Natürliche Sprachen

Sprachfähigkeit - die erstaunliche Fähigkeit, anhand endlicher (verrauschter) sprachlicher Daten ein System aufzubauen, das unendlich viele sprachliche Daten erzeugen und verstehen kann.
Die Sprachfähigkeit bezieht sich auf alle natürlichen Sprachen.
Auch debile Menschen lernen eine Sprache - dabei ist eine natürliche Sprache äußerst komplex!

Sprache - ein System, das unendlich viele, auch bisher nicht gehörte Äußerungen versteht und produziert, wobei die Äußerungen der Kommunikation dienen.
Kompetenz
Kinder lernen eine beliebige Sprache - dabei sind die Sprachen sehr unterschiedlich.

Sprechen - Produktion einer Äußerung (in räumlich-zeitlicher Gleichheit von Hörer und Sprecherin).
Performanz
Die Sprache wird angewandt, um zu kommunizieren. Erstaunlicherweise gelingt dies, obwohl Äußerungen sehr mehrdeutig sind.

Sprachebenen

Syntax:
Struktur von Äußerungen, definiert durch konstruktive Definition einer Strukturbeschreibung und Angabe einer Funktion, die einem beliebigen Satz eine Strukturbeschreibung zuordnet.
Wohlgeformtheit

Semantik:
Bedeutung und Sinn von Äußerungen
Wahrheit, propositionaler Gehalt

Pragmatik:
Bezug zur Situation (Hörer, Sprecherin, Umgebung, räumlich-zeitliche Beziehungen, Handlungsfolgen)
Glücken sprachlicher Handlungen

Lexikalische Mehrdeutigkeiten

Mutter
Schloß
Schlag

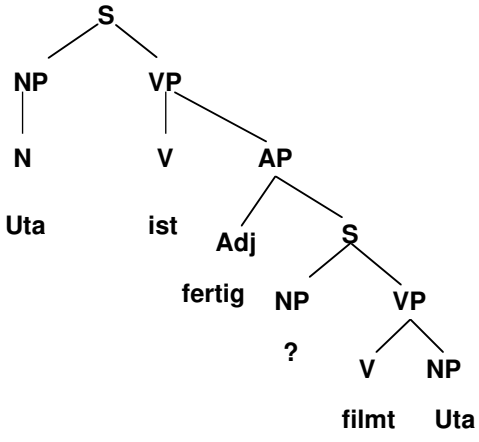
"Ich kann die Mutter nicht finden!"

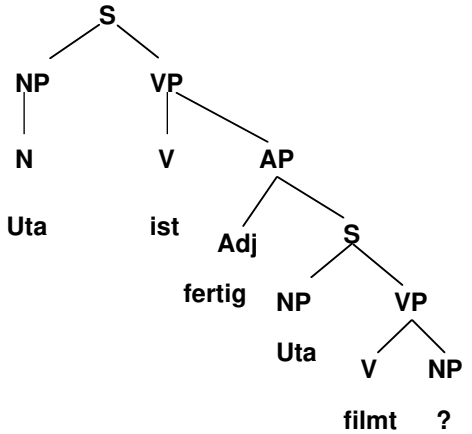
"Der Schlüssel paßt nicht ins Schloß!"

"Die beiden sind vom selben Schlag."

Syntaktische Mehrdeutigkeiten

"Uta ist fertig zum Filmen."





"Ich sah Kraniche auf dem Flug nach New York." **"Die Milch trinkt die Katze."**
"Er schneidet die Salami mit Pfeffer."
"Ich sah das Mädchen mit dem Fernrohr."

Semantische Mehrdeutigkeiten

"Hans will eine Künstlerin heiraten."

∃ X | künstlerin (X) & möchte (hans, heiraten (hans, X)) oder
möchte (hans, (∃ X | künstlerin (X) & heiraten (hans, X)))

Die Frau, die Hans heiraten möchte, ist eine bereits bestimmte. Diese bestimmte Frau hat die Eigenschaft, Künstlerin zu sein.

Die Frau, die Hans heiraten möchte, muß Künstlerin sein - alles andere, insbesondere ihre Identität, steht noch nicht fest.

Pragmatische Mehrdeutigkeiten

"Sie ließ die Schüssel auf die Terrasse fallen und zerbrach sie."
die Schüssel? die Terrasse?

"Bin gleich zurück!"
Wann?

"Wissen Sie, wie spät es ist?"
- "ja"
- "entschuldigen Sie, daß ich jetzt noch anrufe, aber..."
- "viertel vor zwölf."

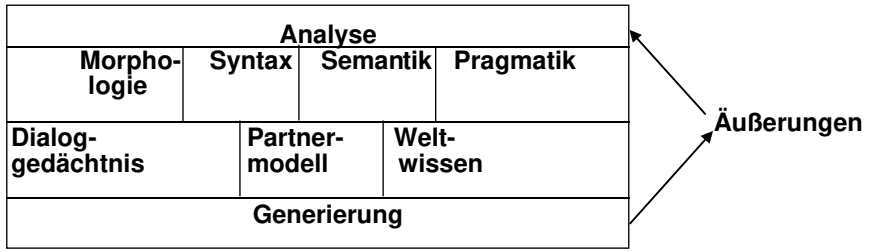
ganz eindeutig aufgrund konversationeller Implikatur:
"Der Kapitän war heute nüchtern." (Nachfolgeadjektive)

Sprachliche Einheiten

Text, Rede
Handlungen
Satz
Phrasen, Kategorien
Wort
Phonem, Morphem

Natürlichsprachliche Systeme

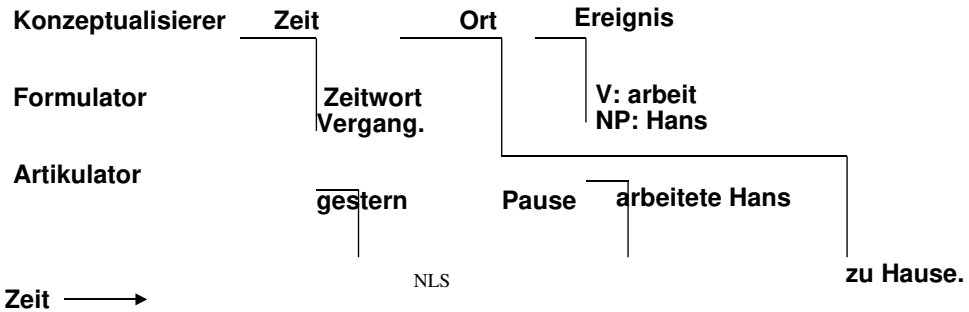
- **kognitiv**
- **anwendungsorientiert**
- **als Beschreibung**



Kognitionsorientierte Arbeiten

Wie werden Äußerungen produziert oder verstanden?
Levelt (1989) Generierung von Sätzen
DeSmedt (1990) Implementierung des Modells von Levelt
Randbedingungen:
2 - 3 Wörter pro Sekunde (ca. 15 Phoneme pro Sekunde)
selten und nur an speziellen Stellen Pausen

* Gestern zu Hause arbeitete Hans.
Gestern im Kino traf ich Ruth.



Module

Konzeptualisierer produziert, WAS gesagt werden soll:
Intention --> Plan der Äußerung
Plan --> lineare Abfolge gemäß Perspektive und Betonung
Selbstbeobachtung --> Korrektur

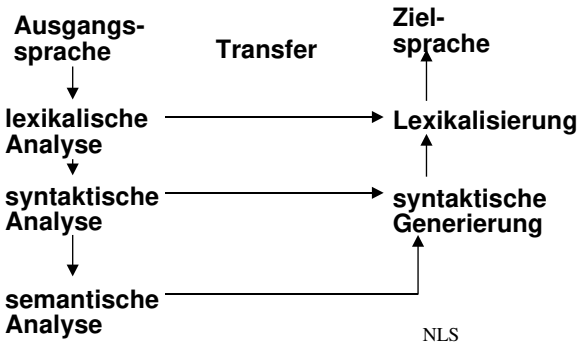
Formulator produziert, WIE es gesagt werden soll (Artikulationsplan):
Lexikon --> grammatische Struktur, phonologische Struktur

Sem. Syn.

Artikulator produziert den gesprochenen Satz:
Atmung,
Sprechwerkzeuge

Anwendungen von NLS
Maschinelle Übersetzung

EUROTRA
kanadischer Wetterdienst
VerbMobil

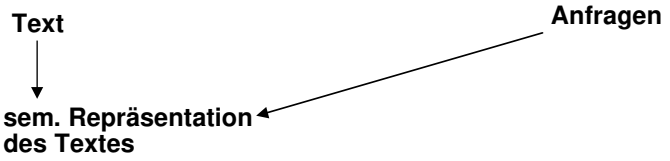


Übersetzungshilfen

Schreibunterstützung

Rechtschreibkorrektur
Stilkorrektur
Vorschläge zur Verbesserung
sachliche Prüfung des Textes

Textverstehen



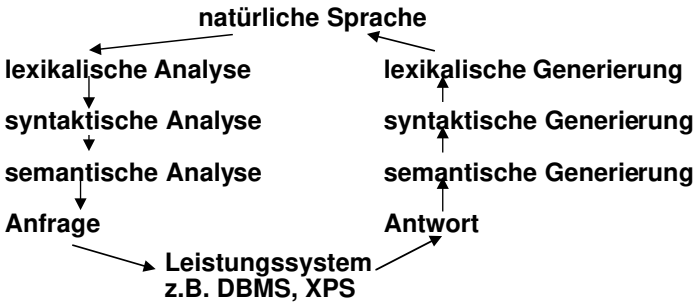
Arbeiten der Schule
von Roger Schank

LILOG

Lernen aus Texten: wit
(Ulrich Reimer)

Zugangssysteme

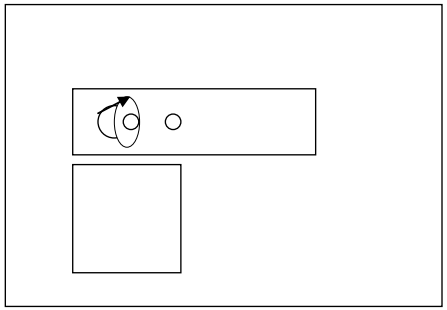
Frage-Antwort-Systeme z.B. PLIDIS, Produkt: Q/A, F/A
Dialogsysteme z.B. HAM-RPM
Beratungssysteme z.B. HAM-ANS



Problem: Sprachgrenzen sind den Benutzern unklar.
"Gib Proben Lauxmann!"

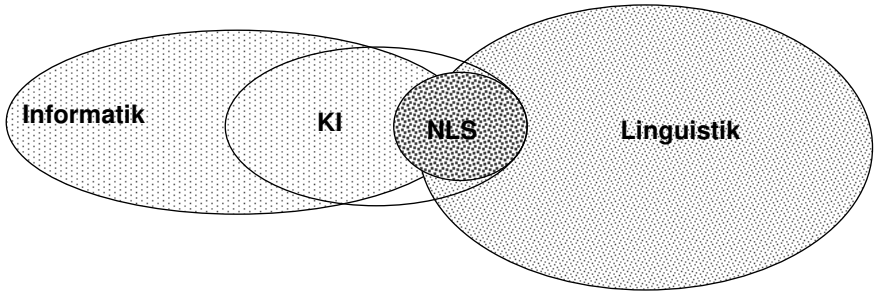
Multimediale Präsentation

Den linken Knopf nach rechts drehen.



WIP (DFKI Saarbrücken)

Indirekte Anwendung



SYS soll nach Regeln menschlicher Kommunikation gestaltet werden!
SYS verändert Kommunikation in Firmen.
e-mail, networks bieten neue Kommunikationsarten.
Blick in die Linguistik ist notwendig!
Linguistik bietet bereits formale Theorie und Methodik
Psychologie hingegen: Methodik zur Sammlung von Daten, die Hypothesen über menschl. Verhalten stützen/verwerfen.

Lexikalische Analyse

lexikalische Analyse im engeren Sinne Eingabesequenz ---> Wörter
morphologische Analyse Wörter ---> bedeutungstragende Einheiten mit Merkmalen

Nicht einmal die Segmentierung in Wörter ist immer leicht!
Ein Junge namens To liebte Pferde. Seine Mutter sagte:
"Da! Blumen, To, Pferde, alles, was dein Herz begehrt! "
"Da! Blumentopferde!"
"Trave, münde!" bei gesprochener Sprache: ...
Komposita:
Blumen-topf-erde, Jäger-schnitzel, Schweine-schnitzel, Esels-ohren,
Morpheme:
Blume-n-topf-erde, Schwein-e-schnitzel, Esel-s-ohr-en
Plural **Genitiv**
Verantwortung-s-bewußt-sein, Bahnhof-s-halle, Schiff- ahrt/Schif-fahrt
Fugen-s **Verbot 3maliger Wiederholung**
30% der Wörter in Zeitungen sind Namen oder spontane Komposita.

Lexikon

Vollformenlexikon:
alle zusammengesetzten Wörter
alle flektierten Wörter
spielen, spiele, spielst, spielt, gespielt, spielte
Man spart die morphologische Analyse, aber man erhält ein sehr großes Lexikon!
ansonsten:
Morpheme mit syntaktischer und semantischer Information,
bei gesprochener Sprache auch phonologische Information
auch so etwa 100 000 Einträge

Syntaktische Information

Grammatik oder Lexikon?
intransitives Verb: VP ohne Objekt Z.B.: fliegt, rennt, arbeitet...

Grammatik:
s --> np (nominativ, Genus, Person, Numerus), vp (Person, Numerus).
vp (Person, Numerus) --> v(vi, Person, Numerus).
Vollformenlexikon:
v(vi, 3,singular) -->[fliegt].
Subjekt ist NP im Kasus Nominativ, hier angegeben durch Grammatikregel.
Übereinstimmung zwischen Person und Numerus durch Unifikation.
Grammatik:
s --> subjekt(SKategorie, Kasus, Genus, Person, Numerus),
vp (SKategorie, Kasus, VKategorie, OKategorie, Person, Numerus).

Vollformenlexikon:
vp (np, nom, vi, 0, 3, singular) --> [fliegt].
Subjekt ist NP im Nominativ, hier angegeben im Lexikon.

Makros bei Vollformenlexika

vp (np, nom, vi, 0, 3, singular) --> [fliegt].
vp (np, nom, vi, 0, 3, singular) --> [rennt].
vp (np, nom, vi, 0, 3, singular) --> [arbeitet].
. . .

Statt die immer gleiche Information für intransitive Verben bei jedem Lexikoneintrag anzugeben,

wird im Lexikoneintrag nur noch ein Makro für intransitive Verben angegeben,

das durch eine Regel expandiert wird.

Lexikon:
wort (macro_vi, 3, singular) --> [fliegt].
wort (macro_vi, 3, singular) --> [rennt].
wort (macro_vi, 3, singular) --> [arbeitet].

Makro und Expansion:
vp (np, nom, vi, 0, Person, Numerus) --> wort (macro_vi, Person, Numerus).

Grammatik:
s --> subjekt(SKategorie, Kasus, Genus, Person, Numerus),
vp (SKategorie, Kasus, VKategorie, OKategorie, Person, Numerus).

Makros für die Flexion

Lexikoneintrag im Vollformenlexikon:
vp (macro_vi, 1, singular) --> [fliege].
vp (macro_vi, 2, singular) --> [fliegst].
vp (macro_vi, 3, singular) --> [fliegt].
vp (macro_vi, 1, plural) --> [fliegen].
vp (macro_vi, 2, plural) --> [fliegt].
vp (macro_vi, 3, plural) --> [fliegen].

Makro für Flexion regulärer Verben:
suffix(e, 1, singular) --> [e].
suffix(st,2,singular) --> [s, t].
suffix(t, 3, singular) --> [t].
suffix(en, 1, plural) --> [e, n].
suffix(t, 2, plural) --> [t].
suffix(en, 3, plural) --> [e, n].

Lexikon:
stamm(flieg, vi) --> [f, l, i, e, g].

Makro für Verbkategorien:
macro_kat(np, nom, vi, 0,0).
macro_kat(np, nom, vt, np, akk).
macro_kat(np, nom, vd, np, dat).

morphologische Analyse:
wort(Stamm, VKat, Suffix, Pers, Num) -->
stamm(Stamm, VKat),
suffix(Suffix, Pers, Num).

Grammatik:
vp(SKat, Kas, VKat, OKat, OKas, Pers, Num) -->
wort (Stamm, VKat, Suffix, Pers, Num),
{macro_kat(SKat, Kas, VKat, OKat, OKas)}.

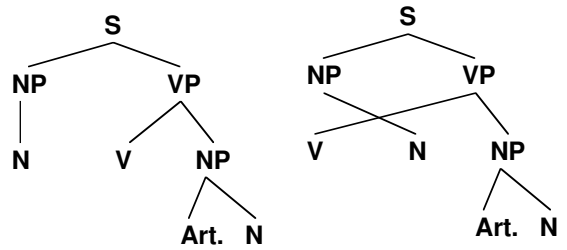
Syntaktische Analyse

Folge von Lexemen mit Informationen ---> syntaktische Struktur oder *
Unendlichkeit von Sätzen der natürlichen Sprachen
Zerlegung eines Satzes in Phrasen (Konsituenten),
die miteinander verbunden sind.
Phrasen lassen sich gemeinsam verschieben.
Am Morgen hole ich Brötchen.
Ich hole am Morgen Brötchen.
Brötchen hole ich am Morgen.
Phrasen werden durch grammatische Kategorien bezeichnet: NP, VP, N, Präp....

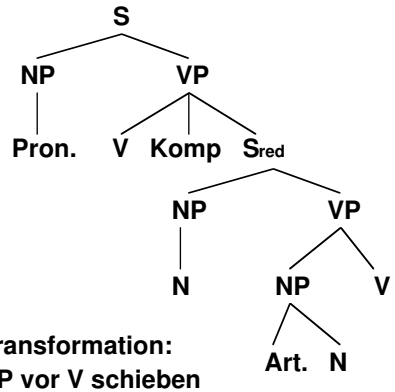
Verbindungen zwischen Phrasen:
lineare Abfolge (linear precedence) z.B.: Verbzweitstellung m Deutschen!
Abhängigkeit (immediate dominance) z.B.: die von VP abhängige NP erhält den durch das Verb bestimmten Kasus.

Transformationen

Anfangs-P-Marker:	Fragesatz:	eingebetteter Satz:
Eva kauft ein Buch.	Kauft Eva ein Buch?	Ich glaube, daß Eva ein Buch kauft.



Transformation: V an den Anfang schieben



Transformationsgrammatik:
eine Basisstruktur, die schrittweise in verschiedene Oberflächenstrukturen überführt wird.

Transformation:
NP vor V schieben

Parsing

Ein Parser ist ein Algorithmus, der einen Satz als Eingabe nimmt und entscheidet, ob er bezüglich einer gegebenen Grammatik wohlgeformt ist und ihm - falls ja - eine Strukturbeschreibung zuordnet.

Top down: beginnend mit S werden die Grammatikregeln angewandt, bis die terminalen Symbole der Grammatik erreicht sind, die die Eingabe des Parsers bilden.

Nichtdeterministisch: wenn es mehrere Regeln mit demselben Kopf gibt, wird vielleicht zunächst eine angewandt, die nicht zum Ziel führt.
Backtracking

Prolog kann direkt als Parser verwendet werden!

Grammatik

Eine Grammatik beschreibt wohlgeformte syntaktische Strukturen.
Damit definiert sie eine Sprache.
Ein Grammatikformalismus beschreibt eine Klasse von Sprachen.
Ein Parser verwendet eine Grammatik, ein Lexikon und einen Eingabesatz zur Erstellung einer syntaktischen Struktur für diesen Satz.
Ein Parserformalismus kann von einer Klasse von Algorithmen benutzt werden.

Bewertung einer Grammatik:

beobachtungsadäquat:
nicht übergeneriert: nicht wohlgeformte Sätze werden zurückgewiesen (korrekt)
nicht untergeneriert: wohlgeformte Sätze werden nicht zurückgewiesen (vollständig)
beschreibungsadäquat:
Grammatik stützt sich auf ein System von Verfahren, das z.B. mit psychologischen und neurologischen Erkenntnissen übereinstimmt.
erklärungsadäquat: Grammatikformalismus kann alle Grammatiken generieren, erklärt also das Erlernen von Sprache.

DCG

Definite Clause Grammar: stets nur ein Kopfliteral bzw. jede Regel expandiert nur ein Nichtterminal.
Notation:
s --> np, vp.
det --> [die].

Aufruf:
:- phrase([die, katze, jagt, die ratte], []).

Prolog übersetzt die DCG-Notation in das interne Format.
Quintus-Prolog überführt
det --> [die]. in
det (L1, L) :- 'C' (L1, die, L).
wobei 'C' ein Prädikat ist, das ein Element vor eine Liste hängt bzw. eine Liste in das erste Element und den Rest zerlegt.
So wird die und L zu L1 = [katze| L].

Grammatik in Prolog

Grammatik	Prolog-Notation
S --> NP VP	s(L1,L):- np(L1,L2),vp(L2,L).
NP --> Det N	np(L1,L):- det(L1,L2),n(L2,L).
Det --> die	det([die|L],L).
N --> Katze	n([katze|L],L).
N --> Ratte	n([ratte|L],L).
VP --> V NP	vp(L1,L):- v(L1,L2),np(L2,L).
V --> jagt	v([jagt|L],L).

Dabei ist L1 die aktuell zu verarbeitende Eingabe,
L2 der durch eine Konstituente erkannte Teil dieser aktuellen Eingabe,
L der noch zu verarbeitende Rest der Eingabe.

Beispiel

Beweisziele	zum Beweis verwendete Klauseln:
:- s([die, katze, jagt, die ratte],[]).	s (L1, L):- np(L1, L2), vp(L2,L).
np([die,katze,jagt,die ratte],_1),vp(_1,[])	np(L1,L) :- det(L1,L2), n(L2,L).
det([die,katze,jagt,die,ratte],_2),n(_2,_1),vp(_1,[])	det([die|L],L).
det([die|katze,jagt,die,ratte],[katze,jagt,die,ratte]) erkannt!	
n([katze,jagt,die,ratte],_1),vp(_1,[])	n([katze|L],L).
n([katze|jagt,die,ratte],[jagt,die,ratte]) erkannt!	vp(L1,L) :- v(L1,L2), np(L2,L).
vp([jagt,die,ratte],[])	v([jagt|L],L).
v([jagt,die,ratte],_3),np(_3,[])	
v([jagt|die,ratte],[die,ratte]) erkannt!	np(L1,L) :- det(L1,L2),n(L2,L).
np([die,ratte],[])	det([die|L],L).
det([die,ratte],_4),n(_4,[])	
det([die|ratte],[ratte]) erkannt!	n([katze|L],L). nicht erkannt!
n([ratte|[]],[])	n([ratte|L], L).
n([ratte|[]],[])	
n([ratte|[]],[]) erkannt! - Eingabe wohlgeformt!	

Wie baut man eine Grammatik, die für einen Ausschnitt beobachtungsadäquat ist?

- **Was für Sätze kommen in der Anwendung vor? - Corpus, Satzschemata**
- **Wie lassen sie sich beschreiben? - syntaktische und semantische Merkmale**
- **Schreiben von Grammatikregeln**
- **Schreiben von Lexikoneinträgen**
- **Testen von Grammatik und Lexikon anhand des Corpus - übergeneriert? - untergeneriert?**

Beispiel

Sachbereich Tiere:
hauttyp: fell, federn, haut, schuppen
aussehen: braun, gestreift, gefleckt, silbrig, hautfarben
nahrung: fleisch, pflanze
nachkommen: gebaehren, eierlegen, laichen
aufzucht: saeugen, fuettern, vernachlaessigen
bewegung: steht, fliegt, schwimmt
ist_ein: saeugetier, vogel, fisch
Beschreibungen von:
Leopard, Tiger,Kuh, Mensch, Strauß, Papagei, Amsel, Karpfen, Hecht

Anwendung:
Benutzer stellt Fragen und erhält Antworten.

Satztypen:
Entscheidungsfragen Antwort: "ja" oder "nein"
Objektfragen Antwort Aussagesatz

Aussagesätze

Der Leopard hat Fell.	Die Amsel hat Federn.	Der Hecht hat Schuppen.
Der Tiger ist gestreift.	Der Hecht ist silbrig.	Leoparden sind gefleckt.
Die Kuh frißt Pflanzen.	Tiger fressen Fleisch.	Menschen essen Pflanzen und Fleisch.
Tiger gebähren ihre Jungen.	Ein Papagei legt Eier.	Der Karpfen laicht.
Der Leopard säugt seine Jungen.	Die Amsel füttert ihre Jungen.	Hechte vernachlässigen die Jungen.
Eine Amsel fliegt.	Ein Mensch steht.	Karpfen schwimmen.
Der Karpfen ist ein Fisch.	Kühe sind Säugetiere.	Ein Strauß ist ein Vogel. ...

Ersetzungsprobe und systematische Erzeugung von Sätzen:
Bei einem Satz für jedes Wort versuchen, jedes andere Wort einzusetzen.
Kann ein neues Wort ein Wort des Ausgangssatzes ersetzen, so wird es der Menge der möglichen Wörter an dieser Position hinzugefügt.
Das Ergebnis ist ein Satzschema, bei dem alle Elemente einer Menge mit allen Elementen anderer Mengen kombinierbar sind.

{Ein, Der} {Leopard, Tiger, Mensch, Strauss, Papagei, Karpfen, Hecht, Vogel, Fisch, Fleischfresser, Pflanzenfresser} hat {Fell, Federn, Haut, Schuppen}.

Kategorien

Satzschema:
{Ein, Der} {Leopard, Tiger, Mensch, Strauss, Papagei, Karpfen, Hecht, Vogel, Fisch, Fleischfresser, Pflanzenfresser} hat {Fell, Federn, Haut, Schuppen}.

Einige Ersetzungen gehen nicht:
*{Ein, Der} {Kuh, Amsel, Säugetier} hat {Fell, Federn, Haut, Schuppen}.
*{Der, Tiger, ist, säugt, seine, Eier, Jungen, ...} {Leopard, Tiger, Mensch, Strauss, ...} hat {Fell, Federn, Haut, Schuppen}.
Nur kleine Abweichungen sind interessant. Wir verwenden linguistisches Wissen:
{Ein, Der, Die, Das, _ } sind Artikel,
{Leopard, Tiger, Mensch, Strauss, Papagei, Karpfen, Hecht, Vogel, Fisch, Fleischfresser, Pflanzenfresse, Fell, Federn, Haut, Schuppen} sind Substantive.
Zusätzliche Satzschemata:
{Eine, Die} {Kuh, Amsel} hat {Fell, Federn, Haut, Schuppen}.
{Ein, Das} {Säugetier} hat {Fell, Federn, Haut, Schuppen}.
{Die, _ } {Leoparden, Tiger, Menschen, Strausse, Papageien, Karpfen, Hechte, Vögel, Fische, Fleischfresser, Pflanzenfresser} haben {Fell, Federn, Haut, Schuppen}.
Art, N1, V, N2.

NP Merkmale und Grammatikregeln

Merkmale bestimmen die möglichen Kombinationen von Artikel und Substantiv:
Genus: masculin, feminin, neutrum
Det: best, unbest
Numerus: sing, plur
Kasus: nom, gen, dat, akk

s --> np (nom, Numerus), vp (Numerus,3).
np(Kasus, Numerus) -->
art(Det, Kasus, Genus, Numerus),
n (Kasus, Genus, Numerus)

Wenn unser Sprachausschnitt keine Pronomen enthält, können wir die Person gleich für die VP festlegen: 3.

Numerus von NP und VP müssen über- einstimmen.
Numerus, Genus und Kasus von Art und N müssen übereinstimmen.

VP Merkmale und Grammatikregeln

{<np(nom, plur, klasse)>}{säugen, füttern, vernachlässigen, ernähren, gebähren} {die, ihre} Jungen.
{<np(nom, plur, klasse)>}legen Eier.
{<np(nom, plur, klasse)>}{fressen, essen} {Fleisch, Pflanzen}.
{<np(nom, sing, klasse)>}ist ein {Tier, Säugetier, Vogel, Fisch, Pflanzenfresser, Fleischfresser}.
Merkmale bestimmen Zusammenhang von V und NP in der VP.
Verbtyp: vt, vi
Gattung: hauttyp, nachkommen, nahrung, klasse
Eigenart: belebt, unbelebt
s --> np (nom, Numerus, SGattung, belebt), vp (Verbtyp, Numerus, 3).
vp(vt, Numerus, 3) --> v(vt, Numerus, 3, SGattung, akk, OGattung, OEigenart), np(akk, Num, OGattung, OEigenart).
vp(vi, Numerus, 3) --> v(vi, Numerus,3, _, _, _).

Fragesätze

Wenn wir keine Transformationen schreiben wollen, geben wir für jeden Satztyp eine eigene Regel an:
s --> efs. s--> ofs. s--> as. Die bisherigen Regeln beziehen sich auf as.
efs --> v(vt, Num, 3, SGatt, akk, OGatt, OEigen), np(nom, Num, SGatt, SEigen), np(akk, Num2, OGatt, OEigen).
efs --> v(vi, Num, 3, _, _, _), np(nom, Num, SGatt, SEigen).
Bei den Objektfragen fällt auf, daß das nicht vorhandene Objekt dennoch Einfluß auf das Fragewort hat:
*Was säugen Menschen?
Wen säugen Menschen.
Wir müssen also die Eigenart des Objektes als Merkmal beim Fragewort angeben.

vorläufige Grammatik

s --> efs. s --> as.
efs --> v(vi, Num, 3,_,_,_), np(nom, Num, SGatt, SEigen).
efs --> v(vt, Num, 3, SGatt, akk, OGatt, OEigen), np(nom, Num, SGatt, SEigen), np(akk, Num2, OGatt, OEigen).
as --> np(nom, Num, SGatt, belebt), vp(VT, Num, 3).

np(Kasus, ohneart, Gatt, Eigen) -->
n(Kasus, Num, Genus, Gatt, Eigen).
np(Kasus, Num, Gatt, Eigen) -->
art(Det, Num, Kasus, Genus),
n(Kasus, Num, Genus, Gatt, Eigen).

vp(vi, Num, 3) --> v(vi,Num,3,_,_,_).
vp(vt, Num, 3) -->
v(vt, Num, 3, SGatt, akk, OGatt, OEigen),
np(akk, Num2, OGatt, OEigen).

Vollformenlexikon dazu

n(akk, ohneart, n, hauttyp, unbelebt) --> "fell". ...
n(nom, sing, m, tier, belebt) --> "leopard". ...
art(best, sing, nom, m) --> "der". ...
v(vt, sing, 3, tier, akk, hauttyp, unbelebt) -->"hat". ...
v(vi, sing, 3, tier, _,_,_) --> "laicht". ...

Aufbau von Strukturen

Bisher haben wir als Ergebnis von der DCG lediglich YES oder NO erhalten für wohlgeformte und nicht wohlgeformte Sätze.
Wir wollen aber eine Strukturbeschreibung erzeugen.

```
s(s(NP,VP)) --> np(NP, kasus:nom, genus:Gen, numerus:Num, gattung:G,
    seigenart: E),  vp(VP, verbtyp:vi, numerus: Num, seigenart:E).
np(np(Art,N), kasus:nom, genus:Gen, numerus:Num, gattung:G, seigenart: E ) -->
     art(Art,kasus: nom, genus:Gen, numerus: Num, det: Det, sem:_),  n(N, kasus:
    nom, genus:Gen, numerus: Num, gattung: G, seigenart:E, sem:_).
vp(vp(V), verbtyp:vi, numerus: Num, seigenart:E ) -->   v(V,verbtyp:vi, person:3,
    numerus:Num, eigenart:E, oeigenart:_, semtyp:_,  sem:_).
art(art(die), kasus:nom, genus:f, numerus:plur, det:best, sem:alle) --> [die].
n(n(amseln),kasus:nom, genus:f, numerus: plur, gattung: tier, seigenart:
    belebt,sem:amsel) --> [amseln] .
v(v(fliegen),verbtyp: vi, person:3, numerus: plur, seigenart: belebt,
    oeigenart:unfragbar, semtyp: bewegung, sem:fliegen) --> [fliegen].
```


Beispiel

Auf die Anfrage

```
phrase(s(S),[die, amseln, fliegen])
```

bekommt man nun nicht nur YES, sondern

```
s(np(art(die), n(amseln)), vp(v(fliegen)))
```

als Lösung.

Testen

Einige mögliche Sätze sind nicht abgedeckt. Die Grammatik ist untergeneriert.
Ja. Nein. sind Aussagesätze.
Der Mensch ist ein Säugetier. ist ein Aussagesatz, und zwar ein Kopulasatz.
Bei Kopulasätzen ist es egal, welche NP vorn und welche hinten steht. Beide NP sind im Nominativ.
Das Verb "ist" verbindet zwei gleichwertige Beschreibungen.
Der Tiger ist gestreift.
Eigentlich müssen beide Verwendungen von "ist" gemeinsam behandelt werden.
Der Einfachheit halber trennen wir hier die Ergänzung durch eine NP von der Ergänzung durch ein Adjektiv und sehen verschiedene Fragen vor:
Wie sieht der Tiger aus?
Was ist der Tiger?
ACHTUNG!
Solche Vereinfachungen kennen die Benutzer nicht!

Semantik

Was bedeuten die Sätze, deren Struktur wir so schön herausgefunden haben?
Intensionale Bedeutung:
die Bedingungen (in der Welt) unter denen der Satz wahr ist.
Ablehnung semantisch nicht wohlgeformter Sätze, d.h. solcher Sätze, die nicht wahr sein können.
Der Berg sieht das Kind.
Auflösung syntaktischer Ambiguitäten.
Das Kind sieht den Berg mit dem Fernglas.

Bedeutung als Referenz

Das Denotat ist bei Wörtern ein Objekt oder eine Menge von Objekten, für die die durch das Wort angegebene Eigenschaft gilt.

rennt **bedeutet die Menge der Rennenden**

Der Bundeskanzler von Deutschland 1995 **bedeutet die Menge der Personen, die 1995 Bundeskanzler von Deutschland sind.**

Der Bundeskanzler rennt **bedeutet den Schnitt dieser beiden Mengen.**

Hans **bedeutet die Menge der Personen mit dem Namen Hans?**

Semantische Repräsentationssprache

Die Bedeutung eines Satzes soll in einer Sprache ausgedrückt werden.
Diese Sprache soll gleichbedeutenden Sätzen der natürlichen Sprache genau einen Ausdruck der semantischen Repräsentationssprache zuordnen.
Sie wird verwendet, um die Wahrheit der Bedeutung eines Satzes gegenüber einer (möglichen) Welt festzustellen.
Praktischerweise kann als mögliche Welt eine Datenbank angenommen werden und die Bedeutung eines Satzes wird über der Datenbank ausgewertet.
Natürlichsprachliche Systeme gehen meist zweistufig vor:

synt. Struktur ---> semantische Analyse --> SRS-Ausdruck

SRS-Ausdruck ---> Anfrageerzeugung -------> Anfrage (in SQL, Prolog, ...)

Anfrage ----------> Auswertung ----------> Antwort

Vereinfachung

Wir können hier vereinfacht ein kleines Produktionensystem mit Aussagen über Tiere als Weltzustand wählen.
Dieses Produktionensystem hat eine Anfragesprache DBQ.

Vereinfacht erzeugen wir mit einem Schritt aus der syntaktischen Struktur eine Anfrage in DBQ.

DBQ

Syntax von DBQ-Ausdrücken:
dbqexp --> referent; aussage.
dbqexp --> connective (dbqexp, dbqexp).
referent --> (quantor menge).
referent --> menge.
referent --> variable.
aussage --> pred (variable, referent).
aussage --> connective (aussage, aussage).
connective --> [und].
connective --> [oder].
quantor --> [es_gibt].
quantor --> [alle].

menge --> [kuh]. ...
pred --> [nahrung]. ...

Beispiele

(alle kuh); nahrung (X, pflanze).
und (
(alle kuh); nahrung (X, pflanze),
(alle kuh) hauttyp (X, fell)).
(es_gibt tier); oder (nahrung (X, pflanze),
nahrung (X, fleisch)).

∀ X kuh (X) --> nahrung (X, pflanze)
∀X kuh (X) --> nahrung (X, pflanze) &
∀X kuh (X) --> hauttyp (X, fell)
∃ X tier (X) | nahrung (X, pflanze) v
nahrung (X, fleisch).

Bezug syntaktische Analyse - DBQ

- beim Parsing wird direkt ein DBQ- Ausdruck aufgebaut

  semantische Grammatik

  NL Satz --> DCG / PrologProgramm --> DBQ-Ausdruck

- Ergebnis des Parsings wird unter Verwendung des Lexikons in DBQ-Ausdruck überführt

  Lexikon ---------->
  NL Satz --> DCG --> synt. Struktur --> PrologProgramm --> DBQ-Ausdruck

- Ergebnis der syntaktischen Analyse wird mit Merkmalen aus dem Lexikon angereichert und dann von einer DCG in DBQ-Ausdrücke überführt.

  synt. Struktur mit LexikonMerkmalen --> DCG --> DBQ-Ausdruck

dbqexp --> referent; aussage:
<NP kasus> = nom
<referent sem > = <NP sem>
<aussage sem> = <VP sem>

referent --> (quantor menge):
<quantor sem> = <art sem>
<menge sem> = <N sem>

referent --> menge:
<menge sem> = <N sem>

aussage --> pred (variable, referent):
<pred sem> = <V Objekt gattung>
<NP kasus> = akk
<referent sem> = <NP sem>

aussage --> pred (variable, referent):
<V1 Objekt eigenart> = fragbar
<pred sem> = <V1 sem>
<V1 semtyp> = <V2 semtyp>
<V2 Objekt eigenart> = unfragbar
<referent sem> = <V2 sem>

Lexikon mit semantischen Merkmalen

Wort fliegen:
<cat> = V
<verbtyp> = vi
<person> = 3
<numerus> = plur
<Subjekt eigenart> = belebt
<Objekt eigenart> = unfragbar
<sem> = fliegen
<semtyp> = bewegung

Wort tun:
<cat> = V,
<verbtyp> = vt
<person> = 3
<numerus> = sing
<Subjekt eigenart> = belebt
<Objekt eigenart> = fragbar
<sem> = #verbvariable
<semtyp> = bewegung

Wort die:
<cat> = art
<numerus> = plur
<det> = best
<sem> = alle

Wort Amseln:
<cat> = N
<numerus> = plur
<genus> = f
<kasus> = nom
<gattung> = tier
<eigenart> = belebt
<sem> = amsel

Beispiel

s(np(art(die),n(amseln)),vp(v(fliegen)))
Da die oberste NP stets die im Nominativ ist, haben wir dies Merkmal nicht in die Strukturbeschreibung aufgenommen. Wir wissen: die oberste NP wird der Referent des DBQ-Ausdrucks.
Die Semantik des Artikels, die auf den Quantor abgebildet wird, finden wir im Lexikon durch eine Lexikonanfrage (in Prolog):

```
art_sem(Wort,Sem):-
    art(art(Wort), _, _, _, _, sem:Sem,_,_)
```

Entsprechend finden wir für alle anderen Wörter der Strukturbeschreibung ihre Semantik, die wir in den DBQ-Ausdruck übernehmen. Ebenso finden wir über den Semtyp zu einem unfragbaren Verb das fragbare, das zum pred des DBQ-Ausdrucks wird.
Die DCG für DBQ-Ausdrücke baut ebenso Strukturen auf wie die DCG für natürlichsprachliche Sätze.
Wir erhalten:
(alle amsel); tun(X,fliegen)

Semantische Repräsentationssprache Probleme

Repräsentationssprache für die intensionale Semantik:
"normale"Präsikatenlogik reicht nicht aus:

- **Sprache enthält vielfältigere Quantoren**
  einige, viele, der typische, relativ wenige... **Defaults!**
- **Wie stellt man Möglichkeiten dar?**
  Es könnte Regen geben. **Defaults!**
- **kollektive und distributive Verben:**
  distributiv: *Hans und Peter schlafen.* Hans schläft. Peter schläft.
  kollektiv: *Hans und Peter streiten um den Apfel.* * Hans streitet. Peter streitet.
  Darstellung durch Prädikat höherer Ordung, das sich auf die Formel für Hans und Peter bezieht?
- **Was macht man mit:**
  Peter besitzt nicht jedes Buch. * Maria liest es.
  Mindestens ein Buch besitzt Peter nicht. Maria liest es.

Weitere Probleme

Kompositionalitätsprinzip:
Die Bedeutung eines Ausdrucks ergibt sich aus seiner syntaktischen Struktur und der Bedeutung seiner Teilausdrücke.
Die Bedeutung eines Teilausdrucks ist sein Denotat (das, was er bezeichnet).

Substitutionsprinzip:
Denotatgleiche Ausdrücke dürfen füreinander eingesetzt werden, ohne daß sich die Bedeutung des Gesamtausdrucks verändert.

Der Bundeskanzler hat die Richtlinienkompetenz.
Der Bundeskanzler ist Gerhard Schröder.
Gerhard Schröder hat die Richtlinienkompetenz.

Der Bundeskanzler hat immer die Richtlinienkompetenz.
Der Bundeskanzler ist Gerhard Schröder.
Gerhard Schröder hat immer die Richtlinienkompetenz. ???

Extension und Intension eines Ausdrucks sind in verschiedenen Welten unterschiedlich. Wir haben nicht (wie in Prädikatenlogik sonst) nur ein Modell.