

Einführung in die Computerlinguistik

Merkmalsgrammatik/ Unifikation/ Semantik

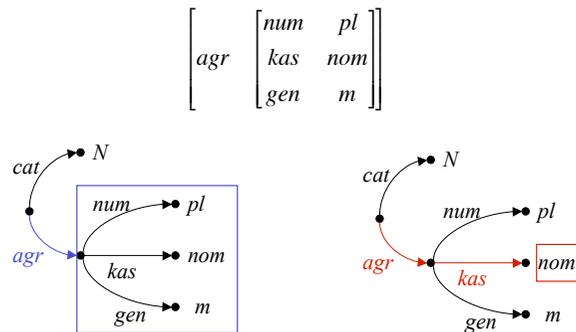
ACHTUNG: Übung 3 von Ü7 braucht (noch) nicht gemacht zu werden!

WS 2008/2009

Manfred Pinkal

Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Merkmalspfade, Pfadgleichungen



Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Merkmalsabhängigkeiten

- Grammatische Merkmale von Ausdrücken in der syntaktischen Struktur hängen in systematischer Weise voneinander ab.
- Die grundlegenden Typen solcher Beziehungen sind
 - Kongruenz und
 - Rektion oder Subkategorisierung

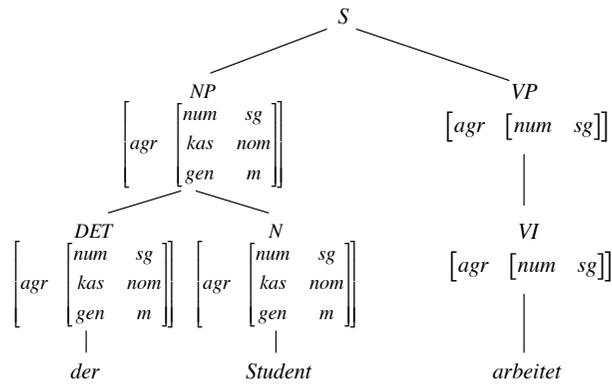
Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

CFG mit Merkmalsconstraints, Beispiel

$S \rightarrow NP VP$ $\langle NP agr num \rangle = \langle VP agr num \rangle$ $\langle NP agr kas \rangle = nom$	$VI \rightarrow arbeitet$ $\langle VI agr num \rangle = sg$
$VP \rightarrow VI$ $\langle VP agr num \rangle = \langle VI agr num \rangle$	$VI \rightarrow arbeiten$ $\langle VI agr num \rangle = pl$
$VP \rightarrow VT NP$ $\langle VP agr num \rangle = \langle VT agr num \rangle$ $\langle NP agr kas \rangle = akk$	$N \rightarrow Student$ $\langle N agr num \rangle = sg$ $\langle N agr gen \rangle = m$ $\langle N agr kas \rangle = nom$
$NP \rightarrow DET N$ $\langle DET agr \rangle = \langle N agr \rangle$ $\langle NP agr \rangle = \langle N agr \rangle$	$DET \rightarrow der$ $\langle DET agr num \rangle = sg$ $\langle DET agr gen \rangle = m$ $\langle DET agr kas \rangle = nom$

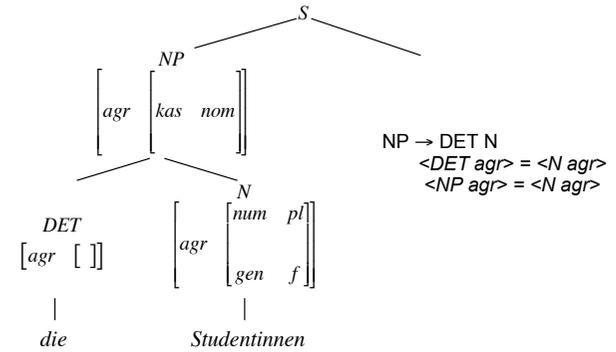
Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Die Anwendung von Merkmals-Constraints



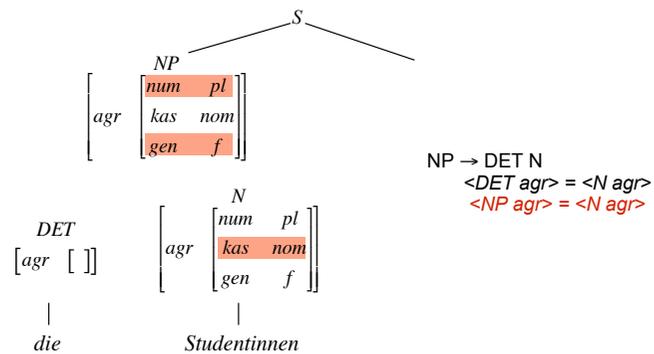
Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Die Anwendung von Merkmals-Constraints



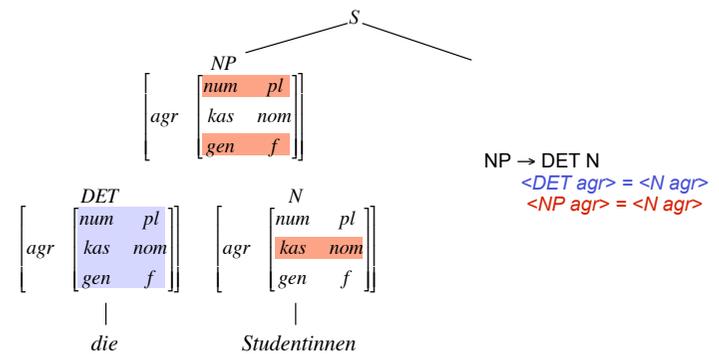
Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Die Anwendung von Merkmals-Constraints



Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Die Anwendung von Merkmalsconstraints

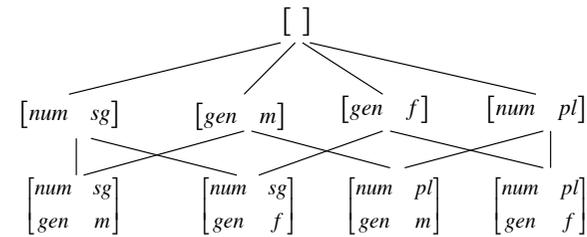


Vorlesung "Einführung in die CL" 2008/2009 © M. Pinkal UdS Computerlinguistik

Unifikation

- Bei der Anwendung von Pfadgleichungen wird nicht die Identität aktueller Werte abgeprüft, sondern
 - die Kompatibilität der vorhandenen Merkmalsinformation getestet, indem
 - die Information auf beiden Seiten der Gleichung zusammengeführt.
- Dies geschieht, indem die entsprechenden Merkmalsinformationen **unifiziert** werden. Wir schreiben $A \sqcup B$.
- Resultat der Unifikation ist die (allgemeinste) Merkmalsstruktur, die die Information aus beiden Merkmalsstrukturen umfasst, wenn es eine solche Struktur gibt. Ansonsten schlägt sie fehl.
- Wenn A allgemeiner ist als B ist, d.h., wenn Merkmalsstruktur B die Information aus Merkmalsstruktur A komplett enthält, sprechen wir von Subsumption: A subsumiert B, oder $A \sqsubseteq B$.
- Das Resultat der Unifikation von A und B: $A \sqcup B$ ist die allgemeinste Struktur C, sodass $A \sqsubseteq C$ und $B \sqsubseteq C$, d.h., die Struktur, die genau die gemeinsame in A und B enthaltene Information enthält.

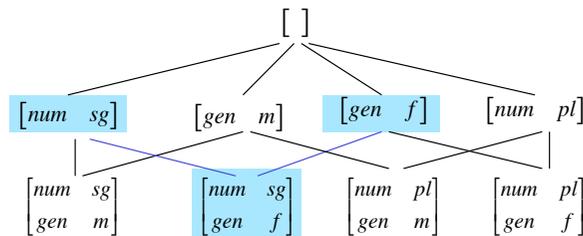
Subsumption



Die Graphik stellt die Halbordnung über Merkmalsstrukturen dar, die durch die Subsumptionsrelation etabliert wird: Die obere Struktur subsumiert jeweils die untere, zum Beispiel

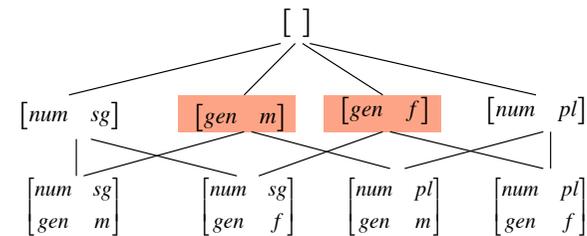
$$[] \sqsubseteq [gen \ m] \sqsubseteq \begin{bmatrix} num & sg \\ gen & m \end{bmatrix}$$

Subsumption und Unifikation



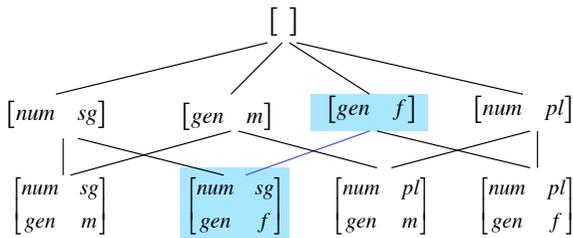
$$[num \ sg] \sqcup [gen \ f] = \begin{bmatrix} num & sg \\ gen & f \end{bmatrix}$$

Subsumption und Unifikation



$$[gen \ m] \sqcup [gen \ f] = \text{fail}$$

Subsumption und Unifikation



$$\begin{bmatrix} num & sg \\ gen & f \end{bmatrix} \sqcup \begin{bmatrix} gen & f \end{bmatrix} = \begin{bmatrix} num & sg \\ gen & f \end{bmatrix}$$

allgemein: wenn $A \sqsubseteq B$, so ist $A \sqcup B = B$

Berechnung von $A \sqcup B$

- Wie berechnet sich das Resultat der Unifikation aus A und B?
- Fall 1: Beide Seiten sind atomar:
 - $a \sqcup a = a$
 - $a \sqcup b = \text{fail}$, wenn $a \neq b$
- Fall 2: Eine Seite ist atomar, die andere komplexe Merkmalsstruktur
 - $a \sqcup F = \text{fail}$
- Fall 3: Beide Seiten sind komplexe Merkmalstrukturen:
 - $F \sqcup G = H$, wobei die Attribute von H Attribute die Vereinigung der Attribute von F und G sind.
 - Der Wert für jedes Attribut f in H ist
 - der Wert von f in F (G), wenn f nur in F (G) instantiiert ist
 - $A \sqcup B$, wenn A und B die Werte von f in F bzw. G sind.

Unifikation, Beispiele

$$sg \cup sg = sg$$

$$sg \cup pl = \text{fail}$$

$$\begin{bmatrix} cat & N \\ agr & \begin{bmatrix} num & sg \\ gen & f \end{bmatrix} \end{bmatrix} \cup \begin{bmatrix} agr & [kas \text{ akk}] \end{bmatrix} = \begin{bmatrix} cat & N \\ agr & \begin{bmatrix} num & sg \\ gen & f \\ kas & akk \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} cat & N \\ agr & \begin{bmatrix} num & sg \\ gen & f \end{bmatrix} \end{bmatrix} \cup \begin{bmatrix} agr & [gen \text{ m}] \end{bmatrix} = \text{fail}$$

CFG mit Merkmalsconstraints

$S \rightarrow NP VP$	$VT \rightarrow \text{wählt}$
$\langle NP \ AGR \ NUM \rangle = \langle VP \ AGR \ NUM \rangle$	$\langle VT \ AGR \ NUM \rangle = sg$
$\langle NP \ AGR \ NUM \rangle = nom$	
$VP \rightarrow VI$	$VI \rightarrow \text{arbeitet}$
$\langle VP \ AGR \ NUM \rangle = \langle VI \ AGR \ NUM \rangle$	$\langle VI \ AGR \ NUM \rangle = sg$
$VP \rightarrow VT \ NP$	$N \rightarrow \text{Studentin}$
$\langle VP \ AGR \ NUM \rangle = \langle VT \ AGR \ NUM \rangle$	$\langle N \ AGR \ NUM \rangle = sg$
$\langle NP \ AGR \ KAS \rangle = akk$	$\langle N \ AGR \ GEN \rangle = f$
$NP \rightarrow \text{DET } N$	$N \rightarrow \text{Fach}$
$\langle \text{DET} \ AGR \rangle = \langle N \ AGR \rangle$	$\langle N \ AGR \ NUM \rangle = sg$
$\langle NP \ AGR \rangle = \langle N \ AGR \rangle$	$\langle N \ AGR \ GEN \rangle = n$
	$\text{DET} \rightarrow \text{die}$
	$\langle \text{DET} \ AGR \ GEN \rangle = f$
	$\text{DET} \rightarrow \text{das}$
	$\langle \text{DET} \ AGR \ NUM \rangle = sg$
	$\langle \text{DET} \ AGR \ GEN \rangle = n$

Alternative Notation von lexikalischen Regeln

- Wir schreiben äquivalent, aber übersichtlicher statt

$$N \rightarrow \textit{Studentin} \quad \textit{Studentin}: \begin{bmatrix} \textit{CAT} & N \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & f \end{bmatrix} \end{bmatrix}$$

$$\begin{aligned} <N \textit{ AGR NUM}> = \textit{sg} \\ <N \textit{ AGR GEN}> = f \end{aligned}$$

- Die in den Merkmalsconstraints enthaltene Information drücken wir in einer Merkmalsstruktur aus.
- Wir fügen ein Attribut (*CAT*) hinzu, das die Kategorie bezeichnet.
- Auch bei nicht-lexikalischen Kategorien nehmen ein *CAT*-Attribut an (eingeführt durch Constraints der Form
 - $<NP \textit{ CAT}> = NP$, allgemein: $<X \textit{ CAT}> = X$, die wir aber nicht explizit aufführen)

Kopfmerkmale

- Bestimmte Merkmale vererben sich „entlang der Kopflinie“.
- Insbesondere teilt die Phrase diese Merkmale mit ihrem lexikalischen Kopf,
 - zum Beispiel die Kongruenzmerkmale (*AGR*),
 - aber auch andere Merkmale, zum Beispiel die „Verbform“ (*VFORM*): Infinitivkonstruktionen enthalten Infinitive, finite Verbphrasen finite (flektierte) Verben, etc.
- Um die Regularität in der Grammatik ausdrücken, führen wir als zusätzliches Attribut *HEAD* ein (unter dem die „Kopfmerkmale“ aufgeführt werden).

CFG mit Merkmalsconstraints

$$S \rightarrow NP VP \quad \begin{aligned} <NP \textit{ AGR NUM}> = <VP \textit{ AGR NUM}> \\ <NP \textit{ AGR KAS}> = \textit{nom} \end{aligned} \quad \begin{aligned} \textit{wählt}: & \begin{bmatrix} \textit{CAT} & \textit{VT} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \\ \textit{arbeitet}: & \begin{bmatrix} \textit{CAT} & \textit{VI} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \end{aligned}$$

$$VP \rightarrow VI \quad <VP \textit{ AGR NUM}> = <VI \textit{ AGR NUM}>$$

$$VP \rightarrow VT NP \quad \begin{aligned} <VP \textit{ AGR NUM}> = <VT \textit{ AGR NUM}> \\ <NP \textit{ AGR KAS}> = \textit{akk} \end{aligned}$$

$$NP \rightarrow DET N \quad \begin{aligned} <DET \textit{ AGR}> = <N \textit{ AGR}> \\ <NP \textit{ AGR}> = <N \textit{ AGR}> \end{aligned} \quad \begin{aligned} \textit{Studentin}: & \begin{bmatrix} \textit{CAT} & N \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & f \end{bmatrix} \end{bmatrix} \\ \textit{Fach}: & \begin{bmatrix} \textit{CAT} & N \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & n \end{bmatrix} \end{bmatrix} \\ \textit{die}: & \begin{bmatrix} \textit{CAT} & \textit{DET} \\ \textit{AGR} & \begin{bmatrix} \textit{GEN} & f \end{bmatrix} \end{bmatrix} \\ \textit{das}: & \begin{bmatrix} \textit{CAT} & \textit{DET} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & n \end{bmatrix} \end{bmatrix} \end{aligned}$$

Grammatik mit Kopfmerkmalen

$$S \rightarrow NP VP \quad \begin{aligned} <S \textit{ HEAD}> = <VP \textit{ HEAD}> \\ <S \textit{ HEAD AGR NUM}> = <NP \textit{ HEAD AGR NUM}> \\ <NP \textit{ HEAD AGR KAS}> = \textit{nom} \end{aligned} \quad \begin{aligned} \textit{wählt}: & \begin{bmatrix} \textit{CAT} & \textit{VT} \\ \textit{HEAD} & \begin{bmatrix} \textit{VFORM} & \textit{finit} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textit{arbeitet}: & \begin{bmatrix} \textit{CAT} & \textit{VI} \\ \textit{HEAD} & \begin{bmatrix} \textit{VFORM} & \textit{finit} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{aligned}$$

$$VP \rightarrow VI \quad <VP \textit{ HEAD}> = <VI \textit{ HEAD}>$$

$$VP \rightarrow VT NP \quad \begin{aligned} <VP \textit{ HEAD}> = <VT \textit{ HEAD}> \\ <NP \textit{ HEAD AGR KAS}> = \textit{akk} \end{aligned}$$

$$NP \rightarrow DET N \quad \begin{aligned} <NP \textit{ HEAD}> = <N \textit{ HEAD}> \\ <DET \textit{ AGR}> = <N \textit{ AGR}> \end{aligned} \quad \begin{aligned} \textit{Studentin}: & \begin{bmatrix} \textit{CAT} & N \\ \textit{HEAD} & \begin{bmatrix} \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & f \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textit{Fach}: & \begin{bmatrix} \textit{CAT} & N \\ \textit{HEAD} & \begin{bmatrix} \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & n \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textit{die}: & \begin{bmatrix} \textit{CAT} & \textit{DET} \\ \textit{AGR} & \begin{bmatrix} \textit{GEN} & f \end{bmatrix} \end{bmatrix} \\ \textit{das}: & \begin{bmatrix} \textit{CAT} & \textit{DET} \\ \textit{AGR} & \begin{bmatrix} \textit{NUM} & \textit{sg} \\ \textit{GEN} & n \end{bmatrix} \end{bmatrix} \end{aligned}$$

Subkategorisierung

- Transitive und intransitive Verben gehören in unserer Beispielgrammatik zu verschiedenen Kategorien. Konsequenterweise müssten auch für ditransitive Verben (*geben*), Verben mit Dativobjekt (*gefallen*), Verben mit Präpositionalen Objekt (*warten auf*), Verben mit *zu*-Infinitiv (*versuchen*) unterschiedliche Kategorien vorgesehen werden. Ebenso für Präsuppositionen, Adjektive, Substantive, die unterschiedliche Argumente nehmen.
- Eleganter ist die Lösung, die Subkategorisierungseigenschaften durch ein Merkmal auszudrücken (traditionell „SUBCAT“).

Grammatik mit Subkategorisierung [1]

$VP \rightarrow V NP$
 $\langle VP HEAD \rangle = \langle V HEAD \rangle$
 $\langle V SUBCAT \rangle = trans$
 $\langle NP HEAD AGR KAS \rangle = akk$

$VP \rightarrow V$
 $\langle VP HEAD \rangle = \langle VI HEAD \rangle$
 $\langle V SUBCAT \rangle = intrans$

wählt: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & trans \end{bmatrix}$

arbeitet: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & intrans \end{bmatrix}$

Grammatik mit Subkategorisierung [2]

$VP \rightarrow V NP$
 $\langle VP HEAD \rangle = \langle V HEAD \rangle$
 $\langle V SUBCAT \rangle = \langle NP \rangle$
 $\langle NP HEAD AGR KAS \rangle = akk$

wählt: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & \begin{bmatrix} HEAD & [AGR [KAS akk]] \end{bmatrix} \end{bmatrix}$

gefällt: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & \begin{bmatrix} HEAD & [AGR [KAS dat]] \end{bmatrix} \end{bmatrix}$

- Wir identifizieren per Pfadgleichung die komplette NP-Merkmalstruktur mit dem SUBCAT-Merkmal des Verbs und schreiben die spezifische Anforderung einzelner Verben in deren Merkmalstruktur. Die obige Regel beschreibt alle „V NP“-Konstruktionen.
- Die tatsächliche Behandlung der Subkategorisierung in Grammatik-Formalismen ist noch allgemeiner und eleganter.

Grammatikfragment 2

$S \rightarrow NP VP$
 $\langle S HEAD \rangle = \langle VP HEAD \rangle$
 $\langle S HEAD AGR NUM \rangle = \langle NP HEAD AGR NUM \rangle$
 $\langle NP HEAD AGR KAS \rangle = nom$

$VP \rightarrow V$
 $\langle VP HEAD \rangle = \langle VI HEAD \rangle$
 $\langle V SUBCAT \rangle = empty$

$VP \rightarrow V NP$
 $\langle VP HEAD \rangle = \langle V HEAD \rangle$
 $\langle V SUBCAT \rangle = \langle NP \rangle$

$NP \rightarrow DET N$
 $\langle NP HEAD \rangle = \langle N HEAD \rangle$
 $\langle NP HEAD AGR \rangle = \langle DET AGR \rangle$

das: $\begin{bmatrix} CAT & DET \\ AGR & [NUM sg] \end{bmatrix}$

die: $\begin{bmatrix} CAT & DET \\ AGR & [GEN f] \end{bmatrix}$

der: $\begin{bmatrix} CAT & DET \\ AGR & [] \end{bmatrix}$

wählt: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & \begin{bmatrix} HEAD & [AGR [KAS akk]] \end{bmatrix} \end{bmatrix}$

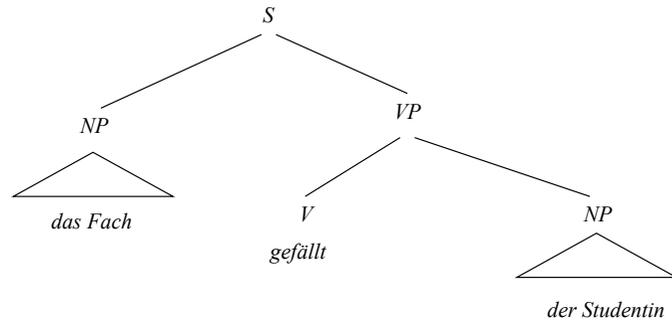
gefällt: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & \begin{bmatrix} HEAD & [AGR [KAS dat]] \end{bmatrix} \end{bmatrix}$

arbeitet: $\begin{bmatrix} CAT & V \\ HEAD & \begin{bmatrix} VFORM & finit \\ AGR & [NUM sg] \end{bmatrix} \\ SUBCAT & empty \end{bmatrix}$

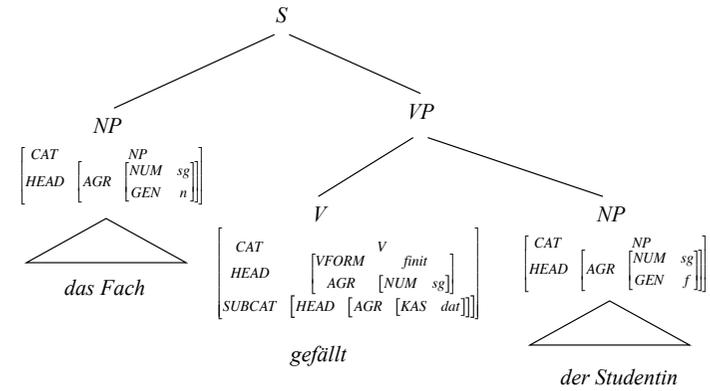
Studentin: $\begin{bmatrix} CAT & N \\ HEAD & \begin{bmatrix} NUM & sg \\ GEN & f \end{bmatrix} \end{bmatrix}$

Fach: $\begin{bmatrix} CAT & N \\ HEAD & \begin{bmatrix} NUM & sg \\ GEN & n \end{bmatrix} \end{bmatrix}$

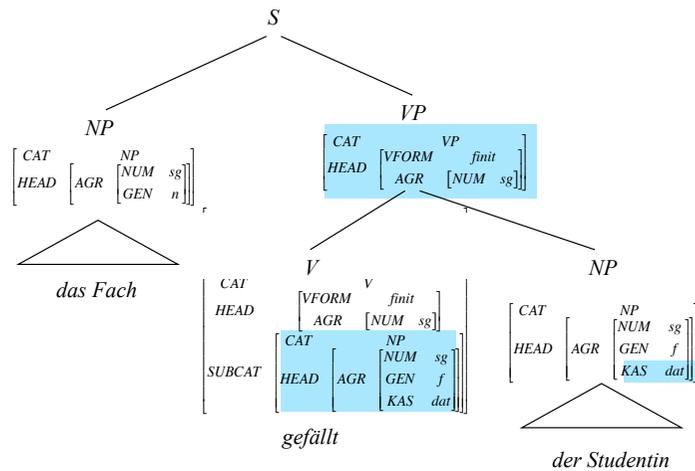
Eine Analyse mit Fragment 2



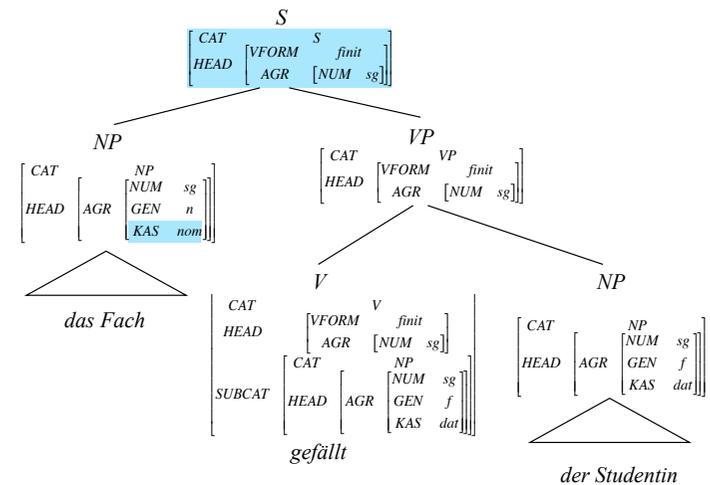
Eine Analyse mit Fragment 2



Eine Analyse mit Fragment 2



Eine Analyse mit Fragment 2



„Reentrancy“

- Subjekt-Verb-Kongruenz kann alternativ so ausgedrückt werden:

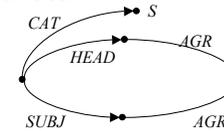
$$\begin{aligned}
 S &\rightarrow NP VP \\
 \langle S \text{ HEAD} \rangle &= \langle VP \text{ HEAD} \rangle \\
 \langle S \text{ SUBJ} \rangle &= \langle NP \text{ HEAD} \rangle \\
 \langle S \text{ HEAD AGR} \rangle &= \langle S \text{ SUBJ AGR} \rangle
 \end{aligned}$$

D.h., wir fügen ein „SUBJ“-Merkmal zu den Merkmalen des Satzes hinzu, kopieren die Kopfmerkmale des Subjekts in dies Merkmal (korrekter: Wir identifizieren das SUBJ-Merkmal mit den Kopfmerkmalen des Subjekts), und verlangen durch die Pfadgleichung $\langle S \text{ HEAD AGR} \rangle = \langle S \text{ HEAD SUBJ AGR} \rangle$, dass die Kongruenzmerkmale des Satzes (die von der VP und damit vom Verb kommen) mit den Kongruenzmerkmalen unter SUBJ, die von der Subjekts-NP kommen, identisch sind.

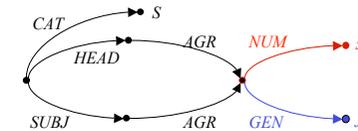
- Dies ist etwas Neues: Der Constraint besagt, dass zwei Pfade des S-Graphen auf denselben Knoten verweisen. Mit anderen Worten: Sie verzweigen an einem Knoten und laufen in einem anderen Knoten wieder zusammen (engl. „reentrancy“).

Reentrancy

- Um die Constraints der S-Regel zu erfüllen, muss der S-Graph die folgende Teilstruktur aufweisen:



- Wenn über den $\langle S \text{ HEAD} \rangle$ -Pfad (durch $\langle S \text{ HEAD} \rangle = \langle VP \text{ HEAD} \rangle$) oder über den $\langle S \text{ SUBJ} \rangle$ -Pfad (durch $\langle S \text{ SUBJ} \rangle = \langle NP \text{ HEAD} \rangle$) neue Merkmalsinformation in das AGR-Attribut kommt, steht sie gleichzeitig auch über den jeweils anderen Pfad zur Verfügung. Unterschiedliche Kongruenz-Information, die vom Subjekt bzw. der VP beigesteuert wird, wird unifiziert. Beispiel: „*sie arbeiten*“ ergibt den folgenden Graph. Analyse von „er arbeiten“ schlägt fehl.



Reentrancy

- In einer Merkmalsmatrix stellen wir Reentrancy durch Koindizierung dar:

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle]
 \end{bmatrix}$$

- Der zweite Merkmalsgraph auf der letzten Folie kann am saubersten (aber etwas umständlich) so dargestellt werden:

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle]
 \end{bmatrix}
 \langle 1 \rangle
 \begin{bmatrix}
 \text{NUM} & \text{sg} \\
 \text{GEN} & f
 \end{bmatrix}$$

- Standardnotation: Die Merkmalsinformation, die von beiden Pfaden geteilt wird, wird an eine Stelle geschrieben, die Koindizierung drückt aus, dass die Information gleichzeitig unter einem anderen Pfad hängt.

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle]
 \end{bmatrix}$$

Reentrancy

- Alle folgenden Darstellungen sind äquivalent:

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle]
 \end{bmatrix}
 \quad
 \begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}]
 \end{bmatrix}$$

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}]
 \end{bmatrix}
 \quad
 \begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}] \\
 \text{SUBJ} & [\text{AGR} \langle 1 \rangle \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}]
 \end{bmatrix}$$

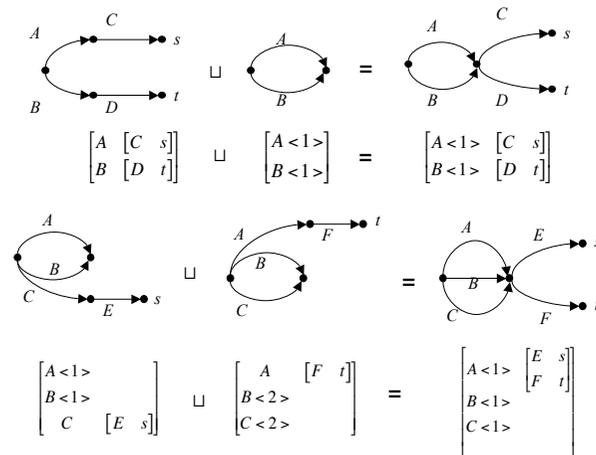
- Nicht äquivalent ist dagegen diese Struktur: Sie ist weniger informativ, weil sie nur zwei Merkmalsstrukturen in Kopie enthält, aber nicht die Information, dass die Werte der Pfade $\langle \text{HEAD AGR} \rangle$ und $\langle \text{SUBJ AGR} \rangle$ immer identisch sein müssen.

$$\begin{bmatrix}
 \text{CAT} & S \\
 \text{HEAD} & [\text{AGR} \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}] \\
 \text{SUBJ} & [\text{AGR} \begin{bmatrix} \text{NUM} & \text{sg} \\ \text{GEN} & f \end{bmatrix}]
 \end{bmatrix}$$

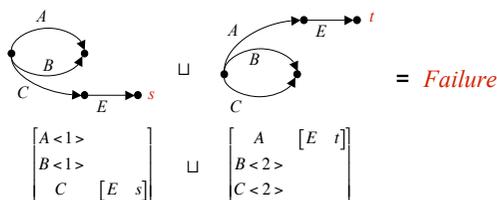
Reentrancy, Subsumption, Unifikation

- A subsumiert B/ A ist mindestens so allgemein wie B/ B ist mindestens so informativ wie A:
 $A \sqsubseteq B$ genau dann, wenn B mindestens die gesamte Merkmals-Information enthält, die A enthält.
- Für Merkmalsstrukturen mit Reentrancy heißt das:
 - B enthält (mindestens) die atomaren Werte für Merkmalspfade, die A enthält
 - B enthält (mindestens) die Pfadverknüpfungen, die A enthält.
- Das Resultat der Unifikation von A und B: $A \sqcup B$ ist die kleinste Merkmalsstruktur C, die von A und B subsumiert wird; das heißt:
 - C enthält genau die Merkmalspfade mit den atomaren Werten, die in A oder in B enthalten sind
 - C enthält genau die Pfadverknüpfungen, die in A oder in B enthalten sind.

Reentrancy und Unifikation: Beispiele



Reentrancy und Unifikation: Beispiele



- Achtung: Die primäre Repräsentation für Merkmalsstrukturen mit Reentrancy sind (gerichtete azyklische) Graphen. Die Matrixdarstellung ist abgeleitet, in gewissem Umfang beliebig und nicht vollständig transparent.
- Im Zweifelsfall deshalb Unifikation anhand von Merkmalsgraphen berechnen (und das Resultat ggf. in Matrixform umschreiben).

Grammatik mit semantischer Information

- Die Benutzung von Pfadverknüpfungen bzw. die Koindizierung von Merkmalsstrukturen ist ein starkes und elegantes Werkzeug, um linguistische Struktur aufzubauen, zu kopieren und zu modifizieren.
- Auf den folgenden Folien wird dazu ein einfaches Beispiel gegeben: Eine Grammatik, deren Constraints parallel zur grammatischen Analyse unter einem „SEM“-Attribut semantische Repräsentationen aufbauen.
- Die komplette hierzu nötige Grammatikerweiterung (blau markiert) findet im Lexikon statt. Die Syntaxregeln sind unverändert die von Fragment 2.
- Vorsicht: Die Grammatik dient nur der Illustration. Die erzeugten semantischen Strukturen sind aus semantischer Sicht nicht korrekt.

Grammatikfragment 3: Lexikon

<i>arbeitet:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{HEAD} \\ \text{SUBCAT} \end{array} \right. \left. \begin{array}{c} \text{V} \\ \text{VFORM} \\ \text{SUBJ} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{finit} \\ \text{AGR} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{NUM sg} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{arbeiten} \\ \text{arg} \end{array} \right]$
<i>wählt:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{HEAD} \\ \text{SUBCAT} \end{array} \right. \left. \begin{array}{c} \text{V} \\ \text{VFORM} \\ \text{SUBJ} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{finit} \\ \text{AGR} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{NUM sg} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{wählen} \\ \text{ARG1} \\ \text{ARG2} \\ \text{SEM} \end{array} \right]$

<i>Studentin:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{HEAD} \end{array} \right. \left. \begin{array}{c} \text{N} \\ \text{AGR} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{NUM sg} \\ \text{GEN f} \\ \text{studentin} \end{array} \right]$
<i>Fach:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{HEAD} \end{array} \right. \left. \begin{array}{c} \text{N} \\ \text{AGR} \\ \text{SEM} \end{array} \right. \left. \begin{array}{c} \text{NUM sg} \\ \text{GEN n} \\ \text{fach} \end{array} \right]$
<i>das:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{AGR} \end{array} \right. \left. \begin{array}{c} \text{DET} \\ \text{NUM sg} \\ \text{GEN n} \end{array} \right]$
<i>die:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{AGR} \end{array} \right. \left. \begin{array}{c} \text{DET} \\ \text{GEN f} \end{array} \right]$
<i>der:</i>	$\left[\begin{array}{c} \text{CAT} \\ \text{AGR} \end{array} \right. \left. \begin{array}{c} \text{DET} \\ \text{[]} \end{array} \right]$

Grammatikfragment 3: Regeln

- $S \rightarrow NP VP$
 - $\langle S \text{ HEAD} \rangle = \langle VP \text{ HEAD} \rangle$
 - $\langle S \text{ HEAD SUBJ} \rangle = \langle NP \text{ HEAD} \rangle$
 - $\langle S \text{ HEAD AGR} \rangle = \langle S \text{ HEAD SUBJ AGR} \rangle$
- $VP \rightarrow V$
 - $\langle VP \text{ HEAD} \rangle = \langle V \text{ HEAD} \rangle$
 - $\langle V \text{ SUBCAT} \rangle = \text{empty}$
- $VP \rightarrow V NP$
 - $\langle VP \text{ HEAD} \rangle = \langle V \text{ HEAD} \rangle$
 - $\langle V \text{ SUBCAT} \rangle = \langle NP \rangle$
- $NP \rightarrow DET N$
 - $\langle NP \text{ HEAD} \rangle = \langle N \text{ HEAD} \rangle$
 - $\langle NP \text{ HEAD AGR} \rangle = \langle DET \text{ AGR} \rangle$