

5 Semantik

Semantik

- Einleitung
- Aufgaben der semantischen Verarbeitung
- Semantische Netze
- Prädikatenlogik 1.Stufe als Semantikformalismus
- Semantikformalismus von MONTAGUE – Typenlogik
- Diskursrepräsentationstheorie (DRT)

5.1 Einleitung

Einleitung

- Die Bedeutung sprachlicher Ausdrücke spielt bei sprachverarbeitenden Systemen offenbar eine ganz zentrale Rolle.
- bestimmte eingeschränkte Aufgabenstellungen
 - Analyse und Synthese gesprochener Sprache
 - Rechtschreibkorrektur

lassen sich in gewissem Umfang und mit einem gewissen Erfolg ohne den Bezug auf den semantischen Aspekt von Sprache bewerkstelligen

Semantik ist wichtig für

- Textverstehungssysteme
- Generierungssysteme
- Dialogsysteme
- Übersetzungsaufgabe

5.2 Aufgaben der semantischen Verarbeitung

Aufgaben der semantischen Verarbeitung

Was ist die Bedeutung des folgenden Satzes?

Er beginnt um fünf im Raum 101.

1. Bedeutung der Wörter und ihre syntaktische Gruppierung:

Etwas fängt an einem bestimmten Tag um fünf Uhr (früh oder nachmittags) in einem Raum 101 an.

oder

Eine männliche Person beginnt unter den genannten zeitlichen und räumlichen Umständen mit einer nicht näher bezeichneten Handlung.

Aufgaben der semantischen Verarbeitung

2. aktueller semantischer Wert des Satzes in einem bestimmten gegebenen **Äußerungskontext**:

Äußerungskontext - Antwort auf die Frage:

Wann fängt der Vortrag von X an?

gestellt am 9.2.2004 mittags

er bedeutet **Vortrag von X**

um fünf bedeutet **9.2.2004, 17 Uhr**



Aufgaben der semantischen Verarbeitung

3. relevante Äußerungsinformation:

Die Kenntnis der Gepflogenheiten an deutschen Hochschulen legt den Schluß nahe, daß der Vortrag zirka um 17.15 Uhr beginnen wird.

Als Zuhörer des Vortrags kann man den nächsten Termin nicht um 18 Uhr, eventuell aber um 19 Uhr einplanen.

Als Dozent einer Vorlesung, die um 15.30 Uhr im selben Raum beginnt, sollte man auf ein pünktliches Ende achten.

Aufgaben der semantischen Verarbeitung

- Semantikkonstruktion
 - ermittelt das semantische Potential auf der Grundlage der lexikalischen und syntaktischen Information, die in der Eingabekette enthalten ist
 - also auf der Grundlage von linguistischen Wissen
- semantische Resolution
 - bestimmt den aktuellen semantischen Wert
 - unter anderem durch die Auflösung von Mehrdeutigkeiten (Disambiguierung)
 - und die Identifikation geeigneter kontextuell gegebener Referenzobjekte für kontextabhängige Ausdrücke
 - dabei wird im wesentlichen Kontextwissen benutzt

Disambiguierung

- Vorgang und Ergebnis der Auflösung lexikalischer oder struktureller Mehrdeutigkeit sprachlicher Ausdrücke durch den sprachlichen oder außersprachlichen Kontext

Lexikalische Mehrdeutigkeit

*Er sah das **Schloß** vor sich liegen.*

Durch den Zusatz – und hob es auf – wird die Mehrdeutigkeit beseitigt.

Die alten schützen die jungen.

Die alten schützen konnten den Schuss nicht mehr hören.

Strukturelle Mehrdeutigkeit

Die Wahl des Vorsitzenden fand Zustimmung.

- zwei mögliche Deutungen:
 - Dass der Vorsitzende gewählt wurde, fand Zustimmung.
 - Die Wahl, die der Vorsitzende getroffen hat, fand Zustimmung.

Ermittlung des Bezuges für Pronomen

*Eva gab dem **Mann** seinen Schlüssel.*

***Eva** gab dem Mann ihr Bild.*

*Hans ging in den Supermarkt. Er holte sich eine Flasche **Milch** in der Lebensmittelabteilung, bezahlte für **sie** und ging.*

***Eva** traf **Susi** in ihrer Wohnung. **Sie** konnten sich woanders nicht treffen, weil Evas Auto kaputt war.*

Ermittlung des Bezuges für Pronomen

*Tom schrie Eugen an, weil er (Eugen)
den Kaffee verschüttet hatte.*

Tom schrie Eugen an, weil er (Tom) Kopfschmerzen hatte.

Aufgaben der semantischen Verarbeitung

- semantische Auswertung
 - extrahiert durch die Anwendung von Deduktions- und Inferenzmechanismen auf den semantischen Wert einer Äußerung die relevante Äußerungsinformation
 - und dabei unter anderem Weltwissen einbezieht
- Bereitstellung eines semantischen Repräsentationsformalismus

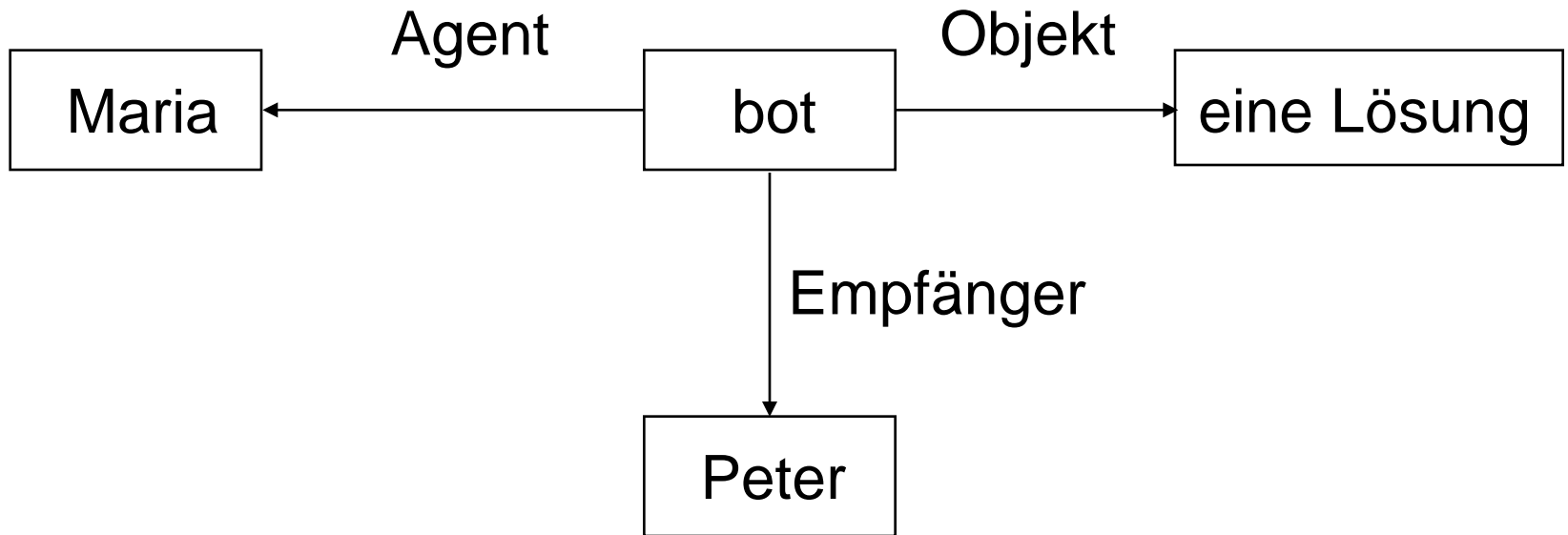
5.3 Semantische Netze

5.3.1 Einfache semantische Netze – CQ-Netze

Semantische Netze

- ein semantisches Netz ist ein gerichteter Graph
- Semantische Netze gehen auf Collins und Quillian (1970) zurück
- nach den Anfangsbuchstaben ihrer Nachnamen nennt man sie auch CQ – Netze
- das Sprachverarbeitungssystem EVAR zum deutschen IC - Zugsystem verwendet ebenfalls semantische Netze zur Wissensrepräsentation. (siehe auch 5.3.3)

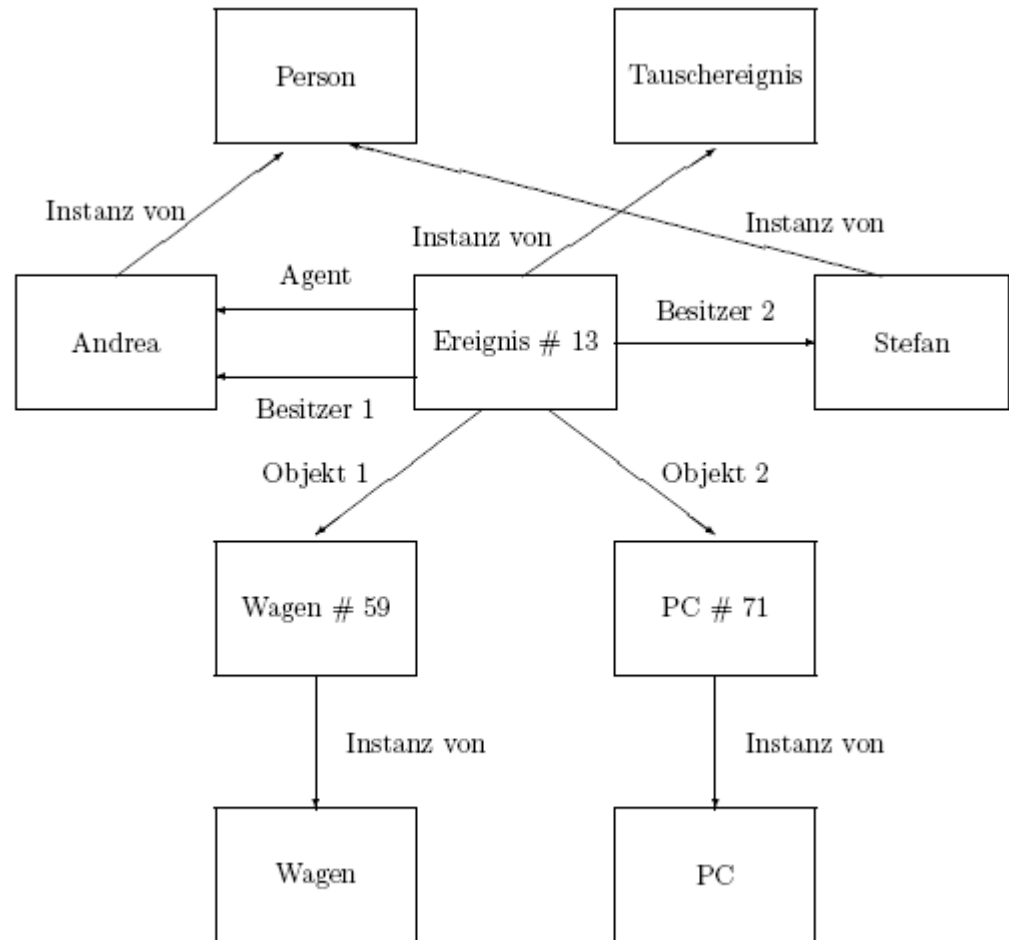
Semantische Netze – Beispiel



Maria bot Peter eine Lösung.

Semantische Netze – Beispiel

Andrea tauschte ihren
Wagen gegen Stefan's
Personalcomputer



5.3.2 CD – Theorie

Conceptual Dependency –
konzeptuelle Abhängigkeitstheorie

CD – Theorie

- wurde um 1970 von Roger Schank entwickelt
- Ziel war die geeignete Unterstützung von Inferenzen (semantische Auswertung)
- die CD – Theorie legt ein semantisches Netz zugrunde
- für die Knoten und Kanten wird ein beschränktes Inventar von Klassen festgelegt

Konzeptklassen

- **ACTs** – Aktionen
- **PPs** – Objekte (picture producers)
- **AAs** – Aktionsmodifikatoren (action aiders)
- **PAs** – Objektmodifikatoren (picture aiders)

Beispiel

$PP \Leftrightarrow ACT$

zeigt an, dass ein Handelnder handelt

$PP \Leftrightarrow PA$

zeigt an, dass ein Objekt ein bestimmtes Attribut hat

$ACT \xleftarrow{o} PP$

zeigt das Objekt einer Aktion an

Aktionen für Handlungsverben

- **ATRANS** – Veränderung einer Relation (z.B. geben)
- **PTRANS** – Veränderung eines physikalischen Orts eines Objekts (z.B. gehen)
- **PROPEL** – Anwenden einer physikalischer Kraft auf ein Objekt (z.B. stossen)
- **MOVE** – Bewegung eines Körperteils durch den Besitzer (z.B. treten)
- **GRASP** – Ergreifen eines Objektes durch einen Handelnden (z.B. packen)

Aktionen für Handlungsverben

- **INGEST** – Nahrungsaufnahme durch ein Tier (z.B. fressen)
- **EXPEL** – mit Kraft ausgeführte Hervorbringung aus dem Körper eines Tieres (z.B. schreien)
- **MTRANS** – Übertragung mentaler Information (z.B. erzählen)
- **MBUILD** – mentale Erzeugung neuer Information – Gedanken erzeugen und modifizieren (z.B. entscheiden)
- **CONC** – Konzeptualisierung einer oder Nachdenken über eine Idee (z.B. denken)
- **SPEAK** – Lauterzeugung (z.B. sagen)
- **ATTEND** – ein Sinnesorgan einem Stimulus zuwenden – Fokussierung auf ein Sinnesorgan (z.B. zuhören)

Beispiel

Hans wirft den Ball.

Hans \Leftrightarrow *PROPEL* $\xleftarrow{\circ}$ Ball

Beispiel

Jim erschoss John.

Die Gesundheit von John beträgt in der Vergangenheit -10 auf einer Skala von -10 bis +10, weil Jim eine Kugel, die sich in einem Gewehr befand, mit einem Impuls in Richtung auf John versehen hat.

Tiefenkasus (thematische Rollen)

- zugrunde liegt die Beobachtung, dass viele Handlungsverben (oder allgemeiner relationale Ausdrücke natürlicher Sprache) eine bestimmte Anzahl von Argumenten benötigen
- Tiefenkasus benennen die semantischen Rollen, die verschiedene Mitspieler in der durch das Verb bezeichneten Situation übernehmen
- jedes Verb selektiert eine bestimmte Menge von Tiefenkasus, die seinen Kasusrahmen bilden
- thematische Rollen sind Modifikatoren an den Kanten

Beispiel – kaufen

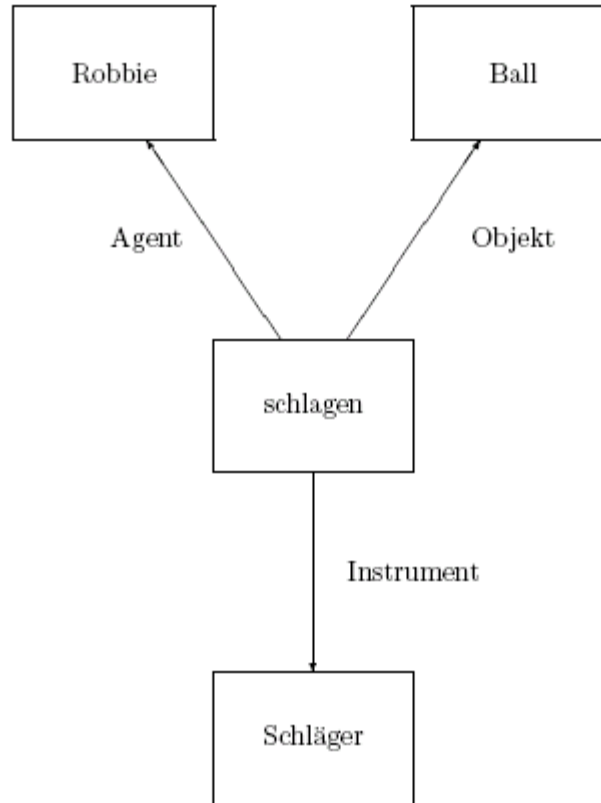
- *Käufer*
- *Verkäufer*
- *Ware*
- *Nutznießer*
- *Preis*

allgemeine thematische Rollen

- **Agent** – ist das Ding, das die Aktion geschehen läßt
- **Coagent** – ein etwas untergeordneter Partner des Agenten, sie führen die Aktion gemeinsam aus
- **Objekt** – ist das Ding, um das es im Satz wirklich geht und das einer Zustandsänderung unterzogen wird
- **Instrument** – ist ein vom Agent benutztes Werkzeug
- **Rezipient** – ist die Person, für die eine Aktion ausgeführt wird

Beispiel

Robbie schlug einen Ball
mit einem Schläger



Vorteile

- Sätze mit gleicher Bedeutung werden auf gleiche Weise dargestellt, z.B.:
 - Patrick gibt John eine Zeitung.
 - Patrick überreicht John eine Zeitung.
 - John erhält eine Zeitung von Patrick.
 - John kriegt eine Zeitung von Patrick
- Sprachunabhängigkeit
- Verminderung von Mehrdeutigkeiten
- Ziehen von Schlussfolgerungen wird erleichtert, nicht jeder Einzelfall muss formuliert werden

Nachteile

- hoher Aufwand
- ungewöhnliche Darstellung
- Nuancen der natürlichen Sprache gehen verloren, z.B. bestehen zwischen
 - überreichen
 - geben
 - erhalten
 - bekommen
 - kriegenfeine Unterschiede
- syntaktische Informationen gehen verloren
- es gibt auch keine Evidenz dafür, dass Menschen ihr Wissen in irgendeiner aus semantischen Primitiven aufgebauten Form speichern

5.3.3 Semantische Netze am Beispiel EVAR

EVAR

- EVAR steht für Erkennen, Verstehen, Antworten, Rückfragen
- ist ein Sprachverarbeitungssystem für einen natürlichsprachlichen Auskunftsdialog über die IC Züge der DB
 - kontinuierlich gesprochene Sprache in Telefonqualität
 - Sprecherunabhängig
- zur Wissensdarstellung wird ein semantisches Netzwerk benutzt

Module

- Worterkennung
- Syntax
- Semantik
- Pragmatik
- Dialog
- Prosodiemodul

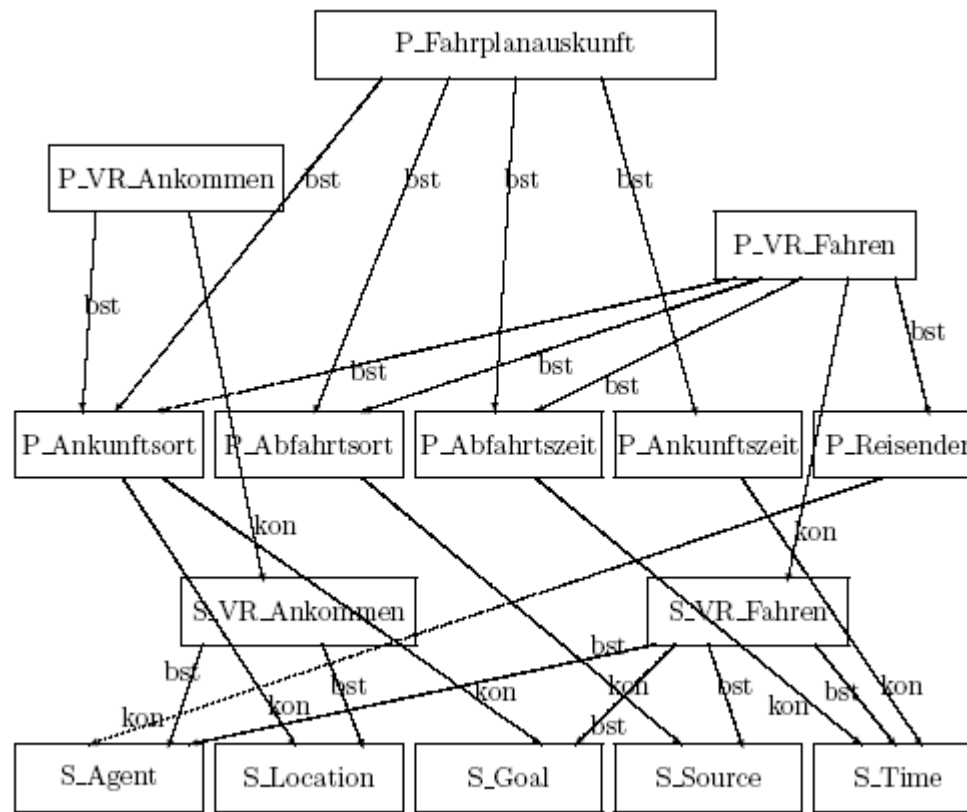
Semantisches Netz

- **Spezialisierungen (spez)**
 - erlauben Verfeinerungen von allgemeinen Konzepten (Knoten des Netzes), wobei die Vererbung von Eigenschaften (Attributen, Bestandteilen, Konkretisierungen) möglich ist
- **Bestandteile (bst)**
 - Beziehungen zwischen einem Konzept und anderen Konzepten der Wissensbasis (semantisches Netz), aus denen sich dieses Konzept zusammensetzt
- **Konkretisierungen (kon)**
 - ermöglichen die Verbindung von Wissen aus höheren Abstraktionsebenen mit niedrigeren Ebenen
 - diese Kanten dienen während der Analyse zur Interpretation (zielgesteuert oder datengesteuert) von signalnahen Konzepten durch abstraktere Konzepte

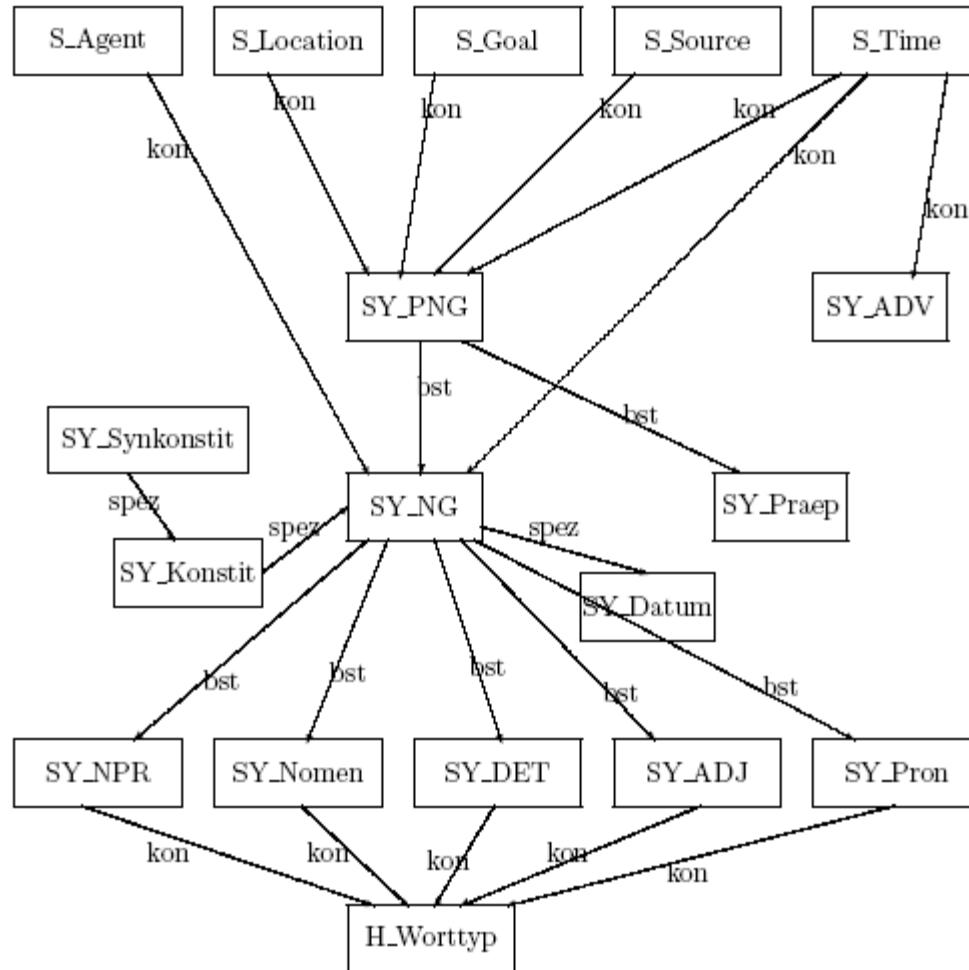
Semantisches Netz

- P – Pragmatik
- S – Semantik
- SY – Syntax
- VR – Verbrahen
- H – Hypothesenebene
- PNG – Präpositionalgruppe
- NG – Nominalgruppe
- NPR – Eigennamen

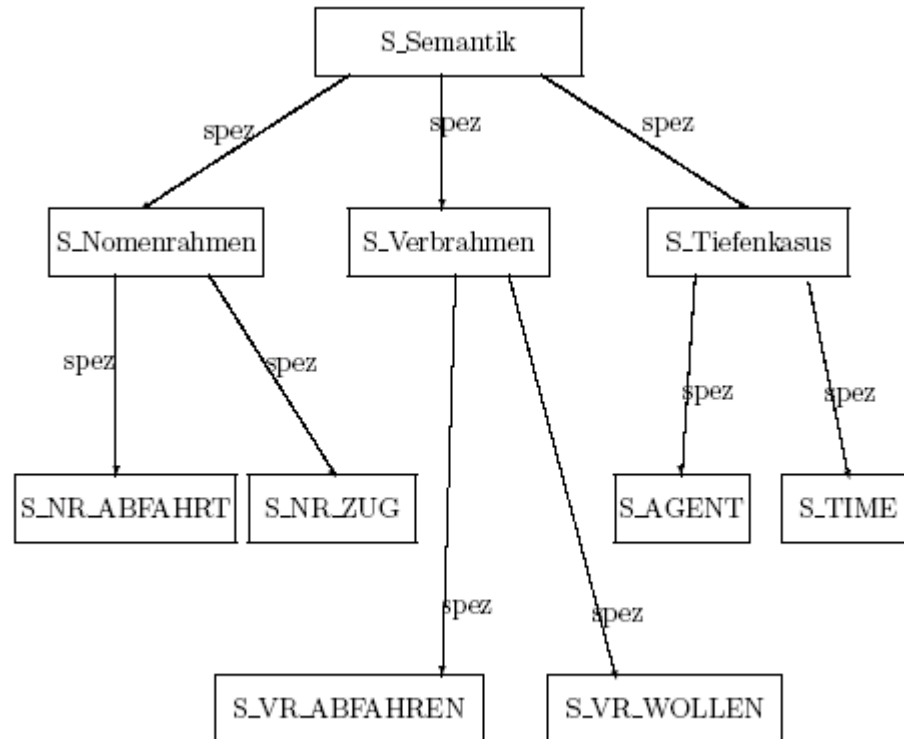
EVAR



EVAR

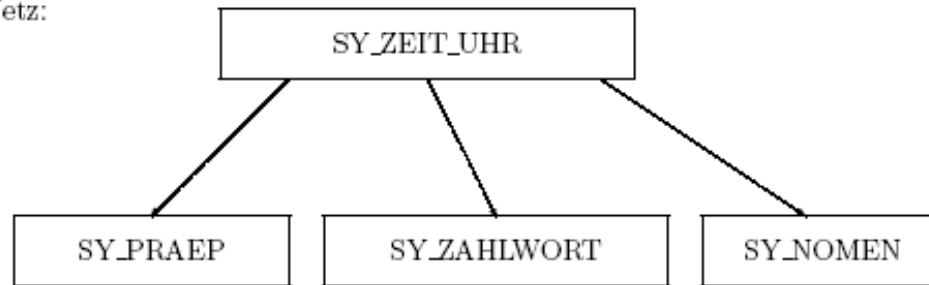


EVAR

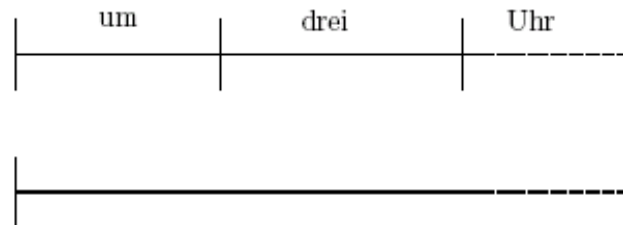


Abarbeitung des semantischen Netzes

Netz:



Hypothesenmenge:

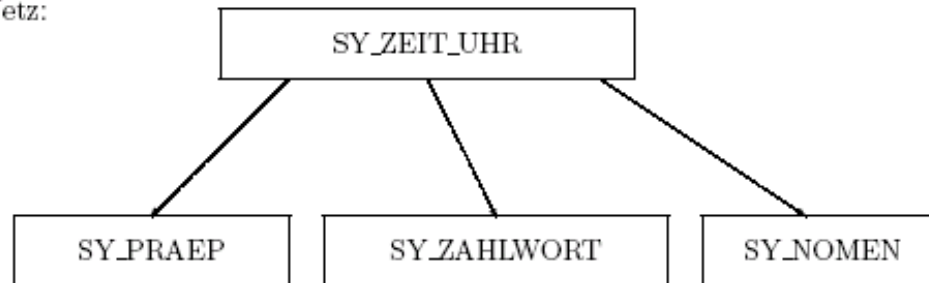


Datengetriebene Analyse (I – Generierung von Instanzen):

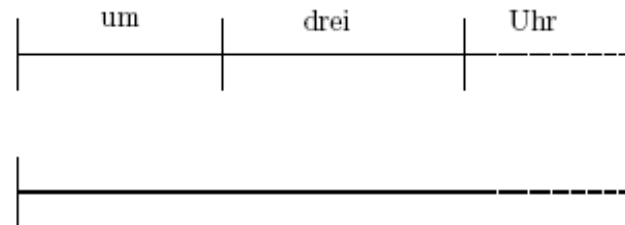
1. I1(SY_PRAEP) → um
2. I1(SY_ZAHLWORT) → drei
3. I1(SY_NOMEN) → Uhr
4. I1(SY_ZEIT_UHR) → um drei
5. I2(SY_ZEIT_UHR) → um drei Uhr

Abarbeitung des semantischen Netzes

Netz:



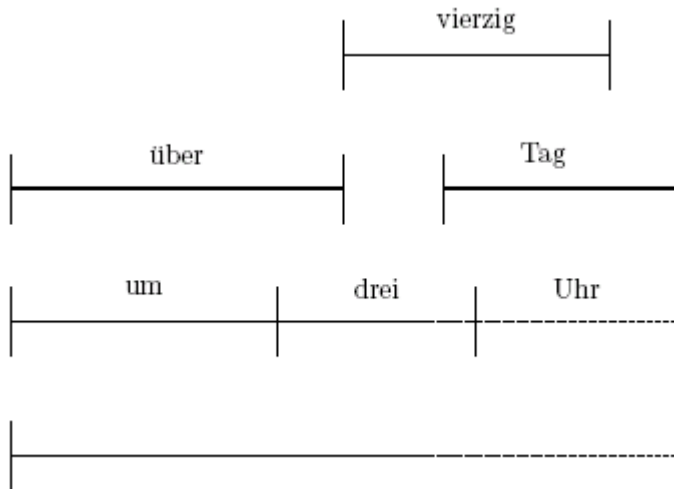
Hypothesenmenge:



Erwartungsgesteuerte Analyse (Q - Erzeugung von Konzepten):

1. Q1(SY_ZEIT_UHR)
2. Q1(SY_PRAEP)
3. I1(SY_PRAEP) → um
4. Q1(SY_ZAHLWORT)
5. I1(SY_ZAHLWORT) → drei
6. I1(SY_ZEIT_UHR) → um drei
7. Q1(SY_NOMEN)
8. I1(SY_NOMEN) → Uhr
9. I2(SY_ZEIT_UHR) → um drei Uhr

Abarbeitung des semantischen Netzes



Sie führt zu folgender datengetriebener Analyse:

1. I1(SY_PRAEP) → um
2. I2(SY_PRAEP) → über
3. I1(SY_ZAHLWORT) → drei
4. I2(SY_ZAHLWORT) → vierzig
5. I1(SY_NOMEN) → Uhr
6. I2(SY_NOMEN) → Tag
7. I1(SY_ZEIT_UHR) → um drei
8. I2(SY_ZEIT_UHR) → um vierzig
9. I3(SY_ZEIT_UHR) → über drei
10. I4(SY_ZEIT_UHR) → über vierzig
11. I5(SY_ZEIT_UHR) → um drei Uhr
12. I6(SY_ZEIT_UHR) → um drei Tag