Lecture 5: Parsing with constraint-based grammars

Providing a more adequate treatment of syntax than simple CFGs: replacing the atomic categories by more complex data structures.

1. Problems with simple CFG encoding: agreement, subcategorisation, long distance dependencies.
2. Feature structures (informally)
3. Encoding agreement
4. Parsing with feature structures
5. Feature stuctures more formally
6. Encoding subcategorisation
7. Interface to morphology

Deficiencies in atomic category CFGs

Overgeneration with lecture 4 grammar:

agreement e.g. subject verb agreement
they fish, *it fishes*, **it fish*, **they fishes*

case pronouns (and maybe *who/whom*)
they like them, **they like they*

Expanding symbols:

```
S -> NP-sg-subj VP-sg
S -> NP-pl-subj VP-pl
VP-sg -> V-sg NP-sg-obj
VP-sg -> V-sg NP-pl-obj
VP-pl -> V-pl NP-sg-obj
VP-pl -> V-pl NP-pl-obj
NP-sg-subj -> he
NP-sg-obj -> him
NP-sg-subj -> fish
NP-pl-subj -> fish
NP-sg-obj -> fish
NP-pl-obj -> fish
```

Intuitive solution for case and agreement

BUT: very large grammar, misses generalizations, no way of saying when we don't care about agreement.

- Have separate slots (features) for case (CASE) and agreement (AGR)
- Allow slot values for CASE to be subj, obj or unspecified
- Allow slot values for AGR to be sg, pl or unspecified
- Subjects must have the same value for AGR as their verbs
- Subjects have CASE subj, objects have CASE obj

can (noun) $\begin{bmatrix} \text{CASE} \; [\,] \\ \text{AGR} \; \textbf{sg} \end{bmatrix}$

fish (noun) $\begin{bmatrix} \text{CASE} \; [\,] \\ \text{AGR} \; [\,] \end{bmatrix}$

she $\begin{bmatrix} \text{CASE} \; \textbf{subj} \\ \text{AGR} \; \textbf{sg} \end{bmatrix}$

them $\begin{bmatrix}\text{CASE } \textbf{obj} \\ \text{AGR } \textbf{pl}\end{bmatrix}$

Subcategorization

- intransitive vs transitive etc
- e.g., verbs have different numbers and types of syntactic arguments:

  **Kim adored*
  **Kim gave Sandy*
  **Kim adored to sleep*
  Kim liked to sleep
  **Kim devoured*
  Kim ate

- Subcategorization is correlated with semantics, but not determined by it.

Overgeneration:

they fish fish it
(S (NP they) (VP (V fish) (VP (V fish) (NP it))))

- Informally: need slots on the verbs for their syntactic arguments.

Long-distance dependencies

1. which problem did you say you don't understand?
2. who do you think Kim asked Sandy to hit?
3. which kids did you say were making all that noise?

'gaps' (underscores below)

1. which problem did you say you don't understand _?
2. who do you think Kim asked Sandy to hit _?
3. which kids did you say _ were making all that noise?

In 3, the verb *were* shows plural agreement.

> * what kid did you say _ were making all that noise?

The gap filler has to be plural.

- Informally: need a 'gap' slot which is to be filled by something that itself has features.

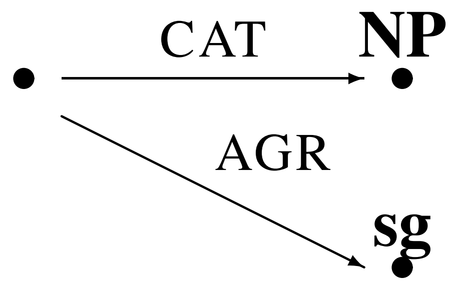
Feature structures

1. Features like AGR with simple values (sg, pl): **atomic-valued**
2. Unspecified values possible on features: compatible with any value.
3. Values for features for subcat and gap themselves have features: **complex-valued**
   - **path**: a sequence of features
4. Method of specifying two paths are the same: **reentrancy**
5. Unification: combining two feature structures, retaining all information from each, or fail if information is incompatible.

- Feature structures are singly-rooted directed acyclic graphs, with arcs labelled by features and terminal nodes associated with values.
- Rules relate FSs — i.e. lexical entries and phrases are represented as FSs
- Rule application by unification

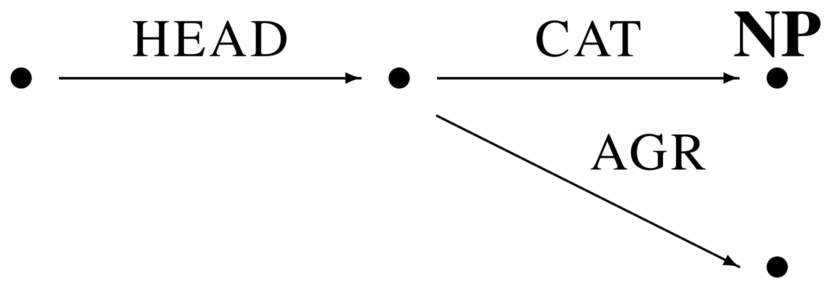
Graphs and AVMs

Example 1:



Here, CAT and AGR are atomic-valued features. **NP** and **sg** are values.

Example 2:



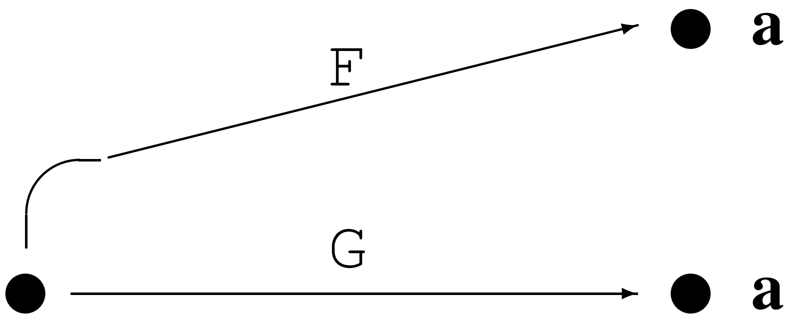
HEAD is complex-valued, AGR is unspecified.

AVM notation:

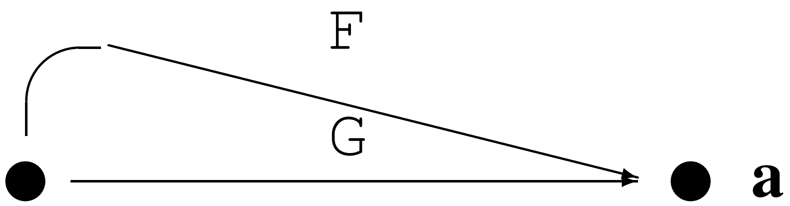
Example 1: $\begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \textbf{sg} \end{bmatrix}$

Example 2: $\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & [\,] \end{bmatrix} \end{bmatrix}$

Reentrancy



$$\begin{bmatrix} \text{F} & \mathbf{a} \\ \text{G} & \mathbf{a} \end{bmatrix}$$



$$\begin{bmatrix} \text{F} & \boxed{0}\ \mathbf{a} \\ \text{G} & \boxed{0} \end{bmatrix}$$

CFG with agreement

```
S -> NP-sg VP-sg
S -> NP-pl VP-pl
VP-sg -> V-sg NP-sg
VP-sg -> V-sg NP-pl
VP-pl -> V-pl NP-sg
VP-pl -> V-pl NP-pl
V-pl -> like
V-sg -> likes
NP-sg -> it
NP-pl -> they
NP-sg -> fish
NP-pl -> fish
```

FS grammar fragment encoding agreement

Grammar rules

Rule1 $\begin{bmatrix} \text{CAT} & \textbf{S} \\ \text{AGR} & \boxed{1} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1} \end{bmatrix}$

Rule2 $\begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & \textbf{V} \\ \text{AGR} & \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & [\,] \end{bmatrix}$

Lexicon:

;;; noun phrases

they $\begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \textbf{pl} \end{bmatrix}$

fish $\left[\begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & [\,] \end{bmatrix}\right]$

it $\begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \textbf{sg} \end{bmatrix}$

;;; verbs

like $\begin{bmatrix} \text{CAT} & \textbf{V} \\ \text{AGR} & \textbf{pl} \end{bmatrix}$

likes $\begin{bmatrix} \text{CAT} & \textbf{V} \\ \text{AGR} & \textbf{sg} \end{bmatrix}$

Root structure:

$\begin{bmatrix} \text{CAT} & \textbf{S} \end{bmatrix}$

Parsing (informally)

they like it

- The lexical structures for *like* and *it* are unified with the corresponding structure to the right hand side of rule 2 (unifications succeed).
- The structure corresponding to the mother of the rule is:

  $$\begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \textbf{pl} \end{bmatrix}$$

- Unifies with the rightmost daughter position of rule 1.
- *they* is unified with the leftmost daughter.
- Result unifies with root structure

Rules as FSs

Rules have features MOTHER, DTR1, DTR2 ... DTRN.

Rule2 (informally):

$$\begin{bmatrix}\text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1}\end{bmatrix} \rightarrow \begin{bmatrix}\text{CAT} & \textbf{V} \\ \text{AGR} & \boxed{1}\end{bmatrix}, \begin{bmatrix}\text{CAT} & \textbf{NP} \\ \text{AGR} & [\,]\end{bmatrix}$$

actually:

$$\begin{bmatrix}\text{MOTHER} & \begin{bmatrix}\text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1}\end{bmatrix} \\ \text{DTR1} & \begin{bmatrix}\text{CAT} & \textbf{V} \\ \text{AGR} & \boxed{1}\end{bmatrix} \\ \text{DTR2} & \begin{bmatrix}\text{CAT} & \textbf{NP} \\ \text{AGR} & [\,]\end{bmatrix}\end{bmatrix}$$

Rule 2 application

like unified with the value of DTR1 in rule 2.

$$\begin{bmatrix} \text{MOTHER} & \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1}\,\textbf{pl} \end{bmatrix} \\ \text{DTR1} & \begin{bmatrix} \text{CAT} & \textbf{V} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & [\,] \end{bmatrix} \end{bmatrix}$$

it is unified with the value for DTR2:

$$\begin{bmatrix} \text{MOTHER} & \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1}\,\textbf{pl} \end{bmatrix} \\ \text{DTR1} & \begin{bmatrix} \text{CAT} & \textbf{V} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \textbf{sg} \end{bmatrix} \end{bmatrix}$$

Rule 1 application

MOTHER value acts as the DTR2 of Rule 1

$$\begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \textbf{pl} \end{bmatrix}$$

is unified with the DTR2 value of:

$$\begin{bmatrix} \text{MOTHER} & \begin{bmatrix} \text{CAT} & \textbf{S} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR1} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

This gives:

$$\begin{bmatrix} \text{MOTHER} & \begin{bmatrix} \text{CAT} & \textbf{S} \\ \text{AGR} & \boxed{1}\,\textbf{pl} \end{bmatrix} \\ \text{DTR1} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Rule 1 application continued

The FS for *they* is:

$$\begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \textbf{pl} \end{bmatrix}$$

The unification of this with the value of DTR1 succeeds but adds no new information:

$$\begin{bmatrix} \text{MOTHER} & \begin{bmatrix} \text{CAT} & \textbf{S} \\ \text{AGR} & \boxed{1}\,\textbf{pl} \end{bmatrix} \\ \text{DTR1} & \begin{bmatrix} \text{CAT} & \textbf{NP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & \textbf{VP} \\ \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Properties of FSs

Connectedness and unique root A FS must have a unique root node: apart from the root node, all nodes have one or more parent nodes.

Unique features Any node may have zero or more arcs leading out of it, but the label on each (that is, the feature) must be unique.

No cycles No node may have an arc that points back to the root node or to a node that intervenes between it and the root node.

Values A node which does not have any arcs leading out of it may have an associated atomic value.

Finiteness A FS must have a finite number of nodes.

Subsumption

Feature structures are ordered by information content — FS1 *subsume*s FS2 if FS2 carries extra information.

FS1 subsumes FS2 if and only if the following conditions hold:

Path values For every path P in FS1 there is a path P in FS2. If P has a value t in FS1, then P also has value t in FS2.

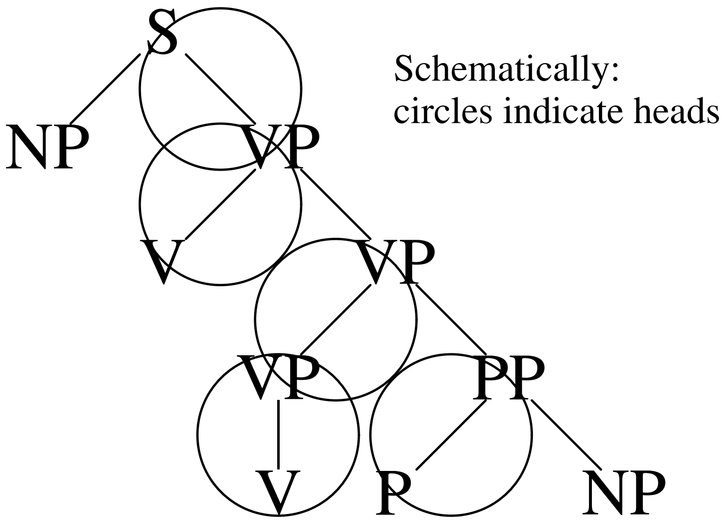
Path equivalences Every pair of paths P and Q which are reentrant in FS1 (i.e., which lead to the same node in the graph) are also reentrant in FS2.

Unification

The unification of two FSs FS1 and FS2 is the most general FS which is subsumed by both FS1 and FS2, if it exists.

Grammar with subcategorisation

HEAD — information shared between a lexical entry and the dominating phrases of the same category (agreement and category)



COMP — subcategorization: arguments that come after the lexical entry in English (e.g., verbs' objects)
Rule 1 unifies the second dtr with the COMP value of the first.

SPR — arguments that come before the lexical entry in English (e.g. verbs' subjects)
Rule 2 unifies the first daughter with the SPR value of the second.

Example rule application: *they fish*

Lexical entry for fish:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{CAT} & \textbf{verb}\\ \text{AGR} & \textbf{pl}\end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \begin{bmatrix}\text{HEAD} & \begin{bmatrix}\text{CAT} & \textbf{noun}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Rule 2:

$$\begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled}\end{bmatrix} \rightarrow$$

$$\boxed{2}\begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{AGR} & \boxed{3}\end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled}\end{bmatrix}, \begin{bmatrix} \text{HEAD} & \boxed{1}\begin{bmatrix}\text{AGR} & \boxed{3}\end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \boxed{2}\end{bmatrix}$$

unification with second dtr position gives:

$$\begin{bmatrix} \text{HEAD } \boxed{1} \begin{bmatrix} \text{CAT } \textbf{verb} \\ \text{AGR } \boxed{3}\ \textbf{pl} \end{bmatrix} \\ \text{COMP } \textbf{filled} \\ \text{SPR } \textbf{filled} \end{bmatrix} \rightarrow \boxed{2} \begin{bmatrix} \text{HEAD } \begin{bmatrix} \text{CAT } \textbf{noun} \\ \text{AGR } \boxed{3} \end{bmatrix} \\ \text{COMP } \textbf{filled} \\ \text{SPR } \textbf{filled} \end{bmatrix}, \begin{bmatrix} \text{HEAD } \boxed{1} \\ \text{COMP } \textbf{filled} \\ \text{SPR } \boxed{2} \end{bmatrix}$$

Lexical entry for *they*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{noun} \\ \text{AGR} & \textbf{pl} \end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled} \end{bmatrix}$$

unify this with first dtr position:

$$\begin{bmatrix} \text{HEAD} & \boxed{1} \begin{bmatrix} \text{CAT} & \textbf{verb} \\ \text{AGR} & \boxed{3}\ \textbf{pl} \end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled} \end{bmatrix} \rightarrow$$

$$\boxed{2} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{noun} \\ \text{AGR} & \boxed{3} \end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled} \end{bmatrix}, \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \boxed{2} \end{bmatrix}$$

Root is:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{verb} \end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \textbf{filled} \end{bmatrix}$$

Mother structure unifies with root, so valid.

Parsing with feature structure grammars

- Naive algorithm: standard chart parser with modified rule application
- Rule application:
  1. copy rule
  2. copy daughters (lexical entries or FSs associated with edges)
  3. unify rule and daughters
  4. if successful, add new edge to chart with rule FS as category
- Efficient algorithms reduce copying.
- Packing involves subsumption.
- Probabilistic FS grammars are complex.

Templates

Capture generalizations in the lexicon:

fish INTRANS_VERB
sleep INTRANS_VERB
snore INTRANS_VERB

INTRANS_VERB

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{verb} \\ \text{AGR} & \textbf{pl} \end{bmatrix} \\ \text{COMP} & \textbf{filled} \\ \text{SPR} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAT} & \textbf{noun} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Interface to morphology

Associate inflectional affixes with templates.

s `PLURAL_NOUN`

`PLURAL_NOUN` $\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CAT} & \textbf{noun} \\ \text{AGR} & \textbf{pl} \end{bmatrix} \end{bmatrix}$

stem is:

$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CAT} & \textbf{noun} \\ \text{AGR} & [\,] \end{bmatrix} \\ \text{COMP} \;\; \textbf{filled} \\ \text{SPR} \;\; \textbf{filled} \end{bmatrix}$$

unify stem with affix template:

$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CAT} & \textbf{noun} \\ \text{AGR} & \textbf{pl} \end{bmatrix} \\ \text{COMP} \;\; \textbf{filled} \\ \text{SPR} \;\; \textbf{filled} \end{bmatrix}$$

Unification failure with verbs etc.