

Ling/CSE 472: Introduction to Computational Linguistics

5/13/15

Feature structures and unification

Overview

- Problems with CFG
- Feature structures
- Unification
- Agreement
- Subcategorization
- Long-distance Dependencies
- Reading questions

Problems with CFG (with atomic node labels)

- Simple rules, with simple category sets overgenerate:
- What are some non-sentences that this CFG licenses?

$S \rightarrow NP VP$

$NP \rightarrow (Det) Noun$

$VP \rightarrow Verb (NP) (NP|PP)$

$PP \rightarrow Prep NP$

$Noun \rightarrow$ cat, cats, dog, dogs, I, you, we, they, he, she, it

$Det \rightarrow$ the, a, this, these, some, many

$Verb \rightarrow$ bark, barks, barked, am, is, are, was, were,
rely, relies, see, sees, saw

$Prep \rightarrow$ on, in, above, before

Problems with CFG (with atomic node labels)

- How could that be fixed, using the CFG formalism?

$S \rightarrow NP_sg VP_sg$

$S \rightarrow NP_pl VP_pl$

$NP_sg \rightarrow (Det_sg) Noun_sg$

$NP_pl \rightarrow (Det_pl) Noun_pl$

$VP_sg \rightarrow V_intrans_sg$

$VP_sg \rightarrow V_trans_sg NP_sg$

$VP_sg \rightarrow V_trans_sg NP_pl$

...

Generalized Phrase Structure Grammar (GPSG)

- Gazdar et al 1982
- Added feature structures to CFG, but stayed CFG-equivalent
- Eventually, it became generally accepted that natural languages are in fact not context free
- GPSG generalized to HPSG (Pollard & Sag 1994)

Feature Structures

- Break ‘atomic’ symbols like ‘V_intrans_sg’ into bundles of information
- Allows for the statement of cross-cutting generalizations

	3rd singular subject	plural subject
direct object NP	<i>denies</i>	<i>deny</i>
no direct object NP	<i>disappears</i>	<i>disappear</i>

Attribute value matrices

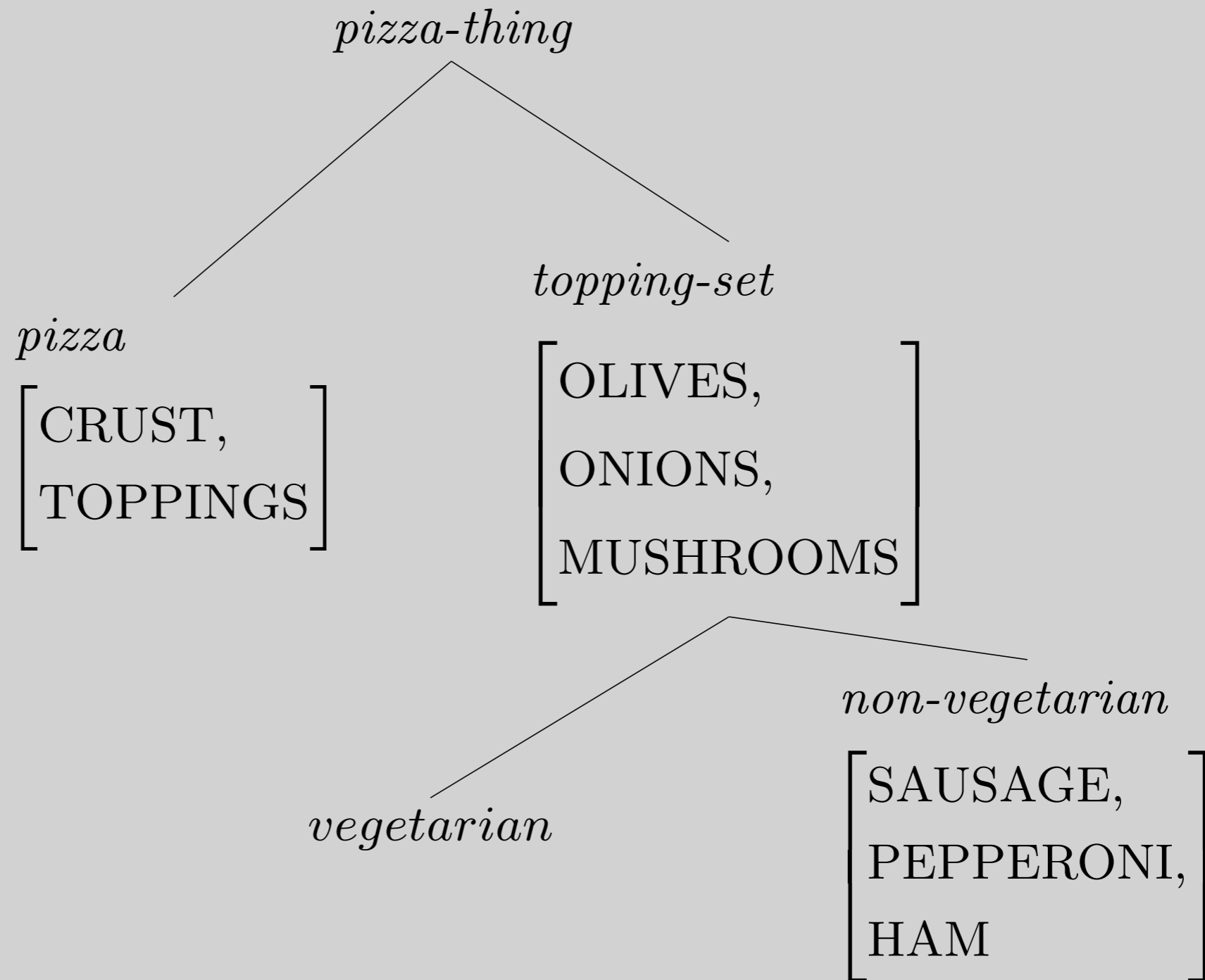
$$\begin{bmatrix} \text{FEATURE}_1 & \text{VALUE}_1 \\ \text{FEATURE}_2 & \text{VALUE}_2 \\ \dots & \\ \text{FEATURE}_n & \text{VALUE}_n \end{bmatrix}$$

- Values can be atomic symbols, or feature structures in their own right.

Unification

- Test whether two feature structures are compatible
- If so, find the most general feature structure that includes all information from both
- Section 15.2 shows unification of untyped feature structures
- Pizza examples (following) add in types (see 15.6)

A Pizza Type Hierarchy



TYPE	FEATURES/VALUES	IST
<i>pizza-thing</i>		
<i>pizza</i>	$\left[\begin{array}{ll} \text{CRUST} & \{ \text{thick, thin, stuffed} \} \\ \text{TOPPINGS} & \text{topping-set} \end{array} \right]$	<i>pizza-thing</i>
<i>topping-set</i>	$\left[\begin{array}{ll} \text{OLIVES} & \{ +, - \} \\ \text{ONIONS} & \{ +, - \} \\ \text{MUSHROOMS} & \{ +, - \} \end{array} \right]$	<i>pizza-thing</i>
<i>vegetarian</i>		<i>topping-set</i>
<i>non-vegetarian</i>	$\left[\begin{array}{ll} \text{SAUSAGE} & \{ +, - \} \\ \text{PEPPERONI} & \{ +, - \} \\ \text{HAM} & \{ +, - \} \end{array} \right]$	<i>topping-set</i>

Type Hierarchies

A type hierarchy....

- ... states what kinds of objects we claim exist (the types)
- ... organizes the objects hierarchically into classes with shared properties (the type hierarchy)
- ... states what general properties each kind of object has (the feature and feature value declarations).

Pizza Descriptions and Pizza Models

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \quad \text{thick} \\ \text{TOPPINGS} \quad \left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right] \end{array} \right]$$

How many pizza models (by definition, fully resolved) satisfy this description?

Answer: 2

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \quad \text{thick} \\ \text{TOPPINGS} \quad \left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right] \end{array} \right]$$

{<CRUST , thick> , <TOPPINGS , { <OLIVES ,
+ > , <ONIONS, +> , <MUSHROOMS, ->}>}

{<CRUST , thick> , <TOPPINGS , { <OLIVES ,
+ > , <ONIONS, +> , <MUSHROOMS, +>}>}

Pizza Descriptions and Pizza Models

pizza
CRUST thick
TOPPINGS $\left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right]$

Pizza Descriptions and Pizza Models

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \quad \text{thick} \\ \text{TOPPINGS} \quad \left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right] \end{array} \right]$$

How many pizzas-in-the-world do the pizza models correspond to?

Pizza Descriptions and Pizza Models

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \quad \text{thick} \\ \text{TOPPINGS} \quad \left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right] \end{array} \right]$$

How many pizzas-in-the-world do the pizza models correspond to?

Answer: A large, constantly-changing number.

Pizza Descriptions and Pizza Models

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \quad \text{thick} \\ \text{TOPPINGS} \quad \left[\begin{array}{l} \textit{vegetarian} \\ \text{OLIVES} \quad + \\ \text{ONIONS} \quad + \end{array} \right] \end{array} \right]$$

‘type’/‘token’ distinction
applies to sentences as well

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \right] \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ - \end{array} \end{array} \quad \& \quad \left[\begin{array}{l} \textit{pizza} \\ \text{TOPPINGS} \end{array} \right] \left[\begin{array}{l} \text{OLIVES} \\ \text{ONIONS} \end{array} \right] \begin{array}{l} + \\ + \end{array} \right]$$

Combining Constraints

<i>pizza</i>							
CRUST	thick						
TOPPINGS	<table><tr><td>OLIVES</td><td>+</td></tr><tr><td>ONIONS</td><td>+</td></tr><tr><td>HAM</td><td>-</td></tr></table>	OLIVES	+	ONIONS	+	HAM	-
OLIVES	+						
ONIONS	+						
HAM	-						

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \begin{array}{l} + \\ - \end{array} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thin} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{ONIONS} \end{array} \begin{array}{l} + \\ + \end{array} \end{array} \right]$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ - \end{array} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thin} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{ONIONS} \end{array} \right] \begin{array}{l} + \\ + \end{array} \end{array} \right] \\ = \emptyset$$

Combining Constraints

$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \right. \left. \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right. \left. \begin{array}{l} + \\ + \end{array} \right] \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \right. \left. \begin{array}{l} \text{thick} \\ \textit{vegetarian} \end{array} \right]$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ + \end{array} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \textit{vegetarian} \end{array} \right]$$

$$= \emptyset$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \right] \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ - \end{array} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \right] \begin{array}{l} \text{thick} \\ \textit{vegetarian} \end{array} \right]$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \left[\begin{array}{l} \text{OLIVES} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ - \end{array} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{CRUST} \\ \text{TOPPINGS} \end{array} \begin{array}{l} \text{thick} \\ \textit{vegetarian} \end{array} \right]$$

$$= \emptyset$$

A New Theory of Pizzas

pizza : $\left[\begin{array}{l} \text{CRUST} \quad \left\{ \text{thick , thin , stuffed} \right\} \\ \text{ONE-HALF} \quad \textit{topping-set} \\ \text{OTHER-HALF} \quad \textit{topping-set} \end{array} \right]$

Combining Constraints

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \end{array} \right] \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \begin{array}{l} + \\ - \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \begin{array}{l} - \\ + \end{array} \right]$$

=

Combining Constraints

$$\begin{array}{l} \left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \end{array} \right] \left[\begin{array}{l} \text{ONIONS} \quad + \\ \text{OLIVES} \quad - \end{array} \right] \quad \& \quad \left[\begin{array}{l} \textit{pizza} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \text{ONIONS} \quad - \\ \text{OLIVES} \quad + \end{array} \right] \\ \\ = \\ \left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \text{ONIONS} \quad + \\ \text{OLIVES} \quad - \\ \text{ONIONS} \quad - \\ \text{OLIVES} \quad + \end{array} \right] \end{array}$$

Identity Constraints (tags)

<i>pizza</i>					
CRUST	thin				
ONE-HALF	<table><tr><td>OLIVES</td><td>1</td></tr><tr><td>ONIONS</td><td>2</td></tr></table>	OLIVES	1	ONIONS	2
OLIVES	1				
ONIONS	2				
OTHER-HALF	<table><tr><td>OLIVES</td><td>1</td></tr><tr><td>ONIONS</td><td>2</td></tr></table>	OLIVES	1	ONIONS	2
OLIVES	1				
ONIONS	2				

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \right] \begin{array}{l} \boxed{1} \\ \boxed{1} \end{array} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \right] \begin{array}{l} + \\ - \end{array} \left. \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \text{MUSHROOMS} \\ \text{OLIVES} \end{array} \right] \begin{array}{l} - \\ - \end{array} \left. \right]$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \begin{array}{l} \boxed{1} \\ \boxed{1} \end{array} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \begin{array}{l} + \\ - \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{OTHER-HALF} \end{array} \left[\begin{array}{l} \text{MUSHROOMS} \\ \text{OLIVES} \end{array} \begin{array}{l} - \\ - \end{array} \right] \right]$$

=

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \begin{array}{l} \\ \boxed{1} \\ \boxed{1} \end{array} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \\ \text{MUSHROOMS} \end{array} \begin{array}{l} + \\ - \\ - \end{array} \right] \right]$$

Note

$$\left[\begin{array}{l} \textit{pizza} \\ \\ \text{ONE-HALF} \\ \\ \text{OTHER-HALF} \end{array} \right] \begin{array}{l} \\ \boxed{1} \\ \\ \boxed{1} \end{array} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \\ \text{MUSHROOMS} \end{array} \right] \begin{array}{l} + \\ - \\ - \end{array}$$

=

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \\ \text{OTHER-HALF} \end{array} \right] \begin{array}{l} \boxed{1} \\ \\ \boxed{1} \end{array} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \\ \text{MUSHROOMS} \end{array} \right] \begin{array}{l} + \\ - \\ - \end{array}$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \boxed{1} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \right] \\ \boxed{1} \textit{vegetarian} \end{array} \right] \begin{array}{l} + \\ + \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \end{array} \right] \left[\begin{array}{l} \text{SAUSAGE} \\ \text{HAM} \end{array} \right] \begin{array}{l} + \\ - \end{array} \right]$$

Combining Constraints

$$\left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \text{OTHER-HALF} \end{array} \right] \left[\begin{array}{l} \boxed{1} \left[\begin{array}{l} \text{ONIONS} \\ \text{OLIVES} \end{array} \right] \\ \boxed{1} \textit{vegetarian} \end{array} \right] \& \left[\begin{array}{l} \textit{pizza} \\ \text{ONE-HALF} \\ \left[\begin{array}{l} \text{SAUSAGE} \\ \text{HAM} \end{array} \right] \end{array} \right]$$

$$= \emptyset$$

Fixing the unwieldy grammar: Agreement

- Awkward CFG analyses

$S \rightarrow NP_sg VP_sg$

$S \rightarrow NP_pl VP_pl$

$NP_sg \rightarrow (Det_sg) Noun_sg$

$NP_pl \rightarrow (Det_pl) Noun_pl$

$VP_sg \rightarrow V_intrans_sg$

$VP_sg \rightarrow V_trans_sg NP_sg$

$VP_sg \rightarrow V_trans_sg NP_pl$

...

Fixing the unwieldy grammar: Agreement

- Better, with unification:

$S \rightarrow NP[AGR \boxed{1}] VP[AGR \boxed{1}]$
 $NP[AGR \boxed{1}] \rightarrow (Det[AGR \boxed{1}]) Noun[AGR \boxed{1}]$
 $VP[AGR \boxed{1}] \rightarrow V_intrans[AGR \boxed{1}]$
 $VP[AGR \boxed{1}] \rightarrow V_trans[AGR \boxed{1}] NP$
 $VP[AGR \boxed{1}] \rightarrow V_pp_trans[AGR \boxed{1}] PP$
 $VP[AGR \boxed{1}] \rightarrow V_ditrans[AGR \boxed{1}] NP NP$
 $VP[AGR \boxed{1}] \rightarrow V_pp_ditrans[AGR \boxed{1}] NP PP$
...

Fixing the unwieldy grammar: Subcategorization

$S \rightarrow NP[AGR \boxed{1}] VP[AGR \boxed{1}]$

$NP[AGR \boxed{1}] \rightarrow (Det[AGR \boxed{1}]) Noun[AGR \boxed{1}]$

$VP[AGR \boxed{1}] \rightarrow V[AGR \boxed{1}, SUBCAT \boxed{A}] \boxed{A}$

$V[AGR \text{ sg}, SUBCAT \langle \rangle] \rightarrow \text{sleeps}$

$V[AGR \text{ pl}, SUBCAT \langle \rangle] \rightarrow \text{sleep}$

$V[SUBCAT \langle \rangle] \rightarrow \text{slept}$

$V[AGR \text{ sg}, SUBCAT \langle NP \rangle] \rightarrow \text{sees}$

$V[AGR \text{ pl}, SUBCAT \langle NP \rangle] \rightarrow \text{see}$

$V[SUBCAT \langle NP \rangle] \rightarrow \text{saw}$

...

Examples

- *wh*-questions:

What did you find?

Tell me who you talked to

- relative clauses:

the item that I found

the guy who(m) I talked to

- topicalization:

The manual, I can't find

Chris, you should talk to.

- *easy*-adjectives:

My house is easy to find.

Pat is hard to talk to.

What these have in common

- There is a ‘gap’: nothing following *find* and *to*, even though both normally require objects.
- Something that fills the role of the element missing from the gap occurs at the beginning of the clause.
- We use topicalization and *easy*-adjectives to illustrate:

The manual, *I can't find* _____

Chris *is easy to talk to* _____

Gaps and their fillers can be far apart:

- *The solution to this problem, Pat said that someone claimed you thought I would never find_____.*
 - *Chris is easy to consider it impossible for anyone but a genius to try to talk to_____.*
- ☞ That's why we call them “long distance dependencies”

Fillers often have syntactic properties associated with their gaps

Him, I haven't met ____.

**He, I haven't met ____.*

The scissors, Pat told us _____ were missing.

**The scissors, Pat told us _____ was missing.*

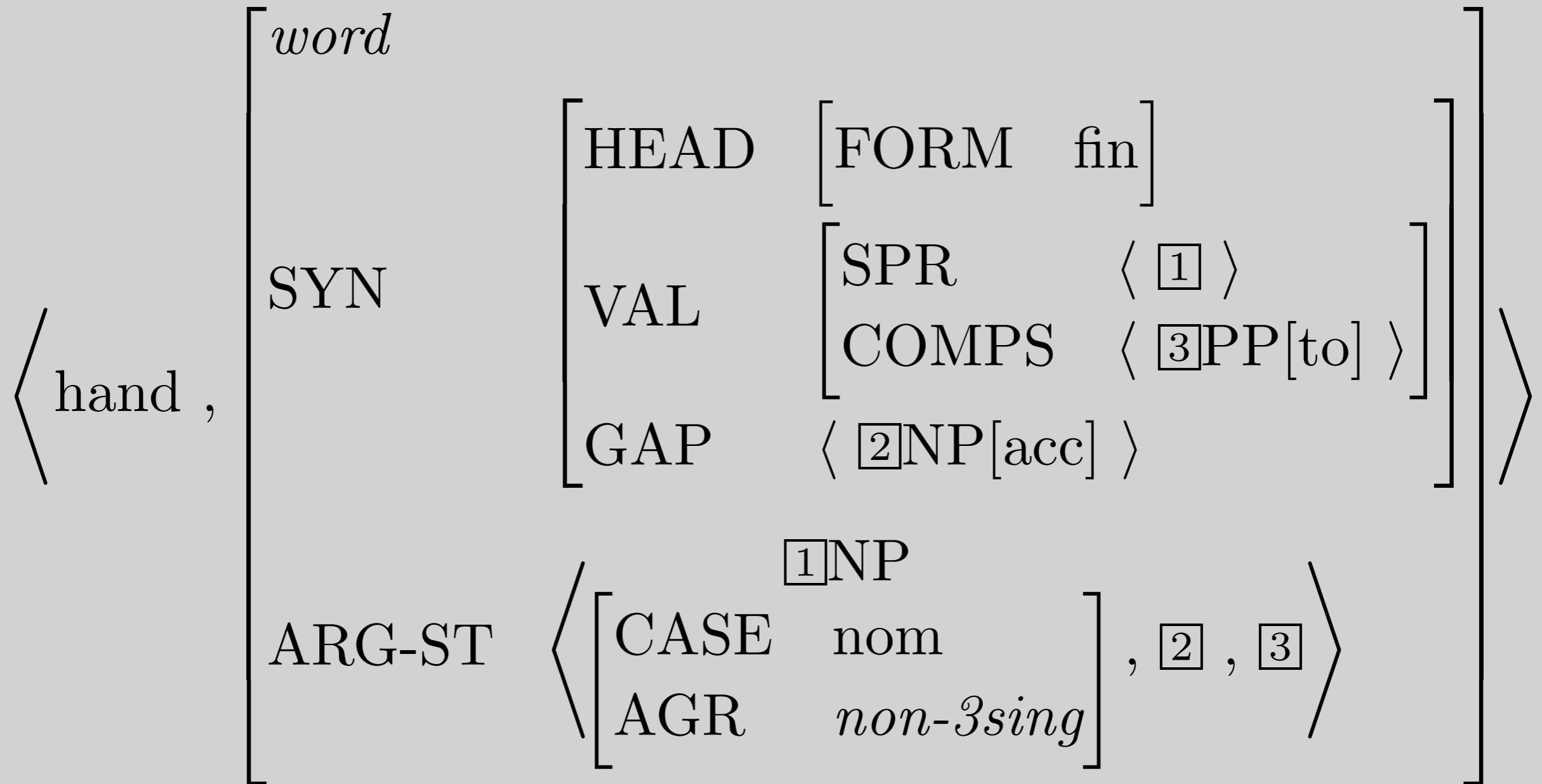
On Pat, you can rely ____.

**To Pat, you can rely ____.*

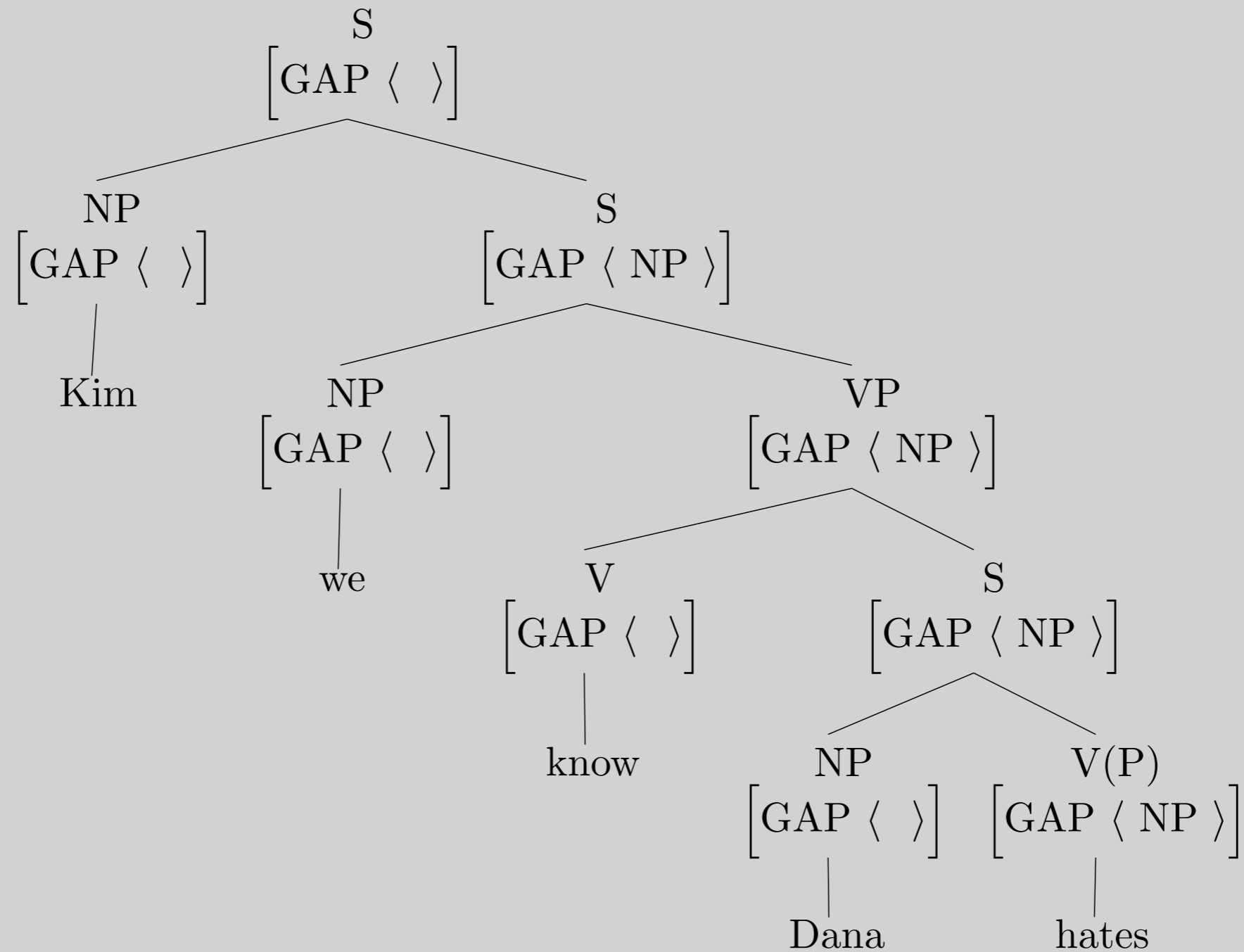
Very Rough Sketch of Our Approach

- A feature GAP records information about a missing constituent.
- The GAP value is passed up the tree by a new principle.
- A new grammar rule expands S as a filler followed by another S whose GAP value matches the filler.
- Caveat: Making the details of this general idea work involves several complications.

A Word with a Non-Empty GAP Value



How We Want GAP to Propagate



The Head-Filler Rule

$$[phrase] \rightarrow \boxed{1} \left[\text{GAP} \quad \langle \rangle \right] \mathbf{H} \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \textit{verb} \\ \text{FORM} \quad \textit{fin} \end{array} \right] \\ \text{VAL} \quad \left[\begin{array}{l} \text{SPR} \quad \langle \rangle \\ \text{COMPS} \quad \langle \rangle \end{array} \right] \\ \text{STOP-GAP} \quad \langle \boxed{1} \rangle \\ \text{GAP} \quad \langle \boxed{1} \rangle \end{array} \right]$$

Overview

- Problems with CFG
- Feature structures
- Unification
- Agreement
- Subcategorization
- Long-distance Dependencies
- Reading questions

Reading questions

- I know what regular and non-regular grammars are, but what makes a grammar context-free?
- There is a distinction made between right and left linear grammar, but I don't really understand why this distinction is too important; is it just for the sake of writing proofs? Shouldn't there also be consideration for terminals appearing on both the left, right and the middle?

Reading questions

- I was not really clear about the two forms of specified constraints:
 - $\langle \text{Bi feature path} \rangle = \text{Atomic value}$
 - $\langle \text{Bi feature path} \rangle = \langle \text{Bi feature path} \rangle$
- How should I understand these two forms? Also, "the notation $\langle \text{Bi feature path} \rangle$ denotes a feature path through the feature structure associated with the Bi component of the context-free part of the rule", this sentence seems helpful for understanding the forms but it is really confusing to me, how should I understand it?

Reading questions

- I'm confused as to how the feature system saves you very much from a proliferation of grammar rules. Don't you still have to write separate grammar rules for each form of each feature?
- Does the order of the features in the attribute-value matrix matter? In the readings, some have [PERSON 3, then NUMBER sg] while others have [NUMBER sg, then PERSON 3]. There is one example on pg 495 where after unification, [PERSON 3, NUMBER sg] becomes [NUMBER sg, PERSON 3]. Do the order differ because of which is more significant or are they just random?
- I'm confused as to how auxiliaries function in these agreement patterns. It seems that auxiliaries take away some of the agreement functionality of the main verbs- in the sentence "He is running," only "is" has 3rd / sg agreement features, not the main verb "running". How would it look for full agreement functionality to be written out for a sentence with an auxiliary?

Reading questions

- I was wondering how feature structures tend to vary with the "type" of language. Would word order based languages lead to more, smaller feature structures, and polysynthetic languages lead to fewer, larger feature structures?
- How do feature structures work in analyzing transcriptions of natural speech? Often times, when we talk to one another, we'll leave out dependencies because they are included in wider context. Can feature structures account for this?

Reading questions

- Is the gap list in long distance dependencies the reason we saw the stacked NP's and such in the parse tree we saw in class on Tuesday?
- How are feature structures encoded? Do they have to be designed by hand for each word? Have there been attempts to use machine learning to deduce argument structures for verbs, for example?