

Antwortgenerierung, flexible Wortwahl und elaborative Inferenzen - ein Regelinventar für LEU/2

Barbara Gängler, Ipke Wachsmuth

Zusammenfassung

Dieser Beitrag befaßt sich mit der Modellierung von regelhaftem Hintergrundwissen für ein textverstehendes System (LEU/2). Solches Wissen spielt eine zentrale Rolle sowohl bei der Konstruktion von semantischer Text-Repräsentationen als auch bei der Beantwortung von Anfragen. Die hier beschriebenen Arbeiten zur Hintergrundaxiomatik richten sich auf Bereiche und Aspekte des Alltagswissens, die über rein räumliche und zeitliche Informationen hinausreichen. Im vorliegenden Bericht werden einerseits Regeln beschrieben, die die Performanz des Systems dadurch verbessern, daß sie Repräsentationen in ein für bestimmte Arten der Informationssuche geeigneteres Format umsetzen. Zum anderen werden Regeln beschrieben, die die konstruierten Text-Repräsentationen zur read- oder question-time elaborieren, so daß implizite Textzusammenhänge für das System verfügbar gemacht werden. Da die hier vorgestellten Arbeiten nicht auf grundlegende theoretische Vorarbeiten aufbauen konnten, sollen vor allem die entwickelten Konzeptionen zur Diskussion gestellt werden.

Übersicht

Ein wichtiger Beitrag zur zweiten LILOG-Entwicklungsumgebung (LEU/2) lag in der Modellierung und Strukturierung des für LEU/2 notwendigen regelhaften Hintergrundwissens der gewählten Anwendungsdomäne. Neben der Ontologie kommt diesem domänenspezifischen Regelinventar zentrale Bedeutung sowohl bei der Konstruktion semantischer Repräsentationen als auch bei der Beantwortung von Anfragen in der Demonstration des aufgebauten Textwissens zu. Im speziellen richteten sich diese Arbeiten auf drei sehr unterschiedliche Teilaufgaben¹: Zum einen war es notwendig, das aus der semantischen Analyse der Sätze eines Textes hervorgehende Wissen für Generierungszwecke aufzubereiten. Diese Arbeiten werden in Abschnitt 2 beschrieben. Zum anderen mußten Möglichkeiten geschaffen werden, die dem Benutzer des Systems in Frage-Antwort-Dialogen eine flexible Wortwahl gestatten. In Abschnitt 3 werden die hierfür entwickelten Vorgehensweisen dargestellt. Darüber hinaus galt es, das aus den analysierten Texten extrahierte (explizite) Wissen durch regelhaftes (implizites) Hintergrundwissen zu vertiefen. Beim menschlichen Verstehen einer Äußerung wird in mehr oder weniger hohem Maße Weltwissen aktiviert, das inferentiell mit der aktuellen Äußerung in Zusammenhang gebracht wird. Von einem intelligenten natürlich-sprachlichen System würde man daher die Fähigkeit erwarten, Fragen, die gerade auf diese inferierten Zusammenhänge zielen, beantworten zu können. Die hier für LEU/2 gewählten Ansätze zur Elaboration von Textrepräsentationen werden in Abschnitt 4 beschrieben.

Ein Regelinventar zur Erfüllung der oben beschriebenen verschiedenartigen Aufgaben war für alle die Bereiche und Aspekte des "Alltagswissens" zu erstellen, die über die Verarbeitung rein

¹Die studentischen Mitarbeiter der Bielefelder LILOG-Arbeitsgruppe, Marianne Greten und Thomas Linke, haben wesentlich zu den hier beschriebenen Arbeiten beigetragen, wofür ihnen die Autoren an dieser Stelle herzlich danken.

räumlicher bzw. zeitlicher Information hinausreichen². Eine wesentliche Schwierigkeit lag jedoch in der Tatsache, daß bislang keine grundsätzlichen theoretischen Ansätze vorliegen, die diese Arbeiten richtungweisend leiten konnten. Deshalb sollen in diesem Bericht vor allem die zur Bewältigung der genannten Aufgaben entwickelten Konzeptionen und die Erfahrungen damit präsentiert werden.

Wissensaufbereitung für die Generierung

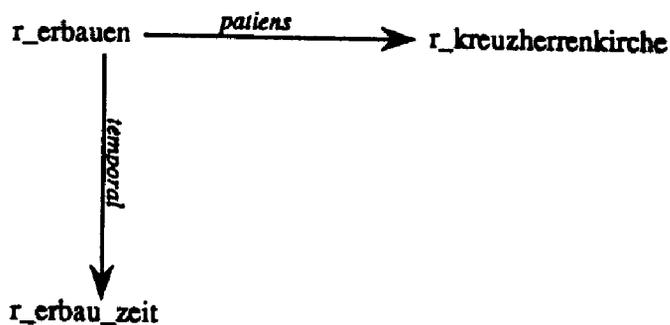
Eine erste Teilaufgabe bestand darin, das aus der semantischen Analyse hervorgehende Wissen in ein Format zu überführen, das es erlaubt, Informationen zur question-time für die Generierung von bestimmten Antworten effizient zur Verfügung zu stellen.

Nachdem ein Satz die linguistische Analyse durchlaufen hat, erscheint die Information in einem Format, bei dem das Referenzobjekt, das dem Prädikat des Satzes entspricht, eine zentrale Stellung einnimmt. Die beteiligten Objekte werden - entsprechend der thematischen Rolle ihrer natürlich-sprachlichen Pendanten - im wesentlichen als Argumente bezogen auf dieses Referenzobjekt repräsentiert. Beispielsweise erscheint der Satz (1) in einer Repräsentation (i), der Satz (2) in einer Repräsentation (ii)³:

(1) Die Kreuzherrenkirche, die im 15. Jahrhundert erbaut wurde, (...)

(i) refo r_erbauen : ERBAUEN.
 refo r_kreuzherrenkirche : KIRCHE.
 refo r_erbau_zeit : ZEIT.
 (...)
 axiom ax_1
 patiens (r_erbauen) = r_kreuzherrenkirche and
 temporal (r_erbauen) = r_erbau_zeit
 (...)

Graphisch kann eine solche Repräsentation wie folgt veranschaulicht werden:



²Arbeiten, die diese Aspekte behandeln, wurden von anderen Teilprojekten geleistet.

³Zur Repräsentationsprache L-LILOG vgl. L-LILOG-Syntax im Anhang. Zur Notation: Mit arabischen Zahlen werden natürlich-sprachliche Sätze oder Fragen bezeichnet, die zugehörigen L-LILOG-Repräsentationen mit entsprechenden kleinen römischen Zahlen (zu Satz (1) Repräsentation (i), zu Satz (2) Repräsentation (ii) usw.); Regeln werden durch (R1), (R2) usw., Literale mit (L1), (L2) usw. bezeichnet. Sorten werden durch Groß-, Attribute durch Kleinschreibung gekennzeichnet. Zugunsten der Lesbarkeit werden für die mit "refo" gekennzeichneten Diskurreferenten "sprechende Namen" gewählt; beim automatischen Aufbau der Repräsentation würden hier Namen wie z.B. drfnt27 erscheinen.

- (2) Die Kirche besitzt zwei parallele Schiffe.
- (ii) refo r_besitzen : BESITZEN.
 refo r_kirchenschiff : KIRCHENSCHIFF+.
 (...)
- axiom ax_2
 thema (r_besitzen) = r_kreuzherrenkirche and
 objekt (r_besitzen) = r_kirchenschiff
 (...)

In diesen Repräsentationen nimmt das Referenzobjekt, das dem Prädikat des analysierten Satzes entspricht (z.B. 'r_erbauen', 'r_besitzen'), die zentrale Stelle ein; die Objekte (z.B. 'r_kreuzherrenkirche', 'r_erbau_zeit') werden als Argumente mithilfe von (ein- oder zweistelligen) Relationen auf dieses zentrale Referenzobjekt bezogen. Da die gemeinsame Obersorte aller Referenzobjekte, die für diese zentrale Stellung in Frage kommen, die Sorte SITUATION ist, wird im folgenden von Repräsentationen wie (i) und (ii) als von "situationszentrierten Repräsentationen" gesprochen.

Für die Informationssuche sind diese situationszentrierten Repräsentationen in bestimmten Fällen ungünstig, etwa wenn eine Anfrage an das System auf bestimmte Eigenschaften eines Objektes (z.B. r_kreuzherrenkirche) zielt, wie in (3):

- (3) Was ist über die Kreuzherrenkirche bekannt?

In einem solchen Fall müßten von der Inferenzmaschine zur Generierung einer adäquaten Antwort sämtliche Axiome daraufhin durchsucht werden, ob in ihnen das fragliche Objekt ('r_kreuzherrenkirche') als Argument irgendeines Prädikates auftaucht. Dies hätte möglicherweise eine L-LILOG-Anfrage wie (iii) zur Folge:

- (iii) axiom ax_f3
 exists S : SITUATION;
 ?- patiens (S) = r_kreuzherrenkirche or
 thema (S) = r_kreuzherrenkirche or
 objekt (S) = r_kreuzherrenkirche or
 ...

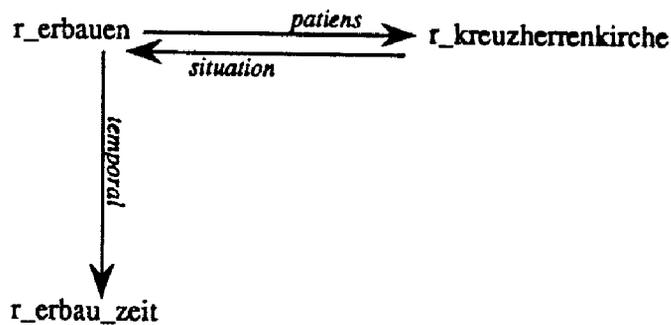
Das würde bedeuten, daß für jedes Referenzobjekt der entsprechenden Sorte (z.B. SITUATION) eine irgendwie definierte Menge von Prädikaten danach durchsucht werden müßte, ob das zugehörige Argument mit dem Referenzobjekt 'r_kreuzherrenkirche' belegt ist. In einem weiteren Schritt müßten dann die gefundenen Instanzierungen auf andere Relationen hin durchsucht werden, die nur indirekt mit dem fraglichen Objekt zusammenhängen, aber für die vollständige sprachliche Darstellung der Information unabdingbar sind, wie beispielsweise eine 'temporal'- oder 'lokal'-Relation der gefundenen Situation.

Da dieses recht aufwendige Verfahren zur question-time erfolgen müßte und dadurch das Antwortverhalten des Systems erheblich verlangsamen würde, ist es offensichtlich, daß eine effizientere Repräsentation gefunden werden mußte. Es sollte bereits zur read-time eine Verbindung zwischen dem situationszentriert repräsentierten Wissen und der für die beteiligten Objekte relevanten Information hergestellt werden. Dafür wurden zwei grundsätzlich verschiedene Herangehensweisen untersucht: Einerseits bestand die Möglichkeit, im Zusammenhang mit den Objekten selbst auf

die Situationen zu verweisen, an denen diese Objekte beteiligt sind. Eine andere Möglichkeit bestand darin, die ursprünglich situationszentrierte Information zusätzlich direkt an den betreffenden Objekten zu repräsentieren. In beiden Fällen jedoch sind es Objekte, die im Zentrum der repräsentierten Information stehen. Aus diesem Grunde wird hier im folgenden auch von "objektzentrierten Repräsentationen" gesprochen.

Die erste Vorgehensweise kann an den beiden Repräsentationen (i) und (i') und an der entsprechenden graphischen Darstellung veranschaulicht werden:

- (i) axiom ax_1
 patiens (r_erbauen) = r_kreuzherrenkirche and
 temporal (r_erbauen) = r_erbau_zeit.
- (i') axiom ax_1'
 situation (r_kreuzherrenkirche, r_erbauen).



Durch axiom ax_1' wird darauf verwiesen, daß das Objekt 'r_kreuzherrenkirche' an einer Situation 'r_erbauen' beteiligt ist; die zeitliche Spezifizierung dieser Situation ist - ebenso wie andere denkbare Information - wiederum in dem die Situation selbst betreffenden axiom ax_1 repräsentiert. Wird eine Frage nach den Eigenschaften des Objekts 'r_kreuzherrenkirche' gestellt, so findet sich im ersten Schritt bei diesem Objekt der Hinweis auf die Situation 'r_erbauen'; im zweiten Schritt kann dann die zeitliche Spezifizierung 'r_erbau_zeit' gefunden werden.

Eine Vorwärtsinferenzregel, die diese Verknüpfung zur read-time leistet, könnte wie (R1) aussehen:

- (R1) axiom ax_r1
 forall E : ERBAUEN,
 G : GEBÄUDE;
 EP⁴ patiens (E) = G
 →
 situation (G, E).

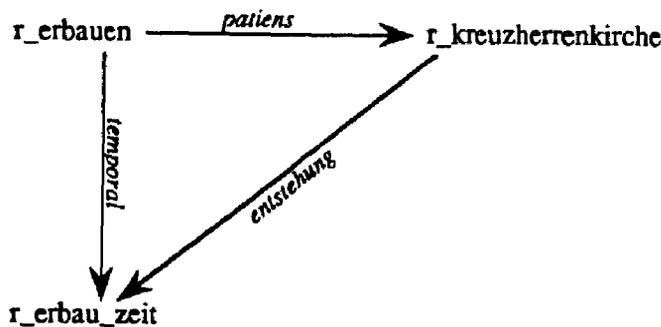
Mit diesem Verfahren würde man eine - gegenüber dem Durchsuchen sämtlicher Axiome - gezieltere Informationssuche bei einer Anfrage wie (3) ermöglichen.

Die zweite untersuchte Möglichkeit, eine Verbindung zwischen situationszentrierter und objektzentrierter Repräsentation zur read-time herzustellen, bestand darin, die situationszentriert repräsentierte Information *vollständig* auch bei den betreffenden Objekten zu repräsentieren. Ein

⁴In LLILOG kennzeichnet der Entrypoint (EP) das Literal, das bei der Abarbeitung der Regel als erstes bewiesen wird; da der Entrypoint hier vor einem Literal der Prämisse steht, handelt es sich um eine Vorwärtsinferenzregel.

der Repräsentation (i) entsprechendes objektzentriertes Pendant wäre beispielsweise (i'') (mit der entsprechenden graphischen Darstellung):

(i'') axiom ax_1''
entstehung (r_kreuzherrenkirche, r_erbau_zeit).



Bei dieser Art der Repräsentation wird dem Objekt 'r_kreuzherrenkirche' ein Attribut zugewiesen, das ein Verhältnis zu einem Zeitobjekt 'r_erbau_zeit' repräsentiert. Die Art dieses Verhältnisses ist durch die einführende Situation vorgegeben und kann sich im Namen der Rolle widerspiegeln.

Regeln, die per Vorwärtsinferenz diese Abbildung von situationszentrierter auf eine objektzentrierte Repräsentation leisten, sehen z.B. wie (R2) aus:

(R2) axiom ax_r2
forall E : ERBAUEN,
G : GEBÄUDE,
Z : ZEIT;
EP temporal (E) = Z and
patiens (E) = G
→
entstehung (G,Z).

Dieses Vorgehen bietet den Vorteil, daß im Falle einer entsprechenden Anfrage die objekt-relevante Information ohne lange Suchverfahren und damit schnell gefunden werden kann. Voraussetzung dafür ist allerdings in einigen Fällen die doppelte Repräsentation von Information.

Nicht sämtliche an Situationen beteiligten Objekte werden derartig behandelt; für die von uns behandelten Kurztexte war es naheliegend, in erster Linie Information über die im Text thematisierten Objekte (wie z.B. das Hetjensmuseum, die Kreuzherrenkirche usw.) objektzentriert darzustellen.

Beide geschilderten Verfahren bieten der Inferenzmaschine die Möglichkeit, der Generierungskomponente Aussagen über Objekte bereitzustellen, ohne die den Satzprädikaten entsprechenden Referenzobjekte zu kennen, mit denen das Objekt im zugrundeliegenden Text ursprünglich eingeführt worden war. Während jedoch bei der ersten Methode der Suchweg in der Regel länger ist und sich negativ auf das Zeitverhalten des Systems auswirkt, belastet das zweite Verfahren durch redundante Repräsentation die Speicherkapazität. Nach Abwägen dieser Vor- und Nachteile wurde aus Performanzgründen zugunsten der letzten Lösung entschieden.⁵

⁵Interessant wären hier empirische Auswertungen, die die Wahl eines der beiden Verfahren abgesehen von theoretischen Überlegungen motivieren könnten.

Flexibilisierung der Wortwahl in Anfragen

Die situationszentrierten Repräsentationen, die nach der satzweise erfolgten semantischen Analyse der Texte aufgebaut werden (vgl. Abschnitt 2), binden die Information an das Referenzobjekt, das das Prädikat des jeweils analysierten Satzes repräsentiert. Die Sortenzugehörigkeit dieses Referenzobjektes schränkt jedoch die Möglichkeiten zur Beantwortung von Fragen nach dieser Information dadurch ein, daß die Situations-Referenzobjekte der Anfragen dieselbe Sortenzugehörigkeit aufweisen müssen. Hieraus ergibt sich das Problem, daß in der Regel zunächst nur solche Anfragen beantwortet werden können, die sich wörtlich an die im vorgegebenen Text auftretenden Prädikate halten. Beim menschlichen Sprachverstehen besteht jedoch die Möglichkeit, durch unterschiedliche Wortwahl auf ein- und denselben Zusammenhang zu referieren. So ist es leicht vorstellbar, daß im Zusammenhang eines vom System verarbeiteten Satzes wie (1) vom Benutzer eine Frage wie (4) an das System gerichtet wird:

- (1) Die Kreuzherrenkirche, die im 15. Jahrhundert erbaut wurde, (...)
- (4) Wann wurde die Kreuzherrenkirche errichtet?

In L-LILOG wird Frage (4) folgendermaßen repräsentiert:

- (iv) axiom ax_f4
exists E : ERRICHTEN,
Z : ZEIT;
?- temporal (E) = Z and
patients (E) = r_kreuzherrenkirche.

Wenn für jedes Prädikat eines Satzes oder einer Frage nur ein Referenzobjekt einer entsprechenden Sorte eingeführt wird, ist die Frage (4) aufgrund des Textwissens allein noch nicht beantwortbar (vgl. Repräsentation (i)). In der Faktenbasis ist kein Referenzobjekt von der Sorte ERRICHTEN eingetragen, das mit der Variablen E unifiziert werden kann. Es wäre natürlich wünschenswert, solche Anfragen aufgrund der vorhandenen (grundsätzlich auch ausreichenden) textuellen Information beantworten zu können. Dazu muß eine Möglichkeit gefunden werden, die verschiedenen möglichen Verbalisierungen gegebener Zusammenhänge miteinander in Verbindung zu bringen.

Im o.g. Beispiel könnte z.B. durch eine Vorwärtsinferenzregel eine Verbindung von ERBAUEN zu ERRICHTEN hergestellt werden:⁶

ERBAUEN> ERRICHTEN

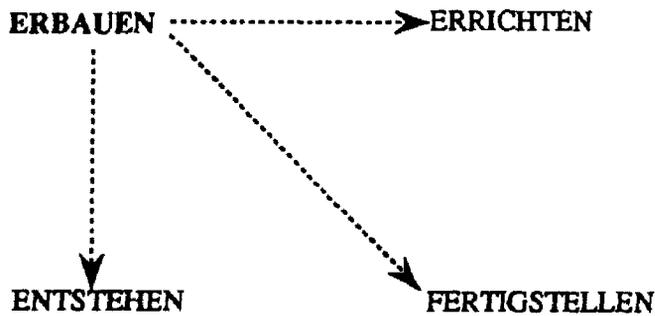
Frage (4) kann jedoch ebenso mit anderen Verben (etwa 'entstehen' oder 'fertigstellen') formuliert werden:⁷

- (4') Wann entstand die Kreuzherrenkirche?/Wann wurde die Kreuzherrenkirche fertiggestellt?

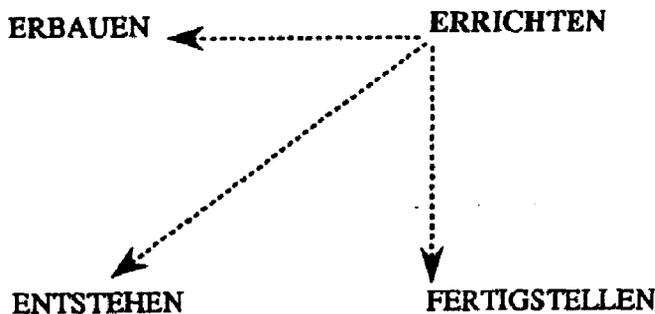
⁶In den folgenden Graphiken entspricht jeder gerichteten Verbindung zwischen Sorten eine zu erstellende Regel. Benötigte Vorwärtsinferenzregeln werden durch $\cdot\cdot>$ angedeutet, Rückwärtsinferenzregeln durch $\cdot\cdot\rightarrow$. Fettgedruckt erscheinen Sorten, von denen ein Referenzobjekt als im Text gegeben angenommen wird.

⁷Dabei handelt es sich bei diesen Verben nicht um Synonyme, wie sie beispielsweise durch ein Synonymlexikon zu behandeln wären; in den hier beschriebenen Verfahren wird lediglich der allen vier Verben gemeinsame Aspekt des Anfangszeitpunktes der Existenz eines Gebäudes betrachtet.

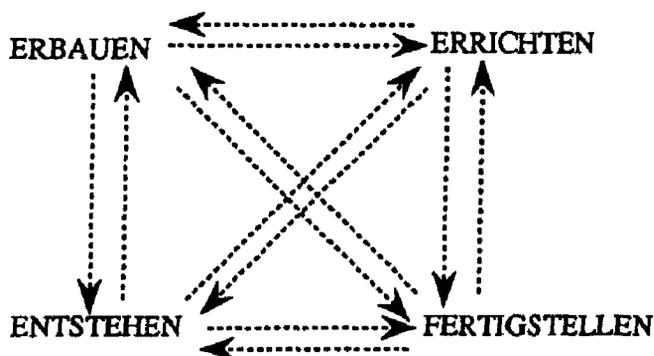
Für diesen Fall sind Verbindungen durch entsprechende Inferenzregeln auch zu diesen Sorten herzustellen:



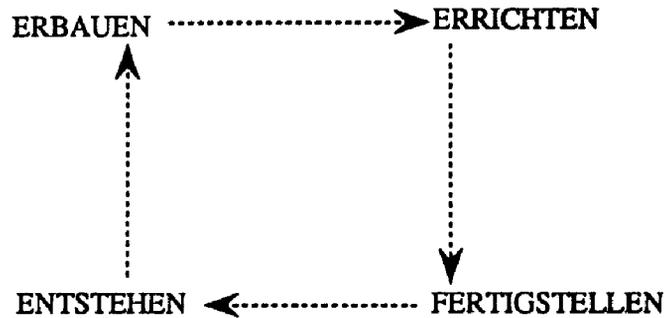
Wird jedoch in einem anderen zu analysierenden Satz eine ähnliche Beziehung von einem Gebäude und einer Zeitangabe mit dem Verb 'errichten' ausgedrückt, sollten die beteiligten Sorten auch mit der entsprechenden Sorte ERRICHTEN durch Regeln in Verbindung gebracht werden:



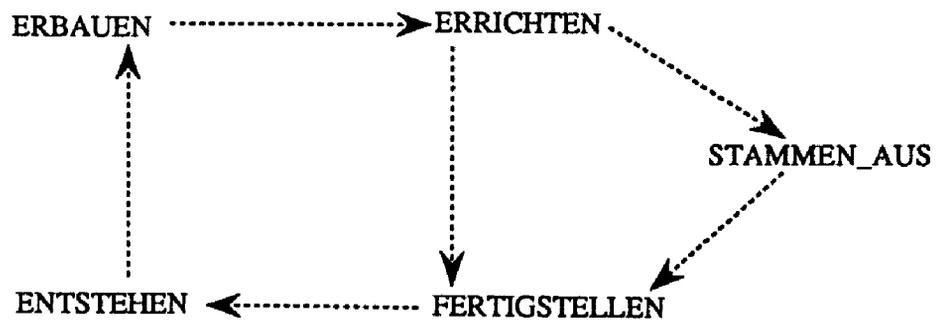
Es wird deutlich, daß letztlich - mehr oder weniger vollständig - alle Sorten aufeinander bezogen werden müssen:



Für diesen Lösungsansatz müßten für die vollständige Vernetzung von n Sorten $n \cdot (n-1)$ Regeln geschrieben werden. Reduzieren ließe sich dieser Aufwand, indem die beteiligten Sorten über Inferenzketten miteinander verbunden werden:



Hier wird die Verknüpfung von n Sorten mit nur n Regeln geleistet, die jedoch als Ketten aufeinanderfolgender Inferenzen realisiert sind. Ein Hinzufügen einer weiteren Sorte in dieses System hat zur Folge, daß zwei neue Regeln geschrieben werden müssen und eine der bisherigen Regeln überflüssig würde:



Gravierender jedoch als das Aufwandsproblem bezüglich der zu erstellenden Axiomatik ist die Tatsache, daß diese Art der Vernetzung der Sorten notwendigerweise zyklisch ist. Selbst wenn die in dieser Veranschaulichung gewählten Vorwärtsinferenzregeln durch Rückwärtsinferenzregeln ersetzt werden, können Zyklen nicht vermieden werden, wenn für ein zu beweisendes goal alle möglichen Lösungen gefunden werden sollen.⁸

Die Möglichkeit einer zyklensfreien Vernetzung der betroffenen Sorten bietet sich unter Rückgriff auf die in Abschnitt 2 beschriebenen Regeln zur Erzeugung objektzentrierter Repräsentationen. Diese Regeln überführen per Vorwärtsinferenz den situationszentriert dargestellten Zusammenhang von Satzprädikat und dessen Argumenten in ein objektzentriertes Format. So wird beispielsweise nach der semantischen Analyse von Satz (1) nach Anwendung von Regel (R2) zur read-time das Literal (L1) abgeleitet, das anschließend der Textwissensbasis hinzugefügt wird:

```

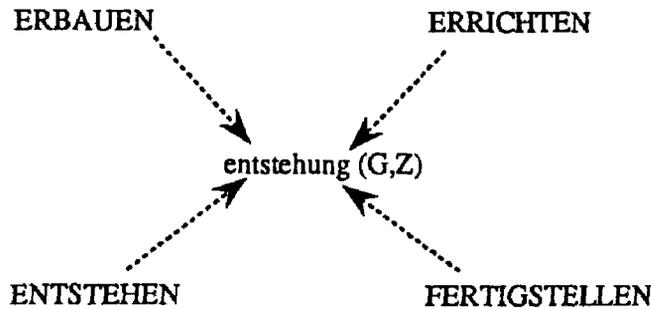
(R2) axiom ax_r2
      forall   E : ERBAUEN,
              G : GEBÄUDE,
              Z : ZEIT;
      EP temporal (E) = Z and
      patiens (E) = G
      →
      entstehung (G,Z).
  
```

⁸Die in LEU/2 verwendete Inferenzmaschine kann keine Zyklen erkennen und vermeiden. Zu vermeiden wären in diesem Modell Zyklen nur durch eine generelle Beschränkung der Inferenztiefe, was jedoch zu anderen - hier nicht zu diskutierenden - Problemen führt.

(L1) entstehung (r_kreuzherrenkirche, r_erbau_zeit)

Die zentrale Idee ist nun, diese auf das Objekt 'r_kreuzherrenkirche' bezogene Repräsentation (L1), die das Gebäude mit seiner Entstehungszeit in Verbindung setzt, als Bezugspunkt für die verschiedenen möglichen Frage-Formulierungen zu nutzen.

Andere Vorwärtsinferenzregeln, die von den Sorten **ERRICHTEN**, **ENTSTEHEN** und **FERTIGSTELLEN** ausgehen, stellen ebenfalls eine solche objektzentrierte Repräsentation her, so daß sich folgendes Bild zeigt:

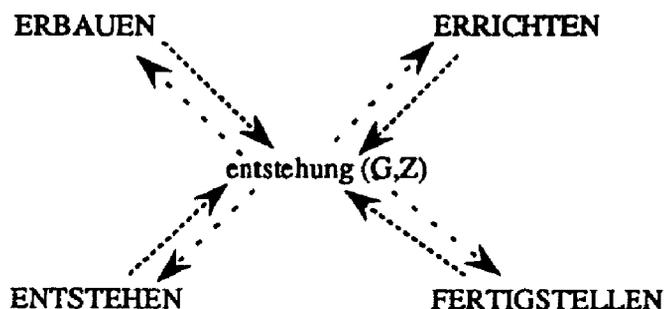


Auf der Basis dieser durch Vorwärtsinferenzregeln abgeleiteten objektzentrierten Repräsentation (L1) kann nun für Frage (4) eine positive Antwort gefunden werden - unabhängig von der Formulierung, mit der das Verhältnis von der Kreuzherrenkirche zu seiner Entstehungszeit im Text beschrieben wurde. Es müssen lediglich die fehlenden Verbindungen eingeführt werden, die von der als Bezugspunkt benutzten objektzentrierten Repräsentation zurück zu den situationszentrierten Repräsentationen führen. Um dabei jedoch erneute Zyklen zu vermeiden, wurden dafür Rückwärtsinferenzregeln gewählt. So kann für goal (iv) eine Lösung aufgrund einer solchen Rückwärtsinferenzregel (R3) gefunden werden:

```

(R3) axiom ax_r3
      forall G : GEBÄUDE,
           Z : ZEIT;
      exists E : ERRICHTEN;
           entstehung (G, Z)
      →
      EP patients (E) = G and
      EP temporal (E) = Z.
  
```

Bei Berücksichtigung der vier Sorten **ERBAUEN**, **ERRICHTEN**, **ENTSTEHEN** und **FERTIGSTELLEN** zeigt sich folgende Vernetzung:



Für die Anbindung jeder neuen Sorte müssen lediglich zwei weitere Regeln erstellt werden. So können - unter Ausnutzung der objektzentrierten Repräsentation - alle möglichen Kombinationen von gegebenen und erfragten Verbalisierungen berücksichtigt werden.

Auf diese Art wurde ein Regelwerk aufgebaut, durch das eine ganze Reihe weiterer Situationssorten aufeinander bezogen werden können:

ERBAUEN, ENTSTEHEN, ERRICHTEN, FERTIGSTELLEN
 ↔ entstehung

ZEIGEN, AUSSTELLEN, GEBEN_IN
 ↔ stellt_aus

UNTERGEBRACHT_SEIN_IN, BEINHALTEN, BEHERBERGEN,
 ↔ lok_sem⁹

usw.

Die Verbindung der hier genannten Situationssorten erfolgt durch den Bezug auf nur ein einziges Attribut. Durch dieses Verfahren sollen die beteiligten Situationssorten *nicht* wie Synonyme behandelt werden, sondern es soll lediglich ein Aspekt erfaßt werden, unter dem eine bedingte Austauschbarkeit möglich ist. Um die Beziehungen der Sorten zueinander präziser zu repräsentieren, müßte über ein differenzierteres System nachgedacht werden. Dennoch kann durch das beschriebene Vorgehen einem Benutzer innerhalb eines gegebenen Rahmens eine gewisse Flexibilität bezüglich seiner Wortwahl ermöglicht werden. Die grundsätzlichen Grenzen dieses Verfahrens liegen jedoch nach wie vor darin, daß in Anfragen nur exakt die Verbalisierungen behandelt werden können, die bereits bei der Entwicklung des Systems eingeplant worden sind. In unserem Fall lag der Wahl der vernetzten Sorten die (introspektive) Überlegung zugrunde, mit welchen Anfrageformulierungen in der Regel zu rechnen sein würde. Es hängt also von der Qualität der vorausgehenden Planung ab, ob die vom Benutzer gewählte Frageformulierung ein Hindernis für die Beantwortung der Frage darstellt oder nicht. Die Frage der Wortwahl-Flexibilisierung grundsätzlich zu lösen, muß weiterer Forschung vorbehalten bleiben.

Regelhaftes Hintergrundwissen für elaborative Inferenzen

Die Modellierung von inferentiellem Hintergrundwissen zur Anreicherung der konstruierten Textrepräsentation stellte eine völlig andersartige, weit schwierigere Aufgabe dar. Ein menschlicher

⁹Das Prädikat 'lok_sem' bezieht sich auf eine ähnliche Behandlung der fraglichen Sorten durch das Hamburger Teilprojekt und repräsentiert u.a. die Lokalisierung von Institutionen in Gebäuden.

Hörer einer Äußerung wird - sofern er die Äußerung versteht - in mehr oder weniger hohem Maße Weltwissen oder Standardannahmen einsetzen, die abhängig von seiner Alltagserfahrung und seinem aktuellen Kontext (Zustand, Interesse, außersprachliche Umgebung usw.) zum Verstehen der Äußerung erheblich beitragen. So läßt beispielsweise Satz (5) eine Reihe von Schlüssen wie etwa (6) plausibel erscheinen:¹⁰

- (5) Das argentinische Restaurant Montedoro am Graf-Adolf-Platz ist wegen seiner saftigen Rindersteaks bekannt.
- (6) (a) Da es sich um ein Restaurant handelt, kann man dort etwas essen.
 (b) In dem Restaurant Montedoro kann man Rindersteaks essen.
 (c) Rindersteak ist wahrscheinlich ein argentinisches Essen.
 (d) Das Montedoro bietet wahrscheinlich gutes Essen an, weil ein Restaurant, das wegen Essen bekannt ist, normalerweise gutes Essen anbietet.
 (e) Saftige Rindersteaks sind etwas Gutes.
 (usw.)

Von einem künstlichen sprachverstehenden System erwartet man - zumindest in Ansätzen - ein "Verständnis" eines eingegebenen Satzes in dem Sinne, daß die unmittelbar aus dem Text extrahierte faktische Information durch plausible Schlüsse angereichert oder *elaboriert* wird (*elaborative Inferenzen*).

Ein Problem bei der Simulation dieses "Verständnisses" stellt die potentiell unbegrenzte, zumindest aber nicht einheitlich bestimmbare Tiefe der Inferenzketten dar, die bei einer Äußerung bezüglich der angesprochenen Objekte aktiv werden. Zudem läßt sich diese Art von regelhaftem Wissen nicht durch funktionale Überlegungen kategorisieren, wie dies bei den in den Abschnitten 2 und 3 beschriebenen Regeln möglich ist. So braucht man (aus Effizienz- oder Kapazitätsgründen) ein Kriterium für die Entscheidung, welche der *inhaltlich* möglichen Inferenzen modelliert werden sollen und welche nicht. Da die Leistung von LEU/2 in erster Linie durch Frage-Antwort-Dialoge demonstriert wird, haben wir diese Entscheidung wiederum aufgrund der Überlegung getroffen, welche Anfragen typischerweise zu erwarten sind. Dabei konnten wir auf vorhandene Fragekataloge zurückgreifen, die zur Evaluation von LEU/2 vorab erarbeitet worden waren, jedoch für die hier betrachtete Thematik nicht ausreichten, da sie im wesentlichen Fragen nach räumlichen und zeitlichen Zusammenhängen enthielten. Die Fragenkataloge wurden deshalb um Fragen erweitert, die sich auf ein "Verständnis" des Textes beziehen, das durch das vorhandene Hintergrundwissen über den Weltausschnitt motiviert ist.

An zwei Fällen wird nachstehend erläutert, wie Textrepräsentationen durch Hinzufügen weiterer Axiome elaboriert werden, um Antworten auf Anfragen zu ermöglichen, die die Kenntnis impliziter Textzusammenhänge voraussetzen.

Bei Kenntnis von Text (7) ist es beispielsweise wünschenswert, eine Frage wie (8) beantworten zu können:

- (7) Das Hetjensmuseum wurde 1909 eröffnet. (...) Die Keramiksammlung des Museums umfaßt zehntausend Objekte. Die Ausstellung ist von 10 Uhr bis 17 Uhr geöffnet.

¹⁰Auf die Frage, welche Schlüsse es im Einzelfall sind, die unter anzugebenden Bedingungen für das Verstehen sprachlicher Äußerungen aktiviert werden und welches Wissen über einzelne Individuen hinaus für einen bestimmten kulturellen Kontext als gültig angenommen werden kann, kann und soll hier nicht eingegangen werden.

(8) Wann ist das Hetjensmuseum geöffnet?

Während die Auflösung des anaphorischen Bezuges von "Hetjensmuseum" und "Museum" ebenso wie die Identifizierung von "Keramiksammlung" und "Ausstellung" bereits bei der semantischen Analyse geleistet wird, muß aus Text (7) geschlossen werden, daß ein Museum, das eine Ausstellung beherbergt, normalerweise dieselbe Öffnungszeit wie die Ausstellung hat. Die Beziehung des Enthaltenseins von Keramiksammlung und Museum, die auf sprachlicher Seite durch die Genitiv-Phrase "Die Keramiksammlung des Museums" ausgedrückt ist, wird - ebenfalls im Zuge der semantischen Analyse - als 'sammlung (r_hetjensmuseum, r_keramiksammlung)' repräsentiert. Die dem Text (7) zugrundeliegende Repräsentation ist (vii):

- (vii) `refo r_geöffnet_sein11: GEÖFFNET_SEIN.`
`refo r_keramiksammlung_hm : KERAMIKSAMMLUNG.`
`refo r_t1 : ZEIT.`

`axiom ax_7`
`sammlung (r_hetjensmuseum,r_keramiksammlung_hm)`
`(...)`
`thema (r_geöffnet_sein) = r_keramiksammlung_hm and`
`temporal (r_geöffnet_sein) = r_t1.`

Regel (R4) erzeugt nach dem in Abschnitt 2 beschriebenen Verfahren eine objektzentrierte Version zu Repräsentation (vii):

- (R4) `axiom ax_r4`
`forall G : GEÖFFNET_SEIN,`
`I : INSTITUTION,`
`Z : ZEIT;`
`EP temporal (G) = Z and`
`thema (G) = I`
`→`
`öffnungszeit (I) = Z.`

Mit Regel (R5) wird die Öffnungszeit einer Institution, die sich innerhalb einer anderen Institution befindet, übertragen:

- (R5) `axiom ax_r5`
`forall I1,I2 : INSTITUTION,`
`Z : ZEIT;`
`EP öffnungszeit (I2) = Z and`
`sammlung (I1, I2)`
`→`
`öffnungszeit (I1) = Z.`

Das durch diese Regel erzeugte Literal ist (L2):

¹¹Zur Verwendung "sprechender Namen" siehe Fußnote 2.

(L2) $\text{öffnungszeit}(\text{r_hetjensmuseum}) = \text{r_t1}$.

Mit Regel (R6) kann - wie in Abschnitt 3 beschrieben - durch Rückwärtsinferenz zur Fragebeantwortung auf dieses Literal zugegriffen werden:

(R6) axiom ax_r6
 forall I : INSTITUTION,
 Z : ZEIT;
 exists G : GEÖFFNET_SEIN;
 $\text{öffnungszeit}(I) = Z$
 →
 EP thema(G) = I and
 EP temporal(G) = Z.

Damit sind die Voraussetzungen für eine erfolgreiche Beantwortung von Frage (8) gegeben.

Durch nur eine weitere Regel kann nun von der Öffnungszeit des Hetjensmuseums auf die Öffnungszeit des Gebäudes, in dem es untergebracht ist, geschlossen werden. Die Beziehung des Untergebracht-Seins ist in Satz (9) ausgedrückt; (ix) ist die zugehörige Repräsentation:

(9) Das Hetjensmuseum ist im Palais Nesselrode untergebracht.

(ix) refo r_unterbringen : UNTERBRINGEN.
 refo r_hetjensmuseum : MUSEUM.
 refo r_raumregion_pn : RAUMREGION.
 refo r_palais_nesselrode : PALAIS.

axiom ax_9
 patiens (r_unterbringen_hm) = r_hetjensmuseum and
 lokal (r_unterbringen_hm, r_raumregion_pn) and
 inlokspec (r_raumregion_pn, r_palais_nesselrode).

Regel (R7) erzeugt die erforderliche objektzentrierte Repräsentation, auf die Regel (R8) aufbaut:

(R7) axiom ax_r7
 forall U : UNTERBRINGEN,
 G : GEBÄUDE,
 I : INSTITUTION,
 K : KONTEXTVEKTOR,
 T : ENT_MIT_ZEITKOMP,
 R : RAUMREGION.
 EP patiens(U) = I and
 lokal(U, R) and
 inlokspec(R, G)
 →
 lok_sem(le:I, reg:R, temp:T, kontext:K).¹²

- (R8) axiom ax_r8
 o_defaultI : INSTITUTION,
 G : GEBÄUDE,
 R : RAUMREGION,
 Z : ZEIT,
 T : ENT_MIT_ZEITKOMP,
 K : KONTEXTVEKTOR;
 EP öffnungszeit (I) = Z and
 lok_sem(le:I, reg:R, temp:T, kontext:K) and
 inlokspec (R,G)
 →
 öffnungszeit (G) = Z.

Regel (R8) besagt, daß die Öffnungszeit einer Institution der Öffnungszeit des Gebäudes entspricht, in dem sich die Institution befindet, sofern keine widersprüchliche Information bekannt ist (o_default)¹³. Das durch (R8) abgeleitete Literal ist (L3):

- (L3) öffnungszeit (r_palais_nesselrode) = r.t1.

Mit einer entsprechenden Rückwärtsinferenz (wie etwa (R6)) kann dann für die Fragebeantwortung auf dieses abgeleitete Literal zugegriffen werden.

Ein anderer Fall, wo die Ergänzung des expliziten Faktenwissens durch regelhaftes Hintergrundwissen notwendig ist, sind Aussagen über Objekte, die sich in einer Teil-Ganzes-Beziehung befinden. Beispielsweise sollte aufgrund eines Textes wie (10) eine Frage wie (11) beantwortbar sein:

- (10) Die Kunstsammlung Nordrhein-Westfalen ist in einem auffälligen Gebäude untergebracht.(...) Die Sammlung umfaßt 180 Bilder. Die Hälfte der Bilder stammt von Paul Klee.
 (11) Zeigt die Kunstsammlung Nordrhein-Westfalen Bilder von Paul Klee?

Die dem Text (10) entsprechende Repräsentation (x) allein erlaubt die erforderlichen Schlüsse nicht:

- (x)
 (a) refo r_kunstsammlung_nrw : KUNSTSAMMLUNG.
 refo r_umfassen_nrw : UMFASSEN.
 refo r_bilder : BILD+.

 axiom ax_10a
 thema (r_umfassen_nrw) = r_kunstsammlung_nrw and
 objekt (r_umfassen_nrw) = r_bilder and
 #r_bilder = 180.

¹²Diese Regel wurde mit dem Teilprojekt in Hamburg abgestimmt.

¹³Durch "o_default" wird das in Regel (R8) abgeleitete Wissen als Standardannahme gekennzeichnet.

- (b) refo r_bilder_paul_klee : BILD+.
 refo r_stammen : STAMMEN_VON.
 refo r_paul_klee : MENSCH.

axiom ax_10b
 r_bilder_paul_klee in r_bilder¹⁴and
 thema (r_stammen) = r_bilder_paul_klee and
 objekt (r_stammen) = r_paul_klee.

Um die Frage (11) beantworten zu können, muß zunächst aus der Teilmengen- Beziehung der beiden Bildermengen ("180 Bilder" und "die Hälfte der Bilder") abgeleitet werden, daß die im ersten Satz erwähnte Sammlung auch die im zweiten Satz erwähnten Bilder von Paul Klee umfaßt. In einem zweiten Schritt muß dann inferiert werden, daß eine Sammlung, die Bilder "umfaßt", diese auch "zeigt".

Regel (R9) erzeugt per Vorwärtsinferenz die zugehörige objektzentrierte Repräsentation:

(R9) axiom ax_r9
 forall K : KUNST_INSTITUTION,
 U : UMFASSEN,
 O : DEPIC_OBJEKT*;
 EP objekt (U) = O and
 thema (U) = K
 →
 stellt_aus (K, O).

In Regel (R10) wird ausgedrückt, daß z.B. eine Sammlung, die eine Menge von Bildern umfaßt, auch die Teilmenge dieser Bilder umfaßt:

(R10) axiom ax_r10
 forall U : UMFASSEN,
 B1, B2 : BILD*;
 EP objekt (U) = B1 and
 B2 in B1
 →
 objekt(U)=B2.

Durch Regel (R10) wird das Literal (L4) abgeleitet:

(L4) objekt (r_umfassen) = r_bilder_paul_klee.

Dieses Literal stößt wiederum die Vorwärtsinferenzregel (R9) an, wodurch ein neues Literal (L5) erzeugt wird:

(L5) stellt_aus (r_kunstsammlung_nrw, r_bilder_paul_klee).

Das Literal (L5) läßt sich - wie in Abschnitt 3 beschrieben - zur Beantwortung der Frage (11) nutzen.

Die beiden in diesem Abschnitt beschriebenen Beispiele zeigen, wie eine Vielzahl von minuziös ineinander greifenden Regeln notwendig ist, um nur ansatzweise das Wissen nachzubilden, das beim Menschen eine entscheidende Rolle für das Verstehen einfachster Textzusammenhänge spielt.

¹⁴Mit der Formulierung "refol in refo2" wird eine Teilmengen-Beziehung ausgedrückt; auf eine präzisere Modellierung von "die Hälfte" wurde verzichtet.

Konkrete Erfahrungen und Ausblick

Der Schwerpunkt der im vorliegenden Dokument zusammengefaßten Beiträge zu LEU/2 lag zusammengefaßt darin, das analysierte Textwissen in verschiedener Hinsicht durch regelhafte Zusammenhänge aufzubereiten: Die semantischen Repräsentationen von Sätzen der LILOG-Texte mußten für die Generierung von Antworten auf Dialoganfragen aufbereitet werden. Für Benutzeranfragen mußte darüberhinaus eine flexible Wortwahl ermöglicht werden. Schließlich sollten durch die Bereitstellung von Hintergrundwissen implizite Textzusammenhänge für das System explizit verfügbar gemacht werden.

Für die Generierung von Antworten auf Anfragen mußte zunächst das sich stellende Problem der effizienten Informationssuche bewältigt werden. Hier wurde entschieden, die in der semantischen Analyse der Texte zunächst erstellten situationszentrierten Repräsentationen mit Hilfe speziell dafür erstellter Regeln in ein objektzentriertes Format zu überführen. Der in Kauf genommene zusätzliche Aufwand für die redundante Informationsdarstellung war dabei von doppeltem Vorteil. Einerseits zahlte er sich durch besseres Zeitverhalten aufgrund kürzerer Suchwege aus. Andererseits war das objektzentrierte Format auch für die Flexibilisierung der Wortwahl von Nutzen: Um bei der Benutzung des Systems in Frage-Antwort-Dialogen in begrenztem Umfang Freiheit bei der Wortwahl zu ermöglichen, mußten verschiedene Verbalisierungen gegebener Zusammenhänge miteinander in Verbindung gesetzt werden. Der zunächst naheliegende Weg einer wechselseitigen Vernetzung aufeinander zu beziehender Sorten durch Inferenzregeln hätte ein zyklisches Verhalten der Inferenzmaschine zur Folge gehabt; durch Rückgriff auf die objektzentrierten Repräsentationen als Bezugspunkt wurde eine zyklenfreie Vernetzung der Sorten bei linearem Aufwand (für jede hinzugefügte Sorte zwei zusätzliche Regeln) gewährleistet.

Schließlich wurde ein Inventar an regelhaftem Hintergrundwissen geschaffen, welches vermittelnde Inferenzen entlang der in den Texten beschriebenen Objekte und damit die Elaboration der unmittelbaren Textrepräsentationen durch Explizierung impliziter Zusammenhänge erlaubt. Der Kernpunkt dieser Arbeiten lag nicht in der Entwicklung neuer Konzeptionen, sondern in der inhaltlichen Modellierung eines Domänenausschnitts durch Standardannahmen des Alltagswissens; es mußte im Grunde vorhergesehen werden, welche Art von Fragen zu erwarten sind, um das System mit einer entsprechenden Hintergrundaxiomatik auszustatten. An dieser Stelle werden Grenzen des grundsätzlichen Vorgehens erkennbar, da die Vielzahl denkbarer impliziter Zusammenhänge, nach denen gefragt werden könnte, sich nur durch ein gegenüber dem tatsächlich erfaßten unvergleichlich reicheres Regelinventar zu bewältigen wäre - ein Massenproblem also.

Die konkreten Erfahrungen bei der Integration der entwickelten Regeln in das Gesamtsystem haben allerdings gezeigt, daß trotz verschiedener Effektivierungsbemühungen selbst bei dem vergleichsweise geringen Umfang der integrierten Regeln eine aufwendige und langwierige Suche oft nicht vermieden werden kann. Häufig wird bei der Antwortsuche zunächst eine unangemessen große Regelmenge berücksichtigt.

Bei der inhaltlichen Konzeption der domänenspezifischen Regeln für elaborative Inferenzen hat sich jedoch herausgestellt, daß in der Regel klar war, für welchen engeren Weltausschnitt die einzelne Regel jeweils geschrieben wurde. Daraus folgt, daß eine thematisch eingeschränkte Regelmenge bereits zu Beginn der Suche betrachtet werden könnte, wenn im Vorfeld der in einer Frage angesprochene Weltausschnitt identifiziert und die Suche überhaupt nur auf die diesem Weltausschnitt zugeordneten Regeln erstreckt wird. Voraussetzung dafür ist eine entsprechende Einteilung des gesamten Regelinventars in bereichsspezifische Module ('Wissenspakete', vgl. LILOG-Report Nr. 12). Die Modularisierung erfolgt aufgrund domänenbezogener Kriterien; für den oben genannten Fall könnte es beispielsweise sinnvoll sein, Wissen über Sehenswürdigkeiten in einem anderen Teilmodul abzulegen als Wissen über Restaurants. Falls sich bei einer anfänglichen Betrachtung des Textes die darin enthalte

Aus diesen Überlegungen ergeben sich eine Reihe von Forschungsfragen, die mit der thematischen Identifikation von Wissenspaketen und ihrer dynamischen Fokussierung zusammenhängen. Ein diesen Fragen zugrundeliegender Gedanke ist der, daß ein erfolgreicher Diskurs darauf beruht, daß alle

Gesprächspartner zu jedem Zeitpunkt wissen, über welchen Weltausschnitt kommuniziert wird; ein Sprecher muß also seinen Gesprächspartnern signalisieren, wenn er sich auf einen anderen Weltausschnitt beziehen will. Ein wechselnder Fokus auf verschiedene Weltbereiche muß sich also in irgendeiner Weise in der Diskursstruktur niederschlagen. Dabei sind es eine Reihe von verschiedenartigen sprachlichen Phänomenen, die diese Diskursstruktur indizieren. Eine Systemkomponente, die eine Fokussteuerung hinsichtlich relevanter Bereiche der Wissensbasis leistet, müßte sich an der Diskursstruktur und diesen sprachlichen Phänomenen orientieren. Was die Identifizierung von Wissenspaketen betrifft, die in verschiedenen Textsegmenten bzw. Anfragen angesprochen sind, so soll u.a. untersucht werden, welche Rolle signifikante Wörter oder Wortkonstellationen für diese Identifizierung spielen.

Mit dem grundlegenden Gedanken einer modular organisierten Wissensbasis für ein sprachverstehendes System würde die Möglichkeit geschaffen, auf Wissen gezielt zugreifen zu können, ohne den gesamten Wissensbestand durchmustern zu müssen. Hierüber weitere Erkenntnisse zu gewinnen, ist ein über die Konzeptionen für LEU/2 hinausgreifendes Forschungsziel.