderen Lexemen und zusätzlich durch Inferenzen, die durch die Bedeutungen anderer Lexeme angestoßen werden, festgelegt. Dies bedeutet, daß Wörter in verschiedenen Kontexten verschiedene Bedeutungen erhalten können, und diese kontextspezifischen Bedeutungen müssen von einem Computersystem erkannt und berechnet werden.

Parallel zu den Ansätzen aus der kognitiven Linguistik und Psycholinguistik wird angenommen, daß zwischen dem sprachlichen Wissen und der Welt eine interne Repräsentation als Weltmodell vermittelt, die sogenannte „projizierte Welt“ (Habel 1986). Diese projizierte Welt wird als Wissensbasis aufgebaut. Die Wortbedeutung ist dann die Menge der Konzepte in der Wissensbasis, auf die die lexikalische Einheit referieren kann. Die Äußerungsbedeutung hingegen ist dasjenige Konzept, auf das

die lexikalische Einheit in einem bestimmten Kontext referiert. Variationen von Wortbedeutungen kommen auf verschiedene Art daher und bilden innerhalb syntaktischer Kategorien bestimmte Muster, die für eine Klasse lexikalischer Einheiten gelten, wie die folgenden Beispiele aus den Bereichen Polysemie, Metonymie und Unterbestimmtheit zeigen.

Nomen wie *Kaffee*, *Museum*, *Tür* und viele andere sind (z.T. hochgradig) polysem. Sie können auf eine Menge von Konzepten referieren, die alle miteinander über Relationen verbunden sind. Charakteristisch ist also für solche Nomen, daß sie nicht auf eine Klasse einheitlicher Objekte referieren, sondern auf ein komplexes System von Konzepten, die eine vernetzte Einheit bilden. In bestimmten Kontexten kann das Nomen dann auf einen Teilbereich dieses Netzes referieren. Dies ist eine Operation, die Bierwisch (1983) konzeptuelle Verschiebung genannt hat:

- (1) *Kaffee tropft auf den Boden.* (GETRÄNK)
- (2) *Kaffee ist im Supermarkt sehr teuer.* (PULVER/BOHNEN)
- (3) *Kaffee wächst in Brasilien.* (STRAUCH/FRUCHT))

Während in (1) aufgrund von Selektionsbeschränkungen *Kaffee* auf das Getränkekonzept referieren muß (denn nur Flüssigkeiten können tropfen), ist die wahrscheinlichste Interpretation des Nomens in (2) durch das Wissen bestimmt, daß im Supermarkt Kaffeepulver oder Kaffeebohnen verkauft werden und nicht das Getränk. Dies ist Wissen, daß durch stereotypische Ereignisabläufe mit ihren involvierten Objekten determiniert wird. Wichtig ist dabei, daß es durchaus möglich ist, daß das Nomen auf das Getränkekonzept referiert, aber es ist nicht wahrscheinlich. In (3) schließlich verweist *Kaffee*, wiederum durch Selektionsrestriktionen gesteuert, entweder auf das Kaffeestrauch- oder das Frucht-konzept. Das konzeptuelle Wissen über Kaffee wird also mindestens durch die Konzepte „Kaffeebohnen“, „Kaffeestrauch“, „Kaffeepulver“ und „Kaffeegetränk“ strukturiert, die durch Wissen über bestimmte Ereignisabläufe miteinander verbunden sind. Die Semantik von *Kaffee* ist aber noch komplexer. *Kaffee* gilt als Massennomen, und als solches unterliegt es der Bedingung der Nicht-Zählbarkeit. Es ist aber möglich, Einheiten der verschiedenen Massen zu messen, wobei die Maßangabe mittels Selektionsrestriktionen wiederum die Art der Masse bestimmen kann:

- (4) *Drei Tassen Kaffee.* (GETRÄNK)
- (5) *Fünf Pfund Kaffee.* (PULVER/BOHNEN)

Unter (4) kann *Kaffee* sowohl auf das Getränkekonzept als auch auf das Kaffeebohnen- oder Pulverkonzept referieren, obwohl auch hier die wahrscheinlichste Bedeutung durch das Getränkekonzept geliefert wird. In (5)

jedoch schließt die Maßangabe das Getränk aus, da Flüssigkeiten üblicherweise nicht in Pfund gewogen werden. Es existiert also generell die Möglichkeit, mithilfe von Maßangaben die Bedeutung vom Massenkonzzept zu Massenquantia zu verschieben. Die inverse Richtung vom individuierten Teil zum Massekonzzept ist ebenfalls möglich, wenn auch eingeschränkter: In dem Satz *Auf dem Teller liegt Lamm* verweist das Nomen *Lamm* auf eine bestimmte Masse.²

Individuativa wie *Tür*, *Buch* und *Museum* haben ein polysemes Verhalten, das in den folgenden Sätzen dargestellt ist:

- (6) *Peter geht durch die Tür/Peter stand in der Tür.* (DURCHLASS)
- (7) *Peter streicht die Tür rot.* (OBJEKT)
- (8) *Peter kauft ein Buch.* (OBJEKT)
- (9) *Peter findet das Buch spannend.* (INFORMATION)
- (10) *Das Museum befindet sich an der Hauptstraße.* (GEBÄUDE)
- (11) *Das Museum beschließt, die Ausstellung zu verlängern.* (INSTITUTION)
- (12) *Das Museum streikt für eine Lohnerhöhung.* (GRUPPE VON MENSCHEN)
- (13) *Der Besuch des Museums war für Peter eine neue Erfahrung.* (AUSSTELLUNGSPROZESSE)

Charakteristisch für diese Klasse von Nomen ist, daß ihre verschiedenen Interpretationen in einer systematischen Beziehung zueinander stehen, die durch bestimmte konzeptbasierte Relationen charakterisierbar sind. Damit unterscheiden sie sich klar von ambigen Wörtern wie *Fliege* oder *Ball*. Ein anderes Problem für die Sprachverarbeitung bereiten metonymische Ausdrücke, die sich durch Übergänge eines Begriffs zu einem anderen mittels bestimmter Relationen wie z.B. Ursache-Wirkung oder Teil-Ganzes auszeichnen. Während also die Wortbedeutung eines polyreferierenden Nomens eine Menge von Konzepten ist, das als Potential für Äußerungsbedeutungen fungiert, ist Metonymie hingegen ein Fall für regelbasierte Uminterpretationen. So ist aus

- (14) *Die Firma hat angerufen.*

zu inferieren, daß ein Mitarbeiter der Firma (Der Gruppe von Mitarbeitern, die die Institution bilden) angerufen hat, nicht das Firmengebäude

² Eine ausführliche Diskussion der Beziehung zwischen Individuativa und Massennomina findet sich in Krifka (1989). Eine weitergehende Darstellung von Bedeutungsvariationsphänomenen, die noch komplexere Eigenschaften von Masse- und Individuumsausdrücken behandelt als in diesem Beitrag angedeutet, ist in Dölling (1992) präsentiert.

oder die Institution selbst. Der Unterschied zwischen Polysemie und Metonymie spiegelt sich in der Organisation der Bedeutungsstruktur der Lexeme wieder: polyseme lexikalische Einheiten referieren auf eine vernetzte Klasse von gleichwertigen Konzepten, die mittels spezifischer Relationen verbunden sind. Metonymische Ausdrücke hingegen besitzen ein festes Bedeutungspotential, das stellvertretend für einen anderen Ausdruck steht und mit diesem in einer bestimmten Beziehung steht. Die klarsten Fälle von Metonymie sind bei Eigennamen auszumachen (*Bonn beschloß Sparmaßnahmen; sie liest van Gulick*), aber (14) zeigt, daß Metonymie in Kombination mit Polysemie auftreten kann: *Firma* ist polysem bzgl. der Gebäude/Institutionslesart und die Institutionslesart wird verwendet, um auf einen Mitarbeiter metonymisch zu verweisen. Eine andere Art von Problemen zeigt sich in der Verbsemantik. Im Gegensatz zur Nominalsemantik, wo – wie gezeigt – ein Nomen auf ontologisch völlig verschiedene Klassen referieren kann, die aber konzeptuell miteinander verbunden sind, können Verben Spielraum in der Bedeutung innerhalb einer Klasse zulassen. Verbkonzepte erlauben dann abhängig von den Konzeptstrukturen der Argumente Spezialisierungen. Dies ist eine Operation, die Bierwisch (1983) konzeptuelle Differenzierung nennt:

(15) *Peter öffnet die Flasche.*

(16) *Peter öffnet die Tür.*

(17) *Peter öffnet seine Jacke.*

In allen drei Fällen wird auf spezielle Aktivitäten referiert, die zwar alle als Öffnen-Ereignisse klassifiziert werden können, aber konzeptuell verschiedene Repräsentationen erhalten müssen. Öffnen lassen sich nur Container im weitesten Sinne, aber die konzeptuelle Repräsentation der verschiedenen Container erlaubt nur bestimmte Arten des Öffnens. Dies zeigt sich an möglichen Fragen, die aufgrund der Information aus den Sätzen (15) bis (17) gestellt werden können: die Frage *Was ist denn da drin?* ist nur bei (15) sinnvoll. Andere Unterbestimmtheitsphänomene in der Wortsemantik zeigen sich bei Präpositionen, bei der Kopula *haben* (*eine Freundin haben* vs. *ein Kind haben* vs. *ein Auto haben*), Adjektiven (*langes Auto* vs. *lange CD*) usw. Es scheint so zu sein, daß es keinen Bereich der Lexik gibt, in dem ein Wort eine kontextunabhängige Bedeutung hat.³

³ Auch Quantoren wie *ein paar*, *die meisten* etc. sind hochgradig kontextabhängig zu interpretieren, wie Hörmann (1983) in einer experimentellen Studie nachweist. Diese Befunde werfen dann die Frage auf, ob die Theorie der generalisierten Quantoren aus der modelltheoretischen Semantik nicht um Kontextabhängigkeit erweitert werden müsste.

Ich werde mich hier aber auf Bereiche der Nominalsemantik, die mit Polysemie und Metonymie zu tun haben, beschränken. Ich möchte Ansätze zur Repräsentation von Wortbedeutungen in sprachverstehenden Systemen diskutieren, die die Kontextabhängigkeit der Bedeutung berücksichtigen und, wichtiger noch, Vorschläge zur Verarbeitung, also zur Berechnung der aktuellen Bedeutung im Kontext. Es muß also sowohl etwas über den Repräsentationsformalismus als auch über die Verarbeitungsstrategien für Wissen über Wortbedeutungen gesagt werden.

2. Repräsentationen von Wortbedeutungen

Mittels eines Repräsentationsformalismus werden Wortbedeutungen formal beschrieben. Die Aufgabe einer Wissensrepräsentation für einen Ausschnitt der Welt besteht darin, ein internes Weltmodell zu konstruieren. Das Weltmodell repräsentiert Wissen, das durch perzeptive und motorischer sowie sprachlich vermittelter Information aufgebaut wird und muß nicht notwendigerweise der Realität entsprechen, wie das Verstehen fiktiver Texte zeigt. Die Struktur der Konzepte ergibt sich somit aus Information über Objekteigenschaften funktionaler und räumlicher Art im weitesten Sinne. Wortbedeutungen sind dann durch das Wissen über die Konzepte, auf die das Wort verweisen kann sowie die interne Struktur der Konzepte bestimmt. Da Wissensrepräsentation eine Rekonstruktion des menschlichen Wissens über die Realität ist, das zu repräsentierende Wissen andererseits sehr komplex ist, kann eine vollständige Repräsentation nicht erreicht werden. Aus diesem Grunde sind Wissensrepräsentationen immer Approximationen an die Strukturen des Weltmodells (Habel 1990). Es sind in der Forschung zur Wissensrepräsentation mehrere Formalismen zur Repräsentation von Wissen vorgeschlagen worden, von denen die semantischen Netze einen Teilbereich bilden. Andere Möglichkeiten zur Wissensrepräsentation werden z.B. von den üblichen Logiken, framebasierten Sprachen oder Produktionsregeln bereitgestellt. Die Diskussion, welche Art von Formalismus geeignet ist zur Repräsentation welcher Art von Wissen muß auf vier verschiedenen Ebenen geführt werden, um die verschiedenen Aspekte, die bei einer Wissensrepräsentationssprache zum Tragen kommen, klar zu trennen.⁴

Die erste Perspektive oder Ebene ist die Implementationsebene. Auf dieser Ebene muß festgelegt werden, wie die Verarbeitung von Ausdrücken

⁴ Die Unterteilung in vier Ebenen ist in Reichgelt (1991) vorgenommen worden. Sie geht auf Brachman (1979) zurück, der allerdings noch eine fünfte Ebene einführt, um die aktuelle Modellierung von der vierten Ebene, der Angabe der zu benutzenden primitiven Bausteine, zu trennen.

der Wissensrepräsentationssprache realisiert werden soll. Dies betrifft Fragen wie etwa die nach der Art der zu benutzenden Datenstrukturen, den Algorithmen zur Implementierung der Inferenzen und Effizienz der Repräsentation. Diese Ebene ist eher für die reine Informatik wichtig und ich gehe auf sie hier nicht weiter ein.

Die zweite Ebene betrifft die logischen Aspekte einer WR-Sprache. Hier muß festgelegt werden, welche Art von Ausdrücken der WR-Sprache was zu bedeuten haben. Ihnen muß also eine Semantik gegeben werden. Die Ausdrucksstärke des Formalismus ist ebenfalls ein Aspekt der logischen Perspektive auf Wissensrepräsentationssprachen: man muß wissen, welche Information in dem Formalismus angegeben werden kann, welche aber nicht mehr. Ein weiterer Aspekt der logischen Perspektive betrifft die Inferenzen, die gezogen werden sollen. Welche Schlüsse möchte man zulassen, welche aber nicht? Reichen deduktive Inferenzen, oder sollen andere (z.B. abduktive) Schlüsse auch möglich sein?

Die dritte Ebene ist die epistemologische Ebene. Hier geht es um die Fragen, welche Arten von konzeptuellen Primitiven angenommen werden sollen und welche Arten von Inferenzen notwendig sind. Es geht auf dieser Ebene noch nicht um konkrete Wissensrepräsentationen, sondern um die Typen, die für eine konkrete Repräsentation notwendig sind. Diese Ebene grenzt sich von der zweiten, der logischen, zum einen dadurch ab, daß die Bausteine, die auf der zweiten Ebene zugelassen sind, hier benutzt werden und zum zweiten, daß entschieden werden muß, wie die auf der zweiten Ebene zugelassenen Schlußarten verwendet werden sollen. Auf die Wortsemantik bezogen heißt dies: man muß wissen, welche Grundbausteine für eine adäquate Repräsentation des Konzeptwissens erforderlich sind und welche Inferenztypen zur Festlegung von Wortbedeutungen im Kontext nötig sind. Gerade die Art der Inferenzen spielt in der Diskussion in der gegenwärtigen Literatur eine große Rolle und ich werde zeigen, daß zwei der drei von mir diskutierten Ansätze auf eben genau der Verwendung bestimmter Inferenztypen beruhen.

Die letzte Ebene schließlich ist die konzeptuelle Ebene. Im Gegensatz zur epistemologischen Ebene muß auf dieser Ebene festgelegt werden, welche konzeptuellen Primitiva und Inferenzen wirklich benutzt werden; es müssen also explizit die Schlüsse und Konzepte angegeben werden. Hier spielt also der Aufbau der Wissenbasis die zentrale Rolle. Wenn man sich also z.B. auf der epistemologischen Ebene entscheidet, daß Wissen am besten in einem semantischen Netz zu repräsentieren ist, daß also Knoten und Kanten als Primitive benutzt werden, dann muß man sich auf der konzeptuellen Ebene entscheiden, was als Knoten und Kanten repräsentiert werden soll und ob bestimmte Kanten als Inferenzen aufzu-

fassen sind. Auf die Wortsemantik bezogen bedeutet dies: die einzelnen Konzepte, auf die ein Wort referieren kann, müssen festgelegt werden (entweder als Knoten oder als Kanten) und die Beziehungen zwischen ihnen ebenfalls.

Die Modellierung des Konzeptwissens beruht hierbei auf der Einsicht, daß nicht wissenschaftliche Theorien, sondern Theorien des alltäglichen Wissens als Konzeptwissen angegeben werden sollen, wenn ein kognitiv adäquates Modell der Konzeptstruktur entworfen werden soll. Es geht also um Alltagstheorien des Menschen über Relationen zwischen Entitäten im allgemeinsten Sinne. Von diesen vier Aspekten ausgehend und sich von ihnen leiten lassend gilt es sich zu entscheiden, welche Art von Repräsentationsformalismus für die Angabe von Wortbedeutungen geeignet ist. Die Art von Repräsentationssprachen, die im folgenden genauer betrachtet werden sollen, sind KL-ONE artige Sprachen, und zwar aus den folgenden Gründen: Die auf dem KL-ONE-Paradigma beruhenden Wissensrepräsentationssprachen – auch Termbeschreibungssprachen oder terminologische Logiken genannt – besitzen eine definierte Syntax mit kompositionaler modelltheoretischer Semantik zum Aufbau von Konzeptstrukturen. Die Semantik für diese Sprachen ist auch die Grundlage bestimmter Inferenzprozesse, die nicht rein operational oder ad-hoc formuliert sind.

Damit erhalten wir eine Wissensrepräsentationssprache, die sich an den Eigenschaften klassischer Logiken orientiert. Die Korrektheit der Repräsentation läßt sich durch die Semantik der Ausdrücke überprüfen. Für die Repräsentation von Wortbedeutungen bedeutet dies, daß alle Bedeutungen, die ein Wort haben kann, in der formalen Sprache, die zum Aufbau des Netzes gegeben ist, formuliert sind und entsprechende mengentheoretische Extensionen besitzen. Die Benutzung einer modelltheoretischen Semantik macht den logischen Aspekt der Wissensrepräsentation aus. Das in einer Wissensbasis angegebene Wissen ist damit also eine logische Rekonstruktion eines Weltausschnitts.

Die Idee, Wissen als vernetzte Repräsentation anzugeben, geht auf Quillian (1968) zurück, ist also schon über 25 Jahre alt. Fast alle Techniken, die später (und auch heute noch) verwendet werden, sind in dieser Dissertation schon vorhanden. Diese grundlegenden Ideen sind:

1. Es werden Knoten eingeführt, die Objektklassen repräsentieren.
2. Die Knoten sind durch Kanten (links) miteinander verbunden, die konzeptuelle Relationen festlegen.
3. Die Konzepte sind in einer Hierarchie angeordnet.
4. Aktivierungsausbreitung wird als Inferenzmechanismus zur assozia-

tiven Verbindung zweier Konzepte angeführt: der Verweis auf ein Konzept führt zur Aktivierung benachbarter Konzepte, die mit dem ursprünglichen Konzept in einer semantischen Verbindung stehen. Zwei Konzepte werden dann durch die Menge der Konzepte, die von beiden erreicht werden können, verglichen. Dabei werden Konzepte nur dann aktiviert, wenn ein bestimmter Schwellenwert überschritten wird.

Aktivierungsausbreitung wird in den meisten heutigen semantischen Netzen nicht mehr als Inferenz verwendet. Einige Ansätze in der Wortsemantik jedoch verwenden die gleiche Idee, die hinter der Aktivierungsausbreitung steht, nämlich die Marker Passing Ansätze, die ich gleich diskutieren werde. Das eigentliche KL-ONE-System wurde ca. 1977 entwickelt und gilt mittlerweile als Stellvertreter einer ganzen Klasse von Wissensrepräsentationssprachen.⁵ Die inferentielle Mächtigkeit von KL-ONE beschränkt sich auf Vererbung von Eigenschaften übergeordneter Konzepte auf speziellere Konzepte. Alle anderen Problemlösungsprozesse liegen außerhalb von KL-ONE und müssen gesondert gehandhabt werden. Mit KL-ONE lassen sich auch Wortbedeutungen repräsentieren. Die Verarbeitungsmechanismen, die notwendig sind, um Wortbedeutungen im Kontext adäquat festzulegen, liegen jedoch außerhalb der Mächtigkeit von KL-ONE, da Vererbung von Eigenschaften nicht zur Festlegung einer Wortbedeutung herangezogen werden kann. Aus diesem Grund soll KL-ONE nur der Aufhänger zur Repräsentation von Wissen über Wortbedeutungen sein. Eine Wissensrepräsentationssprache, die die Festlegung von kontextuell adäquaten Wortbedeutungen ermöglicht, muß neben der Bereitstellung bestimmter Sprachkonstrukte auch die entsprechenden Inferenzen ermöglichen.⁶

KL-ONE ist ein hybrides System, in dem zwischen der Beschreibung terminologischen Wissens über Eigenschaften einzelner Objekte und Klassen und der Beschreibung von Einzelfakten unterschieden wird. Die terminologische Beschreibung wird als TBox, die Beschreibung der Ein-

⁵ Zur Beschreibung von KL-ONE siehe Brachman/Schmolze (1985) oder von Luck/Ownsnicki-Klewe (1990).

⁶ Natürlich kann man sich auch die Frage stellen, ob semantische Netze überhaupt die geeigneten Repräsentationssprachen für Wortbedeutungen sind. Ich glaube, daß zur Verwendung in sprachverstehenden Systemen auch semantische Netze noch nicht die ultimative Repräsentationsform sind und weiterentwickelt werden müssen. Ein Kritikpunkt an netzartigen Repräsentationen ist am Schluß dieses Beitrags angedeutet. Allerdings sind semantische Netze zur Zeit wohl die elaboriertesten Repräsentationsformalismen, so daß ihre Darstellung in diesem Beitrag als Aufhänger gerechtfertigt ist.

zelfakten – die assertorische Komponente der Wissensbasis – als ABox bezeichnet. Beide Komponenten besitzen eine gemeinsame Semantik, die ABox-Einträge als Beschreibungen von Instanzen aus Objektklassen der TBox festlegt, so daß die Beschreibung in TBox und ABox zusammen konsistent ist.⁷

Die Repräsentation der Wortbedeutungen ist Bestandteil der TBox. Deshalb operieren die zu diskutierenden Verarbeitungsstrategien in derselben. Natürlich spielt die Unterscheidung zwischen beiden Wissensarten bei der automatischen Verarbeitung eines Textes eine Rolle, da der Text in eine ABox-Repräsentation übersetzt wird und somit auch die einzelnen Wörter entsprechende Repräsentationen als ABox-Eintrag erhalten. Die Bestimmung einer kontextuell relevanten Bedeutung ist aber ein Prozeß, der in der TBox abläuft, denn auf der Basis möglicher Wortbedeutungen in der TBox wird eine aktuelle in der ABox konstruiert. Obwohl KL-ONE nicht die Inferenzen zur Festlegung von Wortbedeutungen bereitstellt, kann es dennoch zur Repräsentation von Wortbedeutungen zumindest herangezogen werden, denn es bietet eine ganze Menge von Sprachkonstrukten an, die es gestatten, Konzeptwissen, und damit auch Wissen über Wortbedeutungen, elegant zu repräsentieren. Die Sprache erlaubt mit dem wichtigsten Konstrukt die Subsumtion von Begriffen, wodurch sich eine Taxonomie ergibt; es lassen sich also Begriffe unter- bzw. nebenordnen. Mit Kanten zwischen den Knoten können Kollektionen von Fakten vernetzt dargestellt werden. Diese Kanten werden „Rollen“ genannt; mit ihnen lassen sich 2-stellige Relationen zwischen Knoten darstellen. An jede Rolle können zudem Anzahlrestriktionen für deren Werte angegeben werden.

Die Rollen zeigen auf ein Konzept, das selber mit anderen Konzepten über Rollen verbunden ist. Damit lassen sich auch Rollen-Ketten darstellen. Schließlich seien noch Role-Value Maps genannt, die zwei Rollenketten mittels eines binären Vergleichsoperators verbinden. Außerdem können Operatoren eingeführt werden, die es ermöglichen, eine separate Rollenhierarchie aufzubauen und die Disjunktion von Konzepten und Rollen anzugeben.

All diese Sprachkonstrukte besitzen eine präzise Semantik, die eine Übersetzung der Konstrukte in äquivalente Formeln der Prädikatenlogik erster Stufe ermöglicht. Vernetzte Information kann zwar graphisch für Menschen überschaubar und intuitiv einleuchtend repräsentiert werden, ein Computer jedoch kann diese Repräsentationen nicht einlesen und

⁷ Zur Problematik der Abgrenzung beider Wissensbereiche siehe Owsnicki-Klewe (1990).

verarbeiten. Die Repräsentation muß in eine formale Darstellung gebracht werden, die algorithmisch verarbeitet werden kann als Ketten von Symbolen. Dies geschieht mittels der KL-ONE Sprache. Es werden Operatoren zur sequentiellen Angabe vernetzter Information bereitgestellt. Die Semantik für diese Operatoren erlaubt die Übersetzung in die Prädikatenlogik 1. Stufe, so daß eine äquivalente Formel dieser Logik angegeben werden kann. Z.B. läßt sich der Begriff „Biographie“ als ein Buch definieren, das über das Leben eines Menschen informiert. Dies läßt sich in einer KL-ONE Sprache, deren Syntax der Programmiersprache LISP folgt, wie folgt einfach beschreiben: *biographie* < (*and buch* (*all hat_ thema mensch*)). Die Konzepte „buch“ und „mensch“ sowie die Rolle „hat_ thema“ müssen natürlich auch in der Wissensbasis entweder als Primitive eingeführt oder beschrieben bzw. definiert werden. Die semantisch äquivalente prädikatenlogische Formel des Biographie-Eintrags ist: $\forall x,y: \text{biographie}(x) \rightarrow \text{buch}(x) \ \& \ \text{hat-thema}(x,y) \rightarrow \text{mensch}(y)$.

Die Sprachkonstrukte und Operationen stellen Werkzeuge zur Repräsentation von Wissen (auch über Wortbedeutungen) bereit. Wie läßt sich damit nun der „schwammige“ und unklare Bereich der Wortbedeutung in den Griff kriegen? Wie lassen sich unter Rückgriff auf eine Termbeschreibungssprache wie KL-ONE Wortbedeutungen repräsentieren?

Die Aufgabe der Wissensrepräsentationssprache ist es, alle möglichen Wortbedeutungen zu repräsentieren; also eine sehr fein strukturierte Repräsentation anzubieten und die Rolle der Verarbeitungsprozedur ist es, unter Berücksichtigung kontextueller Information die richtige Bedeutung zu bestimmen.

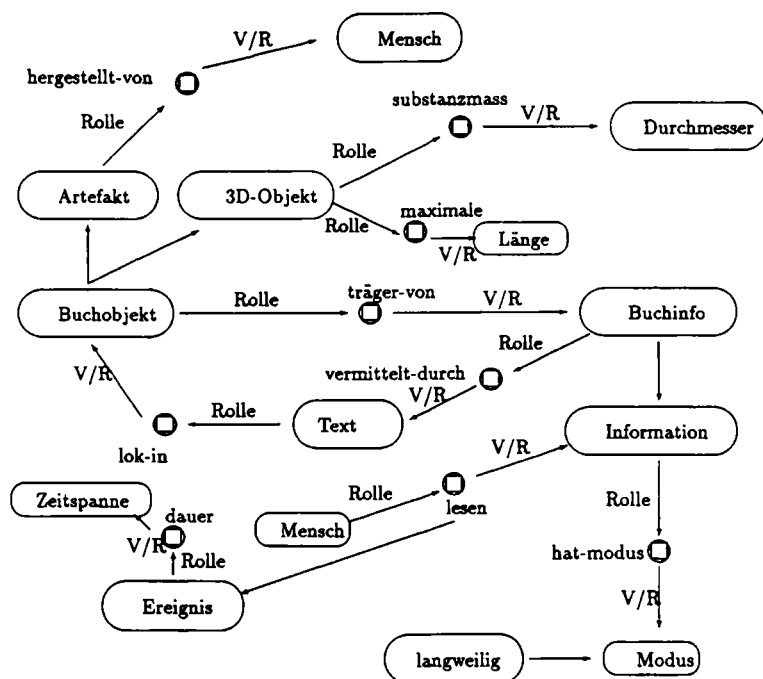
Eine adäquate Repräsentation der Bedeutungen von z.B. *Buch* muß etwa berücksichtigen, daß wir mit *Buch* mindestens auf die Konzepte „Buchobjekt“ und „Buchinformation“ vermittels selektierender Adjektive verweisen, wie die folgenden Phrasen zeigen:

- (1) dickes/teures Buch (OBJEKT)
- (2) langweiliges/lustiges Buch (INFORMATION)
- (3) schweres Buch (OBJEKT/INFORMATION)
- (4) langes Buch (OBJEKT/INFORMATION)

Während aber in (1) und (2) die Adjektive eine spezifische Lesart des Nomens selektieren, sind in (3) und (4) die Phrasen ambig. Abhängig davon, wie das Nomen interpretiert wird, ändert sich auch die Bedeutung des Adjektivs. Wenn in (3) *Buch* auf ein Objekt referiert, verweist das Adjektiv auf einen Punkt oder ein Intervall einer Gewichtsskala. Wird es dagegen als Information interpretiert, dessen Träger das Objekt ist, verweist das Adjektiv auf einen Punkt einer Schwierigkeitsskala

bezüglich des Verstehens, Schreibens oder anderer Prädikate. Ähnlich in (4): ein langes Buch-Objekt wird als überdurchschnittlich bezüglich einer Achse ausgezeichnet; eine lange Buch-Information wird als überdurchschnittlich bezüglich der Länge einer Aktivitätsphase (lesen, schreiben, ...) ausgezeichnet. Wir haben es hier also mit diffizilen Bedeutungsvariationen zu tun, die isoliert einerseits alle in der Wissensbasis vorhanden sein müssen, andererseits aber bei der Verarbeitung der gesamten Phrase erst miteinander kombiniert werden müssen. Das semantische Netz, das diesen Bedeutungsverschiebungen Rechnung trägt, ist in Graphik (1) angegeben.

'Buch'-Konzepte



Auch die Repräsentation der Bedeutungen von *Buch* in Graphik (1) ist natürlich nur eine grobe Annäherung an die komplexe Struktur des konzeptuellen Wissens über Bücher. So ist z.B. klar, daß Zeitspannen und Durchmesser bzw. Längen Werte auf einer Skala sind und nicht einfach Knoten eines Graphen. Wenn dann die Phrase *langes Buch* interpretiert wird, kann sie entweder auf einen Wert auf der Längenskala verweisen, der über dem Normwert liegt oder auf einen Wert auf der Zeitspannenskala des Lesen-Ereignisses, der über dem Normwert liegt. Dies ist mit einer netzartigen Repräsentation nur grob repräsentierbar. Ein weiteres Problem dieser Repräsentation ist folgendes: der konzeptuelle Zusammenhang der verschiedenen kontextabhängigen Bedeutungen wird nur über zweistellige Relationen zwischen zwei disjunkten Konzepten bereitgestellt. Das Problem, das sich nicht nur bei KL-ONE, sondern bei allen mir bekannten Wissensrepräsentationssprachen ergibt, ist die Repräsentation polysemer Wörter als über Relationen miteinander verbundene Menge disjunkter Konzepte. Dadurch wird Polysemie zu einem Spezialfall der Ambiguität. Dies ist aber aus grundsätzlichen Gründen inadäquat, da sich polyseme Wörter ganz anders als ambige Wörter verhalten können. So ist es möglich, daß polyseme Wörter ihre Gesamtheit an Bedeutungen in einen Diskurs mit einführen, und auf diese anaphorisch verwiesen werden kann. Ein Beispiel:

Peter kauft ein Buch (OBJEKT). *Es ist spannend* (INFORMATION).
Der Autor (DER INFORMATION) *hat den Umschlag* (DES OBJEKTS)
selbst entworfen.

Was man also braucht, ist eine Möglichkeit, die Wortbedeutung als Ganzheit zu repräsentieren, aus der ein Teilbereich im Kontext herausgeschnitten werden kann.⁸

3. Verarbeitungsstrategien zur Festlegung kontextadäquater Wortbedeutungen

Wie Wortbedeutungen in erster Näherung repräsentiert werden können, ist oben dargestellt worden. Wie kommt man nun im Kontext von der lexikalischen Einheit zur Äußerungsbedeutung?

⁸ Es gibt allerdings Beschränkungen bei anaphorischer Koreferenz mit polysemen lexikalischen Einheiten. Der Mini-Diskurs *Peter steht in der Tür* (DURCHGLASS). *Er streicht sie rot* (OBJEKT) erlaubt keine Koreferenz des Pronomens *sie* mit der Tür. Zur Zeit sind die Bedingungen für diese Beschränkungen jedoch völlig unbekannt. Sie scheinen auf der Relevanz der konzeptuellen Relation zu beruhen, sind aber noch nicht systematisch erfaßt worden.

Die Fähigkeit, Wortbedeutungen festlegen zu können ist nur ein Spezialfall der allgemeinen Fähigkeit, Probleme lösen zu können. Generell setzt die Fähigkeit, Probleme lösen zu können, voraus, Schlüsse ziehen zu können. Die inferentielle Mächtigkeit ist daher auch der Kern eines jeden wissensbasierten Systems. Ausdrücke der Repräsentationssprache werden aufgrund bestimmter Schlußfolgerungsverfahren verändert. Die unten vorgestellten Inferenzmechanismen gehen über die Mächtigkeit von KL-ONE hinaus. Sie müssen explizit neu als Inferenz definiert werden. Dies ist beim Marker Passing jedoch unproblematisch, da Marker Passing eine Suchstrategie ist, bei der die Semantik der Konstrukte im Grunde genommen irrelevant ist. Gewichtete Abduktion hingegen erfordert eine echte Erweiterung des Repräsentationsformalismus, damit die Werte angegeben werden können. Qualiastrukturen schließlich sind lexikonorientiert und berühren die konzeptuelle Repräsentation nicht.

3.1. Die Rolle des Lexikons bei der Festlegung von Wortbedeutungen

Eine der zentralen Fragen bei der Festlegung von Wortbedeutungen ist die, welche Rolle das Lexikon in der Wortsemantik spielt. Das Lexikon eines sprachverstehenden Systems muß für jeden lexikalischen Eintrag neben syntaktischer und morphologischer auch semantische Information enthalten, die zur Wissensrepräsentation systematisch in Beziehung gesetzt wird. Oftmals ist diese Schnittstelle einfach nur als Zeiger auf ein Konzept realisiert, so daß Polysemie nicht betrachtet wird. Bessere Ansätze arbeiten mit detaillierter Information als semantische Komponente, die beim Aufbau der Textrepräsentation erweitert und/oder modifiziert wird. Die Palette reicht von zweistufigen Ansätzen, die der Zwei-Ebenen-Semantik von Bierwisch/Lang (1987) folgen (z.B. Carstensen/Lang 1990; Meyer 1993), über abgeschwächtere Formen mit semi-autonomen Ebenen (Gust 1991) bis zur allgemein üblichen Sicht in der KI, semantisches Wissen als nicht sinnvoll trennbar vom konzeptuellen Wissen anzusehen, so daß das Lexikon als semantische Komponente Konfigurationen von Elementen der Wissensbasis enthält (z.B. Hobbs et al. 1987).

Wie kann die Verbindung zwischen einem lexikalischen Eintrag und dessen Konzepten genau aussehen? Im Fall der Angabe konzeptuellen Wissens im Lexikon entsteht das Problem, daß entweder eine Bedeutung als präferiert angegeben werden muß und von ihr alle anderen abgeleitet werden, oder es werden alle Bedeutungen disjunktiv repräsentiert. In beiden Fällen führen Selektionsrestriktionen der anderen lexikalischen Einheiten zur aktuellen Lesart. Effektiver ist es, die semantische Komponente eines Lexikoneintrags nur als Vermittler zur aktuellen Bedeutung

anzusehen, was in den (schwach) zweistufigen Ansätzen auch getan wird. Auch hier können mehrere Strategien gewählt werden. Es ist klar, daß dabei eine Methode, bei der im Lexikoneintrag eines Lexems nur ein Zeiger auf ein Konzept steht, in den meisten Fällen völlig inadäquat ist. So kann z.B. nicht in der semantischen Komponente des lexikalischen Eintrags von *Buch* einfach nur „buch“ stehen und in der Wissensbasis dann ein Eintrag der Art „buch = (or buchobj buchinfo)“. Hier würde man das Problem der Wortbedeutung ja nur in die Wissensbasis verschieben, denn es kann dann nicht gesagt werden, welche Obersorte „buch“ nun hat und welche Eigenschaften es vererbt bekommen soll von übergeordneten Konzepten. Zudem ist die interne Struktur der Konzepte als einfache Disjunktion zweier Konzepte angegeben, was – wie oben dargestellt – weder beschreibungs- noch erklärungsadäquat ist.

Wie die Beziehung zwischen Lexikon und Wissensbasis am sinnvollsten zu spezifizieren ist und wie die Unterscheidung zwischen semantischer Information im Lexikon und konzeptuellem Wissen – wenn es sie überhaupt gibt – zu definieren ist, ist ein Problem, zu dem in nächster Zeit keine befriedigende Antwort zu erwarten ist, da im Grunde genommen in der Semantik zu wenig über beide Gebiete bekannt ist. Die zwei unten dargestellten Inferenzmechanismen des Marker Passings (Hirst 1987) und Abduktion als Wortbestimmungsmethode (Hobbs et al. 1993) basieren auf der üblichen KI-Sicht, eine Trennung zwischen semantischer und konzeptueller Information nicht zu ziehen; die Angabe von Qualiastrukturen (Pustejovsky 1991) jedoch präsupponiert die Möglichkeit einer Trennung zwischen beiden Ebenen.

3.2. Marker Passing

Marker Passing ist das Pendant zur Quillianschen Inferenztechnik der Aktivierungsausbreitung. Die Inferenz der Aktivierungsausbreitung wurde eingeführt, um das aus der Psycholinguistik bekannte Verhalten erklären zu können, daß die Aktivierung eines Konzepts immer auch die Aktivierung begrifflich mit dem Konzept verbundener Konzepte beinhaltet. Wenn also ein Konzept in einem semantischen Netz aktiviert wird, werden auch alle benachbarte Konzepte mit aktiviert, abhängig von einem Schwellenwert, der für die Aktivierung postuliert wird.

Entlang den Kanten in einem semantischen Netz werden regelgesteuert einfach Marker passiert und es wird versucht, zwischen zwei Konzepten eine Verbindung zu finden. Wenn z.B. die Wörter *Verlag* und *Buch* gelesen werden, werden die Marker aller Konzepte der Nomen passiert sowie alle mit diesen Konzepten in Verbindung stehenden Konzepte (bis zu einer bestimmten Tiefe). Im Durchschnitt beider Aktivierungen kann

sich dann eine Proposition „Verlagsinstitution veröffentlicht Buchobjekt“ befinden. Der Pfad zu dieser Information geht durch bestimmte Wortbedeutungen, die dann die kontextuell relevanten sind, nämlich *Verlag* als Institution und *Buch* als Objekt.

Entsprechend können auch die beiden Bedeutungen von *langes Buch* (s.o.) mittels Marker Passing ermittelt werden. Die präferierte Bedeutung der Phrase wird dann durch die Konzepte bestimmt, deren Distanz zueinander kürzer ist. Die Distanz zwischen zwei Konzepten wird also indirekt über die Aktivierung von Konzepten gemessen. Wenn beim Marker Passing ein Knoten erreicht wird, der selbst schon einmal passiert worden ist, dann wurde ein Pfad gefunden. Hiermit lassen sich dann Wortbedeutungen mittels Assoziationen festlegen. Dieses Prinzip tritt in vielen Varianten auf. Die einfachste ist, bis zu einer bestimmten Tiefe einfach alle Konzepte zu markieren. Der Marker beinhaltet nur die Information, von welchem Knoten im Netz die Aktivierung her kam. Andere Arten arbeiten mit abnehmender Aktivierung, wo zusätzlich zum Aktivierungsursprung noch abnehmende Werte angegeben werden und zur Aktivierung eines Konzepts darf ein bestimmter Wert nicht unterschritten werden. Dies entspricht im wesentlichen der ursprünglichen Idee von Quillian (1968). Marker Passing ist damit eine elegante Methode, Beziehungen zwischen Konzepten zu finden.⁹

Allerdings tritt beim Marker Passing folgendes Problem auf: Leider sind in einem semantischen Netz immer zwei Konzepte über irgendeinen Pfad miteinander verbunden. Dadurch können natürlich auch völlig unsinnige Bedeutungen über lange Pfade miteinander verbunden werden. Abgesehen von der Angabe von Schwellenwerten ist nicht bekannt, nach welchen Prinzipien noch sinnvolle von sinnlosen Assoziationspfaden unterschieden werden sollen, wenn nur die Länge von Pfaden (bzw. Schwellenwerte) als Kriterium gelten sollen.¹⁰ Was aussteht, ist eine Theorie der qualitativen Verbindung zweier Konzepte. Außerdem funktioniert Marker Passing recht gut für ambige Wörter, wie Hirst (1987) zeigt; die feinen semantischen Unterschiede bei Polysemie lassen sich damit aber nur schwer in den Griff kriegen, da es sich hier um sehr detaillierte semantische Netze handelt, in denen sich teilweise Distanzen der gleichen Länge zwischen zwei verschiedenen Wortbedeutungen befinden.

⁹ Wie elegant Marker Passing ist, zeigt sich auch darin, daß komplexe Inferenzen auf diese Technik zurückgeführt werden können (Norvig 1987) und auch zum Aufbau semantischer Netze während der Verarbeitung von Texten verwendet werden kann (Geurts 1989).

¹⁰ Weitere Kritik findet sich in Geurts (1989).

Ein Beispiel: während in einem semantischen Netz die Wortbedeutungen von *Fliege* als Insekt und als Kleidungsstück weit auseinanderliegen und somit Marker Passing bei der Verarbeitung eines Satzes wie *Die Fliege ist im Kleiderschrank* die Kleidungsstück-Lesart aufgrund der Nähe von „Kleidungsstück“ und „Kleiderschrank“ im semantischen Netz wahrscheinlicher als die andere Lesart macht, sind die verschiedenen Bedeutungen von *Museum* direkt über konzeptuelle Relationen miteinander verbunden: ein Museumsgebäude stellt den Platz für die Institution und die Sammlung bereit; die Institution befindet sich im Gebäude und organisiert die Sammlung etc. Es ist dann z.B. nicht hinreichend klar, wie (bzw. ob) Marker Passing die Unterschiede in der Assoziationstärke bei dem Satz *Peter findet das Museum interessant* zwischen der Gebäude- und der Institutionslesart herausfinden kann. Man kann beides interessant finden, aber die bessere Interpretation ist durch die Gebäudelesart gegeben. Wie dies mittels Marker Passing funktioniert, ist unklar.*

3.3. Abduktion

Eine andere Technik, die immer mehr in den Vordergrund nicht nur in der Wortsemantik rückt, ist (gewichtete) Abduktion. Abduktion ist ein Mechanismus zur Generierung von Erklärungen. Im Gegensatz zur Deduktion, wo gültige Schlüsse zur Gewinnung impliziten Wissens formuliert und angewandt werden, bedeutet Abduktion die Festlegung einer Annahme, die sich unter Umständen als falsch herausstellen könnte.

Um ein plakatives Beispiel zu geben: Aus der Implikation „Wenn eine Person im Lotto gewinnt, ist sie glücklich“ und „Peter hat im Lotto gewonnen“ kann deduktiv mittels des Modus Ponens geschlossen werden, daß Peter glücklich ist. Ein abduktiver Schluß wäre es, aus der Implikation und der Information „Peter ist glücklich“ zu mutmaßen „Peter hat im Lotto gewonnen“. Hier wird also eine Hypothese aufgestellt (es könnte ja auch andere Gründe haben, warum Peter glücklich ist), die evtl. modifiziert werden muß.

Natürlich gibt es plausible Erklärungen für ein Phänomen und weniger plausible. Um diesem Unterschied Rechnung tragen zu können, werden Werte berechnet für den abduktiven Schritt im Falle der gewichteten Abduktion. Abduktion ist also im gewissen Sinne die Einbeziehung von Wahrscheinlichkeiten und Güte der Erklärung eines Datums.

* Bei der Abfassung dieses Artikels war mir jedoch die Arbeit von Mehl (1993) noch nicht bekannt, in der mittels einer speziellen Form von Aktivierungsausbereitung Lesarten polysemer Wörter im Kontext festgelegt werden.

Dieselbe Technik kann jetzt auch in der Verarbeitung von Wortbedeutungen eine Rolle spielen. Dann werden Hypothesen generiert, welche Bedeutung des Wortes gemeint sein kann. Diese Methode ist in Hobbs et al. (1993) genauer formuliert. Allerdings ist Abduktion kein gültiger Schluß, da es mehrere Möglichkeiten geben kann, warum ein Fakt besteht. Aus diesem Grund soll es möglich sein, die Annahmen mit variierenden Kosten zu belegen. Jeder Term der logischen Form eines natürlichsprachlichen Ausdrucks wird mit einer Nummer versehen ($\$n$), die die Kosten angibt, wenn angenommen werden soll, daß der Term wahr ist. Jeder Term muß entweder angenommen oder bewiesen werden. Dies wird durch Regeln in der Wissensbasis bewerkstelligt. Jede Regel ist ebenfalls mit Kostenfaktoren ausgestattet und die Interpretation eines Ausdrucks L der Logiksprache ist dann die Menge der Propositionen mit den minimalen Kosten, aus denen L folgt. Die logische Form des Satzes *Peter sitzt auf der Bank* ist z.B.:

$\text{Peter}(x)^{(\$2)} \ \& \ \text{sitzt}(x)^{(\$2)} \ \& \ \text{auf}(x,y)^{(\$20)} \ \& \ \text{bank}(y)^{(\$10)}$

$P(x)^{(\$n)}$ bedeutet: die endgültige Interpretation muß entweder P für $\$n$ annehmen oder P beweisen, evtl. für geringere Kosten. Die Kosten werden an die Wissensbasis weitergegeben. Dort stehen Regeln der Art $P1^{(a1)} \ \& \ ... \ \& \ Pn^{(an)} \rightarrow Q$. Wenn Q die Kosten c hat, erhält $P1$ die Kosten $a1c$, $P2$ die Kosten $a2c$ usw. Angenommen, die folgenden Beweisregeln sind Bestandteil der Wissensbasis:

$\text{bankgebäude}(x)^{(.7)} \rightarrow \text{bank}(x)$
 $\text{bankschalter}(x)^{(.9)} \rightarrow \text{bank}(x)$
 $\text{bankinstitution}(x)^{(.9)} \rightarrow \text{bank}(x)$
 $\text{banksitzgelegenheit}(x)^{(.5)} \rightarrow \text{bank}(x)$
 $\text{banksitzgelegenheit}(y)^{(.3)} \ \& \ \text{lokation-auf}(x,y)^{(.1)} \ \& \ \text{peter}(x)^{(.2)}$
 $\rightarrow \text{auf}(x,y)$

Die Gewichtungen $(.x)$ sind Präferenzinformationen, die besagen, daß der Term für Kosten $.xn$ abgeleitet werden kann. Bei diesen Werten sind die folgenden Interpretationen zu den entsprechenden Kosten zu erhalten:

- 1) einfach die logische Form anzunehmen, kostet \$34 (nämlich die Addition der einzelnen Kosten).
- 2) Die Interpretation „Peter sitzt auf dem Bankgebäude“ erhält \$31 (Der abduktive Schritt von „bank“ nach „bankgebäude“ verringert die Kosten von \$10 nach \$7).

- 3) „Peter ist lokalisiert auf der Banksitzgelegenheit und zwar sitzt er auf ihr“ erhält die Kosten \$11.¹¹

Um einen korrekten Abduktionsschritt durchführen zu können, muß man sowohl wissen, wie stark ein Faktum für oder gegen eine Konklusion spricht als auch, wie einzelne Fakten kombiniert werden müssen, um für oder gegen eine Konklusion zu sprechen. Wie die Kosten und Gewichte genau anzusetzen sind, ist dann natürlich eine empirische Frage, und hier liegt auch schon ein Problem des gerade beschriebenen Ansatzes: die Gewichtung ist ad-hoc. Um eine entsprechende Wissensbasis aufbauen zu können, die für einen festgelegten Weltausschnitt die Gewichtung der einzelnen Prämissen angibt, muß empirisch festgestellt werden, welche Fakten am ehesten für eine Konklusion sprechen. Wahrscheinlich ist die Gewichtung aber auch kontextabhängig. Kontextuelle Information trägt entscheidend zur Disambiguierung von Wörtern und zur Festlegung von polysemen Lesarten bei. Regeln mit festen Werten können aber nur kontextunabhängig Präferenzen angeben. Dies würde bedeuten, daß man dynamisch während der Interpretation die Gewichtung ändern können muß.

Ein weiteres Problem tritt bei Polysemie auf. In dem eben beschriebenen Ansatz würde die Komplexität des konzeptuellen Netzes, das zum polysemen Lexem gehört, einfach auf die Disjunktion der verschiedenen Lesarten reduziert werden, Polysemie also im Grunde wie Ambiguität behandelt werden. Dies ist aber aus den oben genannten Gründen nicht zulässig, denn ein polysemes Item kann ja seine verschiedenen Lesarten in einem Diskurs einführen, ein ambiges Wort aber nicht. Zusätzliche Probleme des oben umrissenen Ansatzes sind in Norvig/Wilensky (1990) beschrieben. Aber trotz der existierenden Probleme scheint Abduktion ein vielversprechender Ansatz zu sein. Mit dieser Technik lassen sich z.B. die Häufigkeiten von Interpretationen eines Wortes berücksichtigen.

¹¹ Wir würden die folgende abduktiv ermittelte logische Form mit den entsprechenden Werten erhalten:

Peter(x)^(\$2) & sitzt(x)^(\$2) & banksitzgelegenheit(y)^(\$5)
& banksitzgelegenheit(y)^(\$0) & lokation-auf(x,y)^(\$2) & Peter(x)^(\$0).

Die in der Prämisse mit Konklusion „auf(x,y)“ auftauchenden Terme, die vorher schon bewiesen wurden, erhalten keinen Wert (Siehe auch Norvig/Wilensky 1990). Der hier dargestellte Abduktionsmechanismus ist eine stark vereinfachte Variante dessen, was Hobbs et al. (1993) vorgeschlagen haben. In deren Interpretationsprozess spielt ein „etc“-Prädikat noch eine Rolle, das benutzt wird, um das Problem der Partialität von Definitionen zu umgehen, das ich aber aus Gründen der Übersichtlichkeit nicht beschreiben werde.

Wortbedeutungen, die präferiert benutzt werden, erhalten geringere Kosten als Wortbedeutungen, die nicht präferiert sind. Auch hier ist es wieder eine empirische Frage, welche Bedeutungen welche Gewichtungen erhalten sollen.

3.4. Qualiastrukturen

Eine Theorie, bei der die Rolle des Lexikons in der Wortsemantik mehr berücksichtigt wird, ist Pustejovskys (1989), (1991) Theorie der Qualiastrukturen.¹² Diese Theorie wurde von Pustejovsky entwickelt, um die Kontextabhängigkeit der Bedeutung von Artefaktnomen und Adjektiven erklären zu können. Er geht davon aus, daß die semantische Repräsentation der Nomen im Lexikon durch die sogenannte Qualiastruktur bestimmt ist. Die die Qualiastruktur ausmachenden Elemente bestimmen die Bedeutung der Nomen und sind für den Prozeß der Satzkomposition verfügbar. Die Qualiastruktur besteht aus vier Komponenten, die verschiedene Bedeutungsaspekte wiedergeben: die Angabe der Teile, seine unterscheidenden Merkmale von anderen Klassen, die „telische Rolle“; i.e. sein Zweck und seine Funktion und schließlich eine Angabe der Ursprungsfaktoren. Ein Beispiel sei die semantische Repräsentation von *Auto* im Lexikon (nach Boguraev/Pustejovsky 1990, S. 40):

auto(*x*)

[Konstitutiv: {karosserie, motor,...}]

[Form: auto-Gestalt(*x*)]

[Telisch: bewegen(P,*x*), fahren(P,y,*x*)]

[Ursprung: artefakt(*x*)]

Für Adjektivmodifizierungen soll dann gelten, daß Adjektive immer Elemente einer Komponente des Eintrags modifizieren. *Schnell* z.B. spezifiziert immer die telische Rolle eines Nomen-Eintrags, da das Adjektiv immer einen Prozeß spezifiziert. *Schnell* erhält dann die folgende Bedeutungskomponente im Lexikon:

schnell(*x*) - > (Telisch: $\lambda P \exists E$ [schnell(E) & P(E,*x*)])

Angewandt auf *Auto* erhält man zwei Bedeutungen, die *schnelles Auto* haben kann: ein Auto, daß sich inhärent schnell bewegt und ein Auto, daß schnell fährt; je nachdem, welches Prädikat in der telischen Komponente appliziert wird. *Schnell* hat also nur einen Lexikoneintrag, der Aspekte der verschiedenen Nominalrepräsentationen herausgreift. Die unterschiedlichen Bedeutungen von *schnelles Auto*, *schneller Walzer*,

¹² Siehe auch Pustejovsky/Anick (1988), Boguraev/Pustejovsky (1990).

schneller Stenotypist, schneller Leser etc. beruhen alle auf demselben Mechanismus.

Wie sieht es nun mit polysemen Nomen aus? Wie können Qualiastrukturen für polyseme Nomen aussehen und wie werden sie in der Satzsemantik angewandt? Auch bei polysemen Nomen geht Pustejovsky davon aus, daß sie nur einen semantischen Eintrag im Lexikon besitzen und dieser die Information für das polyseme Verhalten bereitstellt. Wenn z.B. die Qualiastruktur einer NP mit dem Argument eines Verbs kombiniert wird, können die Elemente der Qualiastruktur benutzt werden, wenn das Nomen die Belegungsbedingungen für das Argument nicht erfüllt. Hierzu wieder ein Beispiel: Nomen wie *Tür* und *Fenster* sind polysem, da sie einerseits auf das Objekt, andererseits aber auch auf den Durchgang referieren. Die Qualiastruktur von *Tür* ist (Pustejovsky 1991, S. 432):

tür(*x*)

[Konstitutiv: öffnung(*y*)]

[Form: phys-obj(*x*)]

[Telisch: hindurchgehen(T,z,*y*)]

[Ursprung: artefakt(*x*)]

Das Prinzip bei der Behandlung der Polysemie verläuft analog zur Adjektivmodifikation: die lexikalischen Items, deren Argument das Nomen belegt, erfordern einen bestimmten Typ. *Anmalen* muß auf die Form-Rolle von *Tür* angewandt werden, die Phrase *durch x gehen* erfordert die konstitutive Interpretation.

Auch dieser Ansatz birgt Probleme in sich. Der offensichtlichste Schwachpunkt liegt in der ad-hoc Formulierung der telischen Rolle, die ja essentiell zu sein scheint für den ganzen Ansatz. Es ist nicht klar, nach welchen Prinzipien Relationen als Bestandteile der telischen Rolle einzuordnen sind und nach welchen Richtlinien andere Relationen in das semantische Netz einzuordnen sind. Es scheint so zu sein, daß das Hineinpassen einer Relation in die telische Komponente selbst das Kriterium ist, warum sie dort lokalisiert ist, nicht das Ergebnis eines Kriteriums. Die angegebenen Qualiastrukturen zeigen zudem, daß Qualiastrukturen im Grunde genommen Konzeptwissen enthalten (nämlich Teil-Ganzes Relationen, Funktionen von Objekten etc.). Unter diesem Aspekt wird die unklare Unterscheidung zwischen Lexikon und Wissensbasis, die in der KI vorherrscht, auch bei den Qualiastrukturen nicht präzisiert, obwohl sie als eine Theorie gedacht ist, in der die Rolle des Lexikons mit seinen wortspezifischen Repräsentationen wieder in den Vordergrund rücken soll. Eine andere – völlig ungeklärte – Frage ist zuletzt, wie denn nun die Qualia-

strukturen mit Wissensbasen in Verbindung stehen. Hierüber sagt die Theorie nichts: es könnte sein, daß Elemente des Konzeptwissens einfach ins Lexikon übernommen werden, es kann aber auch sein, daß die Qualiastrukturen in einer komplexen Beziehung zum Konzeptwissen stehen. Alles in allem scheinen Qualiastrukturen recht ad-hoc formuliert zu sein und nicht prinzipiengelernt konstruiert zu sein, so daß sie zumindest als Theorie nichts über lexikalische und konzeptuelle Strukturen aussagen.

4. Zusammenfassung

Wir haben gesehen, daß Termbeschreibungssprachen als Subklasse semantischer Netze Formalismen sind, die in erster Näherung für die Repräsentation des Konzeptwissens über Wortbedeutungen herangezogen werden können. Wir haben aber auch gesehen, daß die Inferenzen zur Festlegung von Wortbedeutungen außerhalb der Ausdrucksmächtigkeit dieser Sprachen zu liegen scheinen. Alle diskutierten Ansätze arbeiten mit Inferenzen, die KL-ONE-Sprachen nicht anbieten. Hier scheint also eine Ecke für mögliche Erweiterungen zu sein, was insbesondere, aufgrund seiner Allgemeinheit, für Abduktion, aber auch Marker Passing, gilt. Probleme bereiten alle Ansätze. Marker Passing führt zum Problem bei der Polysemie. Qualiastrukturen sind möglicherweise nicht genügend mit Konzeptwissen verbunden. Abduktion hingegen scheint am ehesten zur Festlegung von Wortbedeutungen geeignet zu sein. Es scheint auch unsere intuitive Überlegungen zur Wortbedeutung widerzuspiegeln: Wörter referieren in einem Kontext nur mit einer gewissen Wahrscheinlichkeit auf eine konzeptuelle Struktureinheit; kommt weitere Information hinzu, kann sich die Bedeutung ändern. Obwohl die genaue abduktive Inferenz für eine Theorie zur Wortsemantik noch aussteht, scheint die Grundüberlegung, Hypothesen zur Bedeutung aufzustellen, ein Schwenk in eine vielversprechende Richtung zu sein.

5. Offene Probleme

Die Beispiele zeigen, daß die lexikalische Semantik und ihre Beziehung zur Satz- und Diskurssemantik im Grunde genommen erst in Ansätzen verstanden wird. Es existieren noch viele offenen Probleme. So existieren zwar recht mächtige und effiziente Formalismen zur Repräsentation von Wissen. Diese sind aber nicht unbedingt auf den komplexen Bereich der Wortsemantik ohne Modifikation anwendbar, wie die Polysemie zeigt. Zudem ist die Modellierung konzeptuellen Wissens mittels dieser Formalismen problematisch. Es kann immer nur unter einem bestimmten Aspekt eine Wissensstruktur angegeben werden, die zudem nur partiell

ist und sicherlich keine unter psychologischen Gesichtspunkten adäquate Repräsentation ist.

Die Verarbeitung von Wortbedeutungen bereitet ebenfalls noch große Probleme. Allein die Polysemie ist schwer in den Griff zu kriegen. Die metonymischen Aspekte sind noch viel schwieriger, wenn man vom Angeben ad-hoc formulierter Inferenzregeln abkommen will. Die Diversität der verschiedenen Ansätze zeigt m.E. zudem, daß man noch nicht weiß, welche Prinzipien hinter der Variabilität der Wortsemantik stehen. Die meisten Ansätze beschränken sich darauf, aus einer Menge vorgegebener Wortbedeutungen die dem Kontext angemessene auszuwählen. Dies bedeutet, daß alle theoretisch möglichen Bedeutungen im semantischen Netz angegeben werden und (in der Regel über Selektionsrestriktionen) eine oder mehrere ausgewählt werden. Diesem Ansatz liegt die Annahme zugrunde, daß jedes Wort auch nur eine endliche Menge von Bedeutungen hat. Während diese Annahme bei der Polysemie noch gerechtfertigt ist, zeigen metonymische Phänomene (und metaphorische, aber sie gehören erst recht in eine gesonderte Sparte), daß es vielleicht angemessener ist, kontextuell adäquate Wortbedeutungen zu generieren statt zu suchen. Generierung wäre ein wesentlich kreativerer Prozeß, der die Variabilität der Wortbedeutung angemessener wiedergibt. Natürlich stellen sich auch bei der Generierung von Wortbedeutungen viele Fragen. So sind die Regeln zur Erzeugung unbekannt; ebenso ist nicht klar, ob die Generierung von Basisbedeutungen ausgehen kann oder soll.

Weitere Probleme betreffen die Wissensrepräsentationsformalismen. Die Orientierung zur Logik in der Wissensrepräsentation gestattet es nicht, viele Merkmale menschlichen Wissens zu berücksichtigen wie etwa Vagheit, Unvollständigkeit und Inkonsistenz (siehe hierzu Delgrande/Mylopoulos 1986). Die präzise mengentheoretische Semantik für semantische Netze läßt es z.B. nicht zu, unscharfe Konzepte anzugeben, so daß Vagheit in dem Formalismus nicht berücksichtigt werden kann. Entsprechendes gilt für die anderen Bereiche. Probleme bereitet der an der klassischen Logik orientierte Ansatz auch für die Inferenzen. Wissensrepräsentationsformalismen wie KL-ONE lassen keine nicht-monotonen Inferenzen zu, mit denen Standardannahmen angegeben werden. Aber gerade nicht-deduktive Inferenzen wie etwa Abduktion scheinen für Wortbedeutungsbestimmungen geeignet zu sein. Diese werden aber nicht von KL-ONE-artigen Netzen unterstützt.

Die Rolle des Lexikons bei der Festlegung von Wortbedeutungen ist ebenfalls nicht hinreichend geklärt. Was steht im Lexikon als semantische Komponente? Dort scheint eher eine Konfiguration von Elementen an-

zugeben zu sein als nur ein Zeiger auf Konzepte, um Probleme wie etwa Unterbestimmtheit behandeln zu können. Das Fazit ist also: eine Theorie zur automatische Festlegung von Wortbedeutungen steht noch aus.

Literatur

- Bierwisch, M. (1983): Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In: Ruzicka, R./Motsch, W. (Hg.): Untersuchungen zur Semantik (studia grammatica XXII). Berlin: Akademie Verlag, S. 61-99.
- Bierwisch, M./Lang, E. (1987): Etwas länger – viel tiefer – immer weiter: Epilog zum Dimensionsadjektiveprojekt. In: Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven (studia grammatica XXVI + XXVII). Berlin: Akademie Verlag, S. 649-700.
- Boguraev, B./Pustejovsky, J. (1990): Lexical Ambiguity and the Role of Knowledge Representation in Lexicon Design. In: COLING 90. Helsinki: University of Helsinki, S. 36-41.
- Brachman, R. (1979): On the epistemological status of semantic networks. In: Findler, N. (ed.): Associative networks: Representation and use of knowledge by computers. New York: Academic Press, S. 3-50.
- Brachman, R./Schmolze, J.G. (1985): An overview of the KL-ONE knowledge representation system. In: Cognitive Science 9 (2), S. 171-216.
- Carstensen, K.-U./Lang, E. (1990): OSCAR – A PROLOG Program for Modelling Dimensional Designation and Positional Variation of Objects in Space. Stuttgart, IBM Deutschland GmbH, Institut für wissenschaftsbasierte Systeme, IWBS Report 109.
- Delgrande, J.P./Mylopoulos, J. (1986): Knowledge Representation: Features of Knowledge. In: Bibel, W./Jorrand, Ph. (eds.): Fundamentals of Artificial Intelligence. Berlin: Springer, S. 3-36.
- Dölling, J. (1992): Flexible Interpretationen durch Sortenverschiebung. In: Zimmermann, I./Strigin, A. (Hg.): Fügungspotenzen (studia grammatica XXXI). Berlin: Akademie Verlag, S. 23-62.
- Geurts, B. (1989): Processing Contrast Relations. In: Metzinger, D. (Hg.): GWA I 89. 13th German Workshop on Artificial Intelligence. Berlin: Springer, S. 196-201.
- Gust, H. (1991): Representing Word Meanings. In: Herzog, O./Rollinger, C.-R. (Hg.): Text Understanding in LILOG. Berlin: Springer, S. 127-142.
- Habel, C. (1986): Prinzipien der Referentialität. Berlin: Springer.
- Habel, C. (1990): Repräsentation von Wissen. In: Informatik-Spektrum 13, S. 126-136.
- Hirst, G. (1987): Semantic interpretation and the resolution of ambiguity. Cambridge: Cambridge University Press.
- Hobbs, J.R./Croft, W./Davies, T. (1987): Commonsense Metaphysics and Lexical Semantics. In: Computational Linguistics 13 (3,4), S. 241-250.

- Hobbs, J.R./Stickel, M./Appelt, D./Martin, P. (1993): Interpretation as Abduction. In: *Artificial Intelligence* 63; S. 69-142.
- Hörmann, H. (1983): Was tun die Wörter miteinander im Satz? oder Wieviele sind einige, mehrere, ein paar? Göttingen: Hogrefe Verlag für Psychologie.
- Krifka, M. (1989): Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen. München: Fink.
- Luck, K. von/Ownicki-Klewe, B. (1990): KL-ONE: Eine Einführung. In: Struß, P. (Hg.): *Wissensrepräsentation*. München: Oldenbourg Verlag, S. 103-121.
- Mehl, S. (1993): *Dynamische semantische Netze. Zur Kontextabhängigkeit von Wortbedeutungen*. Sankt Augustin: Infix Verlag.
- Meyer, R. (1993): *Compound Comprehension in Isolation and in Context*. Tübingen: Niemeyer.
- Norvig, P. (1987): Inference in Text Understanding. In: *Proceedings AAAI*, Seattle, Washington, S. 561-565.
- Norvig, P./Wilensky, R. (1990): A Critical Evaluation of Commensurable Abduction Models for Semantic Interpretation. In: *COLING 90*. Helsinki: University of Helsinki, S. 225-230.
- Ownicki-Klewe, B. (1990): A General Characterization of Term Description Languages. In: Bläsius, K.-H./Hedstüch, U./Rollinger, C.-R. (Hg.): *Sorts and Types in Artificial Intelligence*. Berlin: Springer, S. 183-189.
- Pustejovsky, J. (1989): Current Issues in Computational Lexical Semantics. In: *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics (EACL 89)*; S. xviii-xxv.
- Pustejovsky, J. (1991): The Generative Lexicon. In: *Computational Linguistics* 17 (4), S. 409-441.
- Pustejovsky, J./Anick, P. (1988): On the Semantic Interpretation of Nominals. In: *COLING 88*. Budapest: John von Neumann Society for Computing Science, S. 518-523.
- Pustejovsky, J./Bergler, S. (Eds.) (1992): *Lexical Semantics and Knowledge Representation*. Berlin: Springer.
- Quillian, M. (1968): Semantic Memory. In: Minsky, M. (ed.): *Semantic Information Processing*. Cambridge, Ma.: MIT Press, S. 227-270.
- Reichgelt, H. (1991): *Knowledge Representation. An AI Perspective*. Norwood, New Jersey: Ablex Publishing Corporation.