Einführung in die Computerlinguistik

Merkmalstrukturen und Unifikation

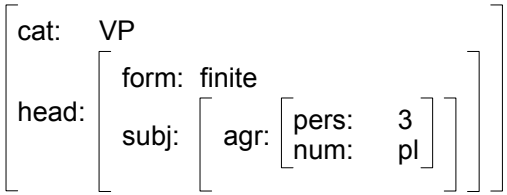
Dozentin: Wiebke Petersen

WS 2004/2005

Universität Potsdam
Institut für Linguistik
Computerlinguistik

Einführung in unifikationsbasierte Grammatikformalismen

Thomas Hanneforth



Merkmalsstrukturen: Eine Übersicht

Illustration: Wozu braucht man Merkmalsstrukturen?

Gegeben sei folgende Grammatik G_1:

$S \rightarrow NP\ VP$
$NP \rightarrow Art\ N$
$VP \rightarrow V$
$VP \rightarrow V\ NP$

$N \rightarrow Frau \mid Mann \mid Kind \mid Frauen \mid Männer \mid Kinder$
$Art \rightarrow der \mid die \mid das \mid den \mid dem$
$V \rightarrow schläft \mid schlafen \mid kennt \mid kennen \mid hilft \mid helfen$

Mit G_1 sind z.B. folgende Sätze ableitbar:	aber auch:
Die Frau kennt den Mann	*Den Männer hilft der Kind*
Die Kinder helfen der Frau	*Die Kinder schläft dem Frau*
Das Kind schläft	*Dem Mann kennt*

Probleme der Grammatik G_1

Die Grammatik G_1 ist nicht in der Lage

1. Kongruenz zwischen Subjekt und Prädikat auszudrücken
2. Abhängigkeit zwischen Artikelform und Numerus und Genus des Nomens auszudrücken
3. Abhängigkeit zwischen Verb und Anzahl und Kasus der Argumente auszudrücken

Wollte man diese Beziehungen in der Grammatik zum Ausdruck bringen, so müßte man neue Kategorien schaffen, u.a.:

N_{sg}	für Nomen im Singular
N_{pl}	für Nomen im Plural
V_{sg}	für Verben im Singular
V_{pl}	für Verben im Plural

aber auch:

N_{sg_mask}	für Nomen im Singular Maskulin
$V_{2\_sg\_dat\_akk}$	für Verben in der 2. Person Singular mit einem Dativ- und einem Akkusativargument (z.B. *gibst*)

G_1 wird dann durch eine Grammatik G_2 ersetzt, die diese Art von Abhängigkeiten ausdrückt, beispielsweise:

$$
\begin{aligned}
S &\rightarrow NP_{1\_sg}\ \ VP_{1\_sg} \\
S &\rightarrow NP_{2\_sg}\ \ VP_{2\_sg} \\
VP_{3\_sg} &\rightarrow V_{3\_sg\_dat\_akk}\ \ NP_{dat}\ \ NP_{akk}
\end{aligned}
$$

Diese Art der Abbildung grammatischer Beziehungen führt zu folgenden Problemen:

1. Die neue Grammatik G_2 ist einem hohen Maße redundant und um Größenordnungen umfangreicher als G_1
2. Strukturelle Information, wie sie durch die Regeln ausgedrückt wird, wird gekoppelt mit Informationen über Kongruenz etc.
3. Die Informationen über die syntaktischen Merkmale einer Kategorie ist im Namen dieser Kategorie kodiert und daher nicht ohne weiteres verfügbar

⇒ Statt Kategorien als atomare Entitäten zu betrachten, in denen Informationen verschiedenster Natur (morphologische, syntaktische, phonologische, semantische Informationen) quasi fest verdrahtet sind, erscheint es sinnvoller, Kategorien als analysierbar, noch besser, als strukturiert aufzufassen.

Merkmalsstrukturen

Merkmalstrukturen (*feature structures*, *Attribut-Wert-Terme*) erlauben, strukturierte Kategorien aufzubauen. Hierzu benötigt man:

1. eine Menge von *Merkmalen* (Merkmalsnamen), z.B. *cat*, *pers*, *num*
2. eine Menge von *atomaren Merkmalswerten*, z.B. *NP*, *3*, *pl*

Beispiel:
Eine Nominalphrase in der 3. Person Plural könnte als Merkmalsstruktur folgendermaßen aussehen:

$$\begin{bmatrix} \text{cat:} & \text{NP} \\ \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix} = D_{NP\_3\_pl}$$

Die Klammernotation ist eine Mengen-Notation.

Beispiel:
Die Merkmalsstruktur für NP_{3_pl} kann als Menge auch so notiert werden:

{ cat: NP, pers: 3, num: pl }

Häufig beziehen sich grammatische Vorgänge auf Gruppen von Merkmalen.
Kongruenz (engl. *agreement*) zwischen Subjekt und Prädikat bezieht sich auf die Übereinstimmung von Person und Numerus.
Hierfür liegt es nahe, die Merkmale hierfür unter einem gemeinsamen Merkmal *agr* zusammenzufassen.

Beispiel:

$$\begin{bmatrix} \text{cat:} & \text{NP} \\ \text{agr:} & \begin{bmatrix} \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix} \end{bmatrix}$$

Technisch läßt man also als Werte von Merkmalen nicht nur atomare Merkmalswerte zu, sondern seinerseits wieder Merkmalsstrukturen.

Hierdurch wird es möglich, Merkmalsstrukturen von beliebiger Komplexität aufzubauen.

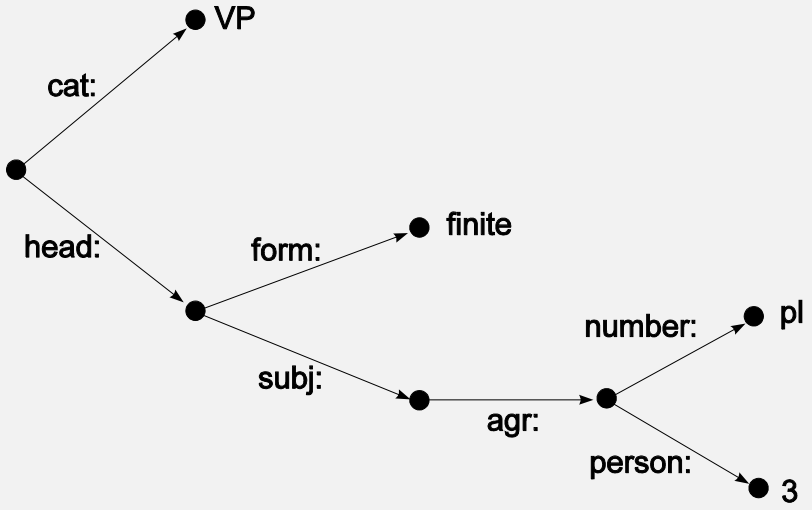
Beispiel: Eine Merkmalsstruktur für eine VP mit einem finiten Verb

$$\begin{bmatrix} \text{cat:} & \text{VP} \\ \text{head:} & \begin{bmatrix} \text{form:} & \text{finite} \\ \text{subj:} & \begin{bmatrix} \text{agr:} & \begin{bmatrix} \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} = D_{VP}$$

Merkmalsstrukturen als Graphen

Alle Merkmalsstrukturen lassen sich auch als gerichtete Graphen darstellen.

Beispiel: D_{VP} hat folgende Darstellung als Graph:



Pfade

Eine Folge von Merkmalsnamen zwischen zwei verbundenen Graphknoten heißt *Pfad*.

Beispiel:
Die Folge *head: subj: agr: person:* ist ein Pfad in D_{VP}.

Reentranz

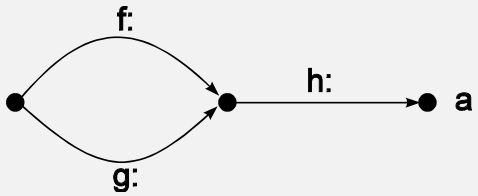
Eine Merkmalsstruktur ist *reentrant*, wenn zwei Merkmale in der Struktur einen gemeinsamen Wert aufweisen.
Reentranz wird in der AVM-Notation durch sog. *tags* (z.B. [1]) wiedergegeben.

Beispiel:

$$\begin{bmatrix} f: \boxed{1} \begin{bmatrix} h: a \end{bmatrix} \\ g: \boxed{1} \end{bmatrix} = D_{id}$$

Reentrante Merkmalsstrukturen in der Graphdarstellung weisen mindestens einen Knoten auf, zu dem zwei Pfeile zeigen.

Beispiel:



Reentranz darf nicht mit Wertgleichheit verwechselt werden.

Beispiel:

$$\begin{bmatrix} f: \begin{bmatrix} h: a \end{bmatrix} \\ g: \begin{bmatrix} h: a \end{bmatrix} \end{bmatrix} = D_{sim}$$

Subsumption

Merkmalsstrukturen können nach ihrem Informationsgehalt geordnet werden.

Beispiel:
Der Informationsgehalt der Merkmalsstruktur D_{NP} ist kleiner als der Informationsgehalt der Merkmalsstruktur D_{NP3_pl}:

$$\begin{bmatrix} \text{cat:} & \text{NP} \end{bmatrix} = D_{NP} \qquad \begin{bmatrix} \text{cat:} & \text{NP} \\ \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix} = D_{NP_{3\_pl}}$$

- Eine Merkmalsstruktur *A subsumiert* eine Merkmalsstruktur *B* (symbolisch $A \sqsubseteq B$), wenn der Informationsgehalt von *A kleiner* oder *gleich* als der von *B* ist.

- Eine andere Sichtweise auf Merkmalsstrukturen ist, sie danach zu klassifizieren, welche Kardinalität die Menge der durch sie charakterisierten Objekte hat.

- Ist der Informationsgehalt einer Merkmalsstruktur *A kleiner/gleich* gegenüber dem einer Merkmalsstruktur *B*, so ist die Zahl der durch *A* charakterisierten Objekte *größer/gleich* als die Zahl der durch *B* charakterisierten Objekte.

🔔 *Vorsicht*: Subsumption wird in der Literatur manchmal auch gerade umgekehrt definiert. 🔔

Die Subsumptionsrelation definiert eine partielle Ordnung über allen möglichen Merkmalsstrukturen.
Die Merkmalsstruktur mit dem geringsten Informationsgehalt ist

$$\begin{bmatrix} \end{bmatrix} = D_{var}$$

Beispiel: Es gelten u.a. folgende Beziehungen:

$$\begin{bmatrix} \end{bmatrix} \sqsubseteq \begin{bmatrix} \text{cat:} & \text{NP} \end{bmatrix} \sqsubseteq \begin{bmatrix} \text{cat:} & \text{NP} \\ \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix}$$

und

$$\begin{bmatrix} \text{f:} & \begin{bmatrix} \text{h:} & \text{a} \end{bmatrix} \\ \text{g:} & \begin{bmatrix} \text{h:} & \text{a} \end{bmatrix} \end{bmatrix} \sqsubseteq \begin{bmatrix} \text{f:} & \boxed{1} \begin{bmatrix} \text{h:} & \text{a} \end{bmatrix} \\ \text{g:} & \boxed{1} \end{bmatrix}$$

Es gilt hingegen nicht:

$$\begin{bmatrix} \text{cat:} & \text{NP} \end{bmatrix} \sqsubseteq \begin{bmatrix} \text{cat:} & \text{VP} \end{bmatrix}$$

D_{Var} subsumiert alle anderen Merkmalsstrukturen.

Unifikation

Unifikation ist eine Operation, die den Informationsgehalt zweier Merkmalsstrukturen miteinander kombiniert, sofern beide miteinander kompatibel sind.

Beispiel:

Die Unifikation der Merkmalsstrukturen

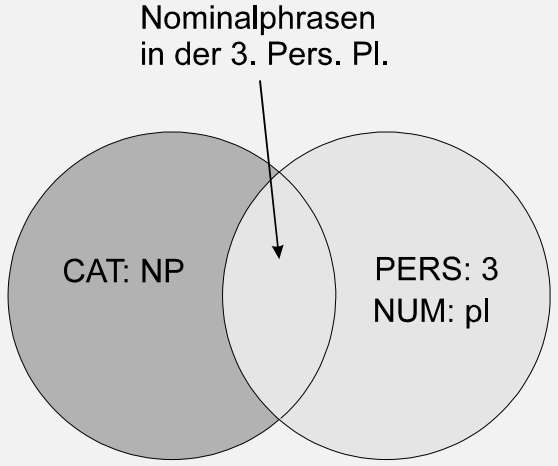
$$\begin{bmatrix} \text{cat:} & \text{NP} \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix}$$

ergibt:

$$\begin{bmatrix} \text{cat:} & \text{NP} \\ \text{pers:} & 3 \\ \text{num:} & \text{pl} \end{bmatrix}$$

Mengentheoretisch kann die Unifikation als Schnittmengenbildung definiert werden:

Beispiel:



Nicht immer können die Informationen zweier Merkmalsstrukturen miteinander kombiniert werden.

Beispiel:
Die Informationen der folgender Merkmalstrukturen sind nicht miteinander kompatibel:

$$\begin{bmatrix} \text{cat: NP} \end{bmatrix} \qquad \begin{bmatrix} \text{cat: VP} \end{bmatrix}$$

In diesen Fällen sagt man auch, daß die Unifikation scheitert.

(Halb)Formale Definition der Unifikation:

Die Unifikation zweier Merkmalsstrukturen *A* und *B* (symbolisch $A \sqcup B$) ist die kleinste (d.h. allgemeinste) Merkmalsstruktur *C*, so daß gilt: $A \sqsubseteq C$ und $B \sqsubseteq C$.

Beispiele für Unifikation:

♦ *monotone Hinzufügung von Information:*

$$\begin{bmatrix} \text{cat: NP} \end{bmatrix} \sqcup \begin{bmatrix} \text{agr:} \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \text{cat: NP} \\ \text{agr:} \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix}$$

♦ *Idempotenz:*

$$\begin{bmatrix} \text{f:} \begin{bmatrix} \text{h: a} \end{bmatrix} \\ \text{g:} \begin{bmatrix} \text{h: b} \end{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{f:} \begin{bmatrix} \text{h: a} \end{bmatrix} \\ \text{g:} \begin{bmatrix} \text{h: b} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \text{f:} \begin{bmatrix} \text{h: a} \end{bmatrix} \\ \text{g:} \begin{bmatrix} \text{h: b} \end{bmatrix} \end{bmatrix}$$

♦ *Die leere Merkmalsstruktur ist das neutrale Element der Unifikation:*

$$\begin{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{cat: NP} \\ \text{agr: } \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \text{cat: NP} \\ \text{agr: } \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix} \\ \text{agr: } \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \begin{bmatrix} \text{pers: 3} \end{bmatrix} \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \begin{bmatrix} \text{num: sg} \\ \text{pers: 3} \end{bmatrix} \end{bmatrix} \\ \text{agr: } \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \boxed{1} \begin{bmatrix} \text{num: sg} \end{bmatrix} \end{bmatrix} \\ \text{agr: } \boxed{1} \end{bmatrix} \sqcup \begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \begin{bmatrix} \text{pers: 3} \end{bmatrix} \end{bmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} \text{subj: } \begin{bmatrix} \text{agr: } \boxed{1} \begin{bmatrix} \text{num: sg} \\ \text{pers: 3} \end{bmatrix} \end{bmatrix} \\ \text{agr: } \boxed{1} \end{bmatrix}$$

$$\begin{bmatrix} \text{f: } \begin{bmatrix} \text{h: a} \end{bmatrix} \\ \text{g: } \begin{bmatrix} \text{h: b} \end{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{f: } \boxed{1} \\ \text{g: } \boxed{1} \end{bmatrix} = \textit{fail}$$

Generalisierung

Die Generalisierung zweier Merkmalsstrukturen *A* und *B* ist die größte (d.h. spezifischste) Merkmalsstruktur *C*, so daß gilt:

$$C \sqsubseteq A \text{ und } C \sqsubseteq B.$$

Verbandsstrukturen

Unifikation, Generalisierung und Subsumption kann man sich leicht in Form einer Verbandsstruktur (engl. *lattice*) veranschaulichen.