Parsing Beyond Context-Free Grammars: Tree Adjoining Grammars

Laura Kallmeyer
Heinrich-Heine-Universität Düsseldorf

Winter 2021/22

Overview

1. **Adjunction and substitution**
2. **Tree Adjoining Grammar**
3. **Adjunction constraints**
4. **Derivation trees**
5. **TAG and natural languages**
6. **Feature Structure Based TAG**

[Kal10]

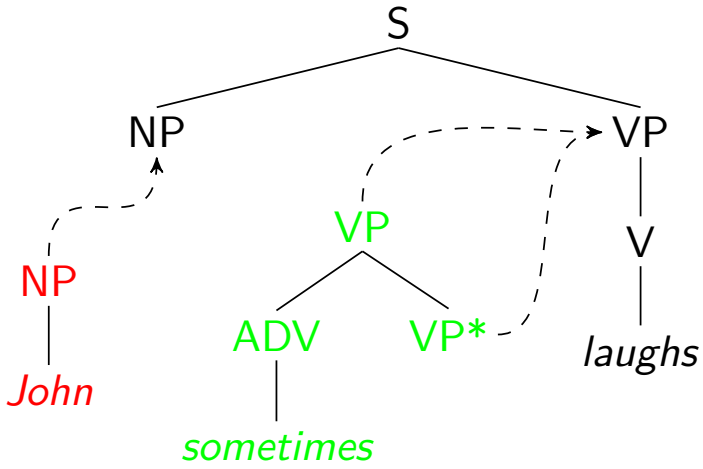
Adjunction and substitution (1)

Tree Adjoining Grammars (TAG) [JLT75, JS97]: Tree-rewriting system: set of *elementary* trees with two operations:

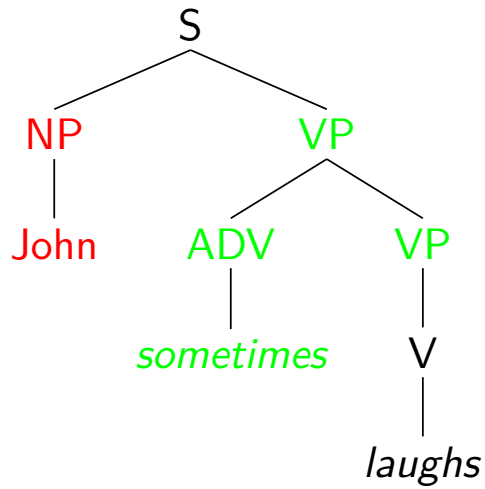
- adjunction: replacing an internal node with a new tree.
  The new tree is an auxiliary tree and has a special leaf, the foot node.
- substitution: replacing a leaf with a new tree.
  The new tree is an initial tree

Adjunction and substitution (2)

(1) John sometimes laughs



derived tree:



Adjunction and substitution (3)

Definition 1 (Auxiliary and initial trees)

1. An auxiliary tree is a syntactic tree $\langle V, E, r \rangle$ such that there is a unique leaf f marked as foot node with $l(r) = l(f)$. We write this tree as $\langle V, E, r, f \rangle$.
2. An initial tree is a non-auxiliary syntactic tree.

As a convention, the foot node is marked with a "*".

Adjunction and substitution (4)

Definition 2 (Substitution)

Let $\gamma = \langle V, E, r \rangle$ and $\gamma' = \langle V', E', r' \rangle$ be syntactic trees with $V \cap V' = \emptyset$ and $v \in V$. $\gamma[v, \gamma']$, the result of substituting γ' into γ at node v is defined as follows:

- if either v is not a leaf, or v is a foot, or $l(v) \neq l(r')$, then $\gamma[v, \gamma']$ is undefined;
- otherwise, $\gamma[v, \gamma'] = \langle V'', E'', r'' \rangle$ with $V'' = V \cup V' \setminus \{v\}$ and $E'' = (E \setminus \{\langle v_1, v_2 \rangle \mid v_2 = v\}) \cup E' \cup \{\langle v_1, r' \rangle \mid \langle v_1, v \rangle \in E\}$. Furthermore, $v_1 \prec v_2$ in $\gamma[v, \gamma']$ iff either $v_1 \prec v_2$ in γ or $v_1 \prec v_2$ in γ' or $v_1 \in V'$ and $v \prec v_2$ in γ or $v_2 \in V'$ and $v_1 \prec v$ in γ.

A leaf that is not a foot node and that has a non-terminal label is called a substitution node.

Adjunction and substitution (5)

Definition 3 (Adjunction)

Let $\gamma = \langle V, E, r \rangle$ be a syntactic tree, $\gamma' = \langle V', E', r', f \rangle$ an auxiliary tree and $v \in V$. $\gamma[v, \gamma']$, the result of *adjoining* γ' into γ at node v is defined as follows:

- if $l(v) \neq l(r')$, then $\gamma[v, \gamma']$ is undefined;
- otherwise, $\gamma[v, \gamma'] = \langle V'', E'', r'' \rangle$ with $V'' = V \cup V' \setminus \{v\}$ and $E'' = (E \setminus \{\langle v_1, v_2 \rangle \mid v_1 = v \text{ or } v_2 = v\}) \cup E' \cup \{\langle v_1, r' \rangle \mid \langle v_1, v \rangle \in E\} \cup \{\langle f, v_2 \rangle \mid \langle v, v_2 \rangle \in E\}$.

Tree Adjoining Grammar (1)

Definition 4 (Tree Adjoining Grammar)

A Tree Adjoining Grammar (TAG) is a tuple $G = \langle N, T, S, I, A \rangle$ such that

- T and N are disjoint alphabets, the terminals and nonterminals,
- $S \in N$ is the start symbol,
- I is a finite set of initial trees, and
- A is a finite set of auxiliary trees.

The trees in $I \cup A$ are called elementary trees.
G is lexicalized iff each elementary tree has at least one leaf with a terminal label.

Tree Adjoining Grammar (2)

- Every elementary tree is considered a derived tree in a TAG.
- In every derivation step, we pick a fresh instance of an elementary tree from the grammar and we add derived trees (by substitution or adjunction) to certain nodes in this tree.
- The trees in the tree language are the derived initial trees with root label S.

Tree Adjoining Grammar (3)

Definition 5 (Derived tree)

Let $G = \langle N, T, S, I, A \rangle$ be a TAG.

1. Every instance γ of a $\gamma_e \in I \cup A$ is a derived tree in G.
2. For pairwise disjoint $\gamma_1, \ldots, \gamma_n, \gamma$ such that $\gamma_1, \ldots, \gamma_n$ are derived trees in G $(1 \leq i \leq n)$ and $\gamma = \langle V, E, r \rangle$ is an instance of a $\gamma_e \in I \cup A$ such that $v_1, \ldots, v_n \in V$ are pairwise different: if $\gamma' = \gamma[v_1, \gamma_1] \ldots [v_n, \gamma_n]$ is defined then γ' is a derived tree in G.
3. These are all derived trees in G.

Note that this definition adopts a bottom-up perspective: derived trees are added to elementary trees.

Tree Adjoining Grammar (4)

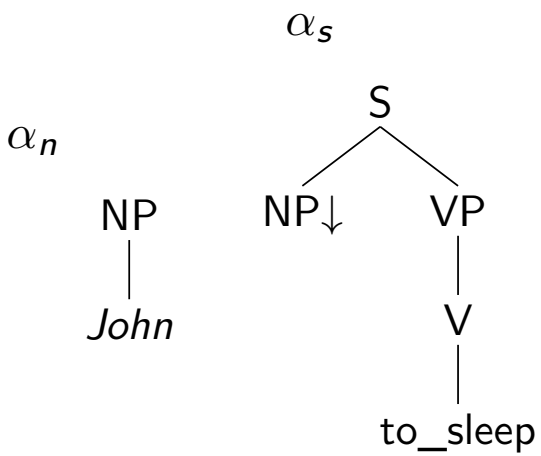
Definition 6 (TAG language)

Let $G = \langle N, T, S, I, A \rangle$ be a TAG.

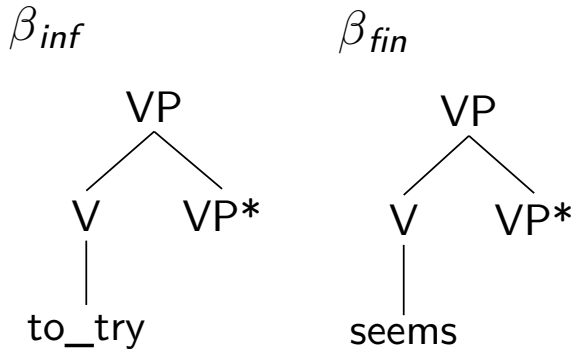
1. The tree language of G is the set of all derived trees $\gamma = \langle V, E, r \rangle$ in G such that
   - $l(r) = S$, and
   - $l(v) \in T \cup \{\varepsilon\}$ for every leaf $v \in V$.
2. The string language of G is $\{w \mid \text{there is a } \gamma \in L_T(G) \text{ such that } w = \mathit{yield}(\gamma)\}$.

Tree Adjoining Grammar (5)

Initial trees:



Auxiliary trees:



Adjunction constraints (1)

TAGs as defined above are more powerful than CFG but they cannot generate the copy language.

In order to increase the expressive power, adjunction constraints are introduced that specify for each node

1. whether adjunction is mandatory and
2. which trees can be adjoined.

For a given node,

1. the function f_{OA} specifies whether adjunction is obligatory (value 1) or not (value 0) and
2. the function f_{SA} gives the set of auxiliary trees that can be adjoined.

Adjunction constraints (2)

Definition 7 (TAG with adjunction constraints)

A TAG with adjunction constraints is a tuple $G = \langle N, T, S, I, A, f_{OA}, f_{SA} \rangle$ where

- $\langle N, T, S, I, A \rangle$ is a TAG as defined above and
- $f_{OA} : \{v \mid v \text{ is a node in some } \gamma \in I \cup A\} \to \{0, 1\}$ and
  $f_{SA} : \{v \mid v \text{ is a node in some } \gamma \in I \cup A\} \to \mathcal{P}(A)$
  where $\mathcal{P}(A)$ is the set of subsets of A are functions such that $f_{OA}(v) = 0$ and $f_{SA}(v) = \emptyset$ for every leaf v.

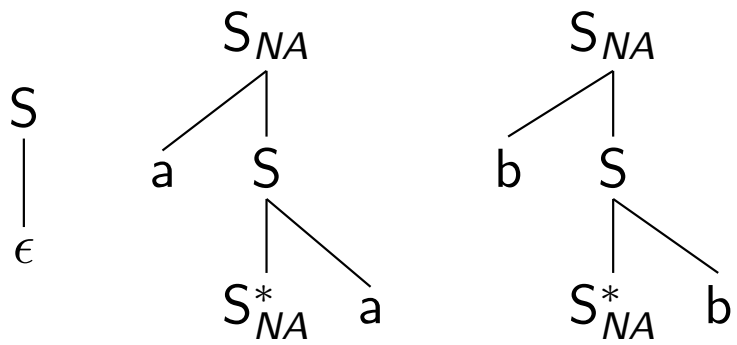
Adjunction constraints (3)

Three types of constraints are distinguished:

- A node v with $f_{OA}(v) = 1$ is said to carry a obligatory adjunction (OA) constraint.
- A node v with $f_{OA}(v) = 0$ and $f_{SA}(v) = \emptyset$ is said to carry a null adjunction (NA) constraint.
- A node v with $f_{OA}(v) = 0$ and $f_{SA}(v) \neq \emptyset$ and $f_{SA}(v) \neq A$ is said to carry a selective adjunction (SA) constraint.

Adjunction constraints (4)

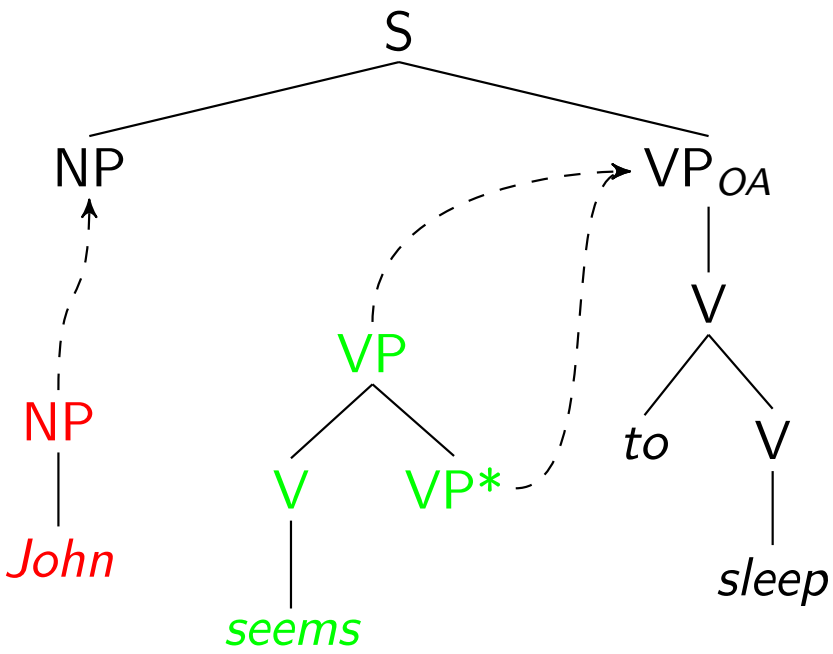
Example: TAG for the copy language:



Adjunction constraints (5)

Example:

(2) John seems to sleep



Adjunction constraints (6)

Definition 8 (Derived tree)

Let $G = \langle N, T, S, I, A, f_{OA}, f_{SA} \rangle$ be a TAG with adjunction constraints.

1. Every instance of a $\gamma_e \in I \cup A$ is a derived tree obtained from γ_e in G.
2. For pairwise disjoint $\gamma_1, \ldots, \gamma_n, \gamma$ such that a) $\gamma_1, \ldots, \gamma_n$ are derived trees obtained from $\gamma_1^e, \ldots, \gamma_n^e$ in G respectively, and b) $\gamma = \langle V, E, r \rangle$ is an instance of a $\gamma_e \in I \cup A$ such that $v_1, \ldots, v_n \in V$ are pairwise different nodes: If
   - $\gamma' = \gamma[v_1, \gamma_1] \ldots [v_n, \gamma_n]$ is defined, and
   - $\gamma_i^e \in f_{SA}(v_i)$ for all $1 \leq i \leq n$ with γ_i^e an auxiliary tree

   then γ' is a derived tree obtained from γ_e in G
3. These are all derived trees in G.

Adjunction constraints (7)

Definition 9 (Tree language)

Let $G = \langle N, T, S, I, A, f_{OA}, f_{SA} \rangle$ be a TAG with adjunction constraints. The tree language of G is the set of all derived trees $\gamma = \langle V, E, r \rangle$ in G such that

- $l(r) = S$,
- $f_{OA}(v) = 0$ for all $v \in V$, and
- $l(v) \in T \cup \{\varepsilon\}$ for every leaf $v \in V$.

In the following, whenever we use the term "TAG", this means "TAG with adjunction constraints".

Derivation trees (1)

TAG derivations are described by derivation trees:
For each derivation in a TAG there is a corresponding derivation tree.
This tree contains

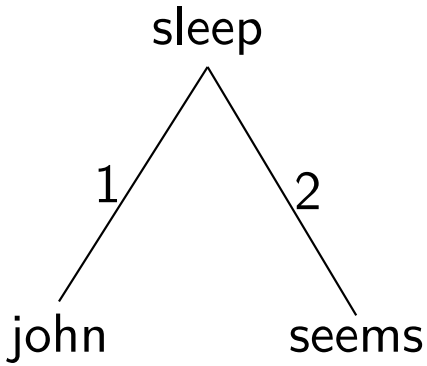
- nodes for all elementary trees used in the derivation, and
- edges for all adjunctions and substitutions performed throughout the derivation.

Whenever an elementary tree γ was attached to the node at address p in the elementary tree γ', there is an edge from γ' to γ labeled with p.

Derivation trees (2)

Example:

derivation tree for the derivation of (2) *John seems to sleep*



TAG and natural languages (1)

Important features of TAG when used for natural languages:

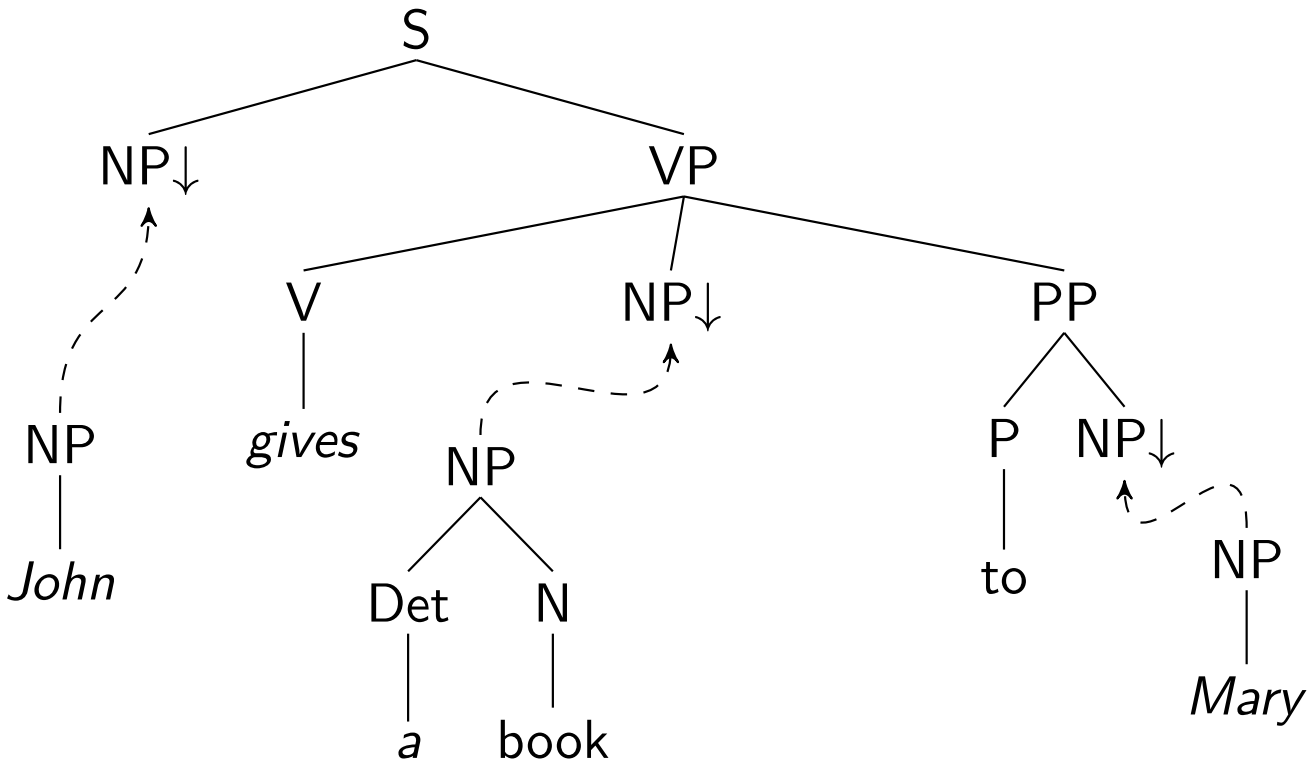
- Grammar is lexicalized
- Recursive parts are put into separate elementary trees that can be adjoined (Factoring of recursion, FR)
- The elementary tree of a lexical predicate contains slots for all arguments of the predicate, for nothing more (Elementary tree minimality).
- Elementary trees can be arbitrarily large, in particular (because of FR) they can contain elements that are far apart in the final derived tree (Extended domain of locality)

For natural language syntax and TAG see
[Kro87, Abe88, Abe02, Fra02, XTA01].

TAG and natural languages (2)

Example:

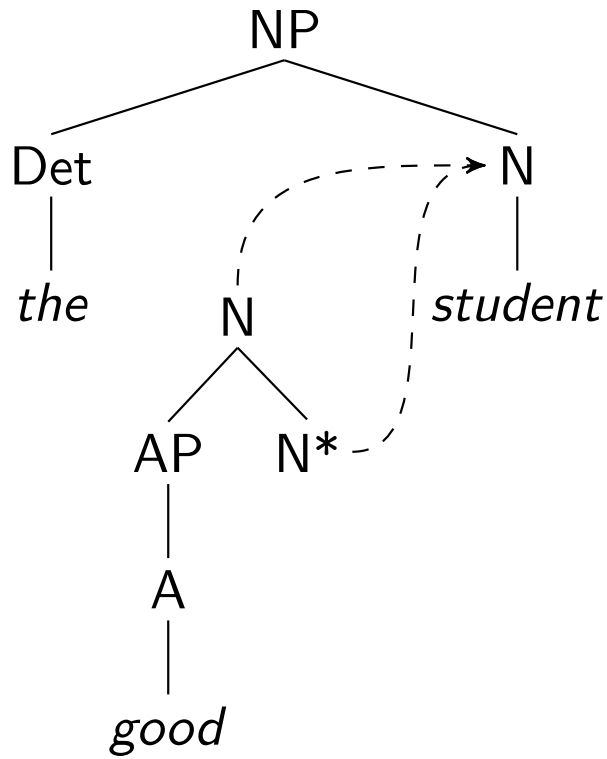
(3) John gives book to Mary



TAG and natural languages (3)

Example:

(4) the good student

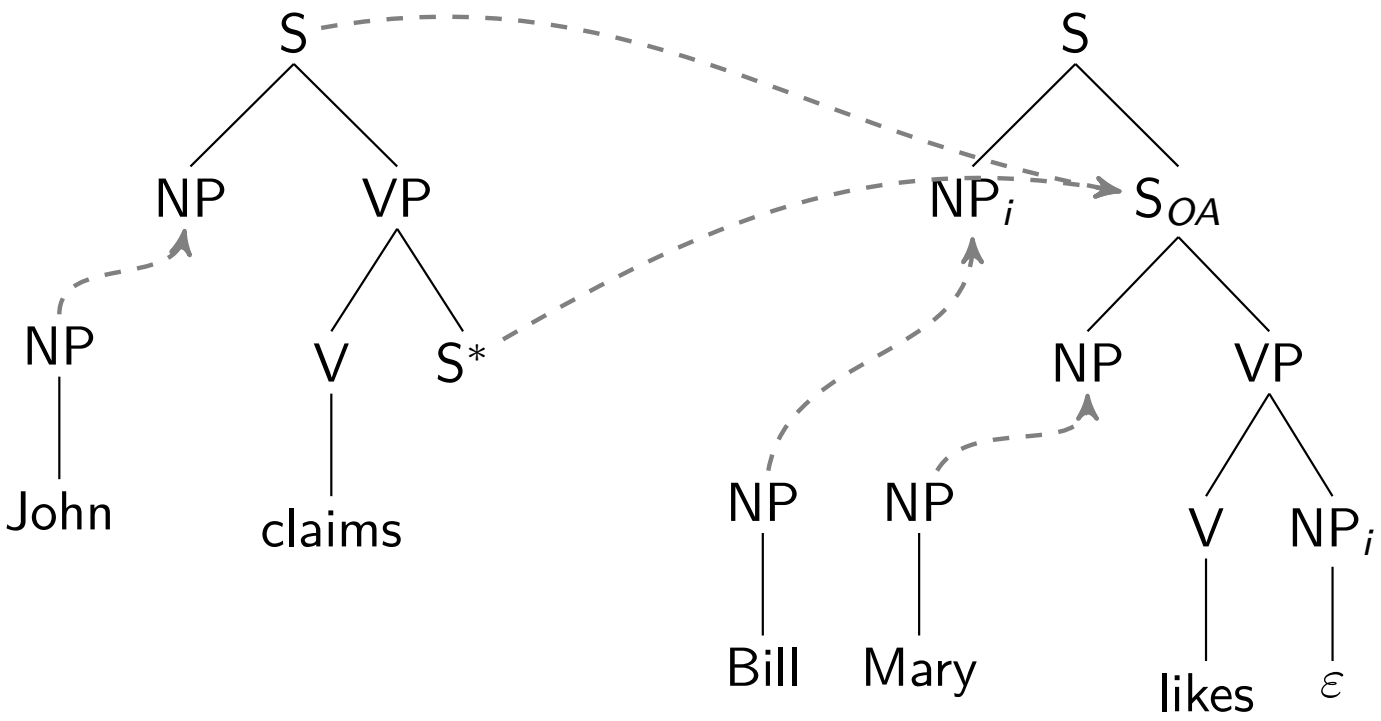


TAG and natural languages (4)

Example of a long-distance dependency:

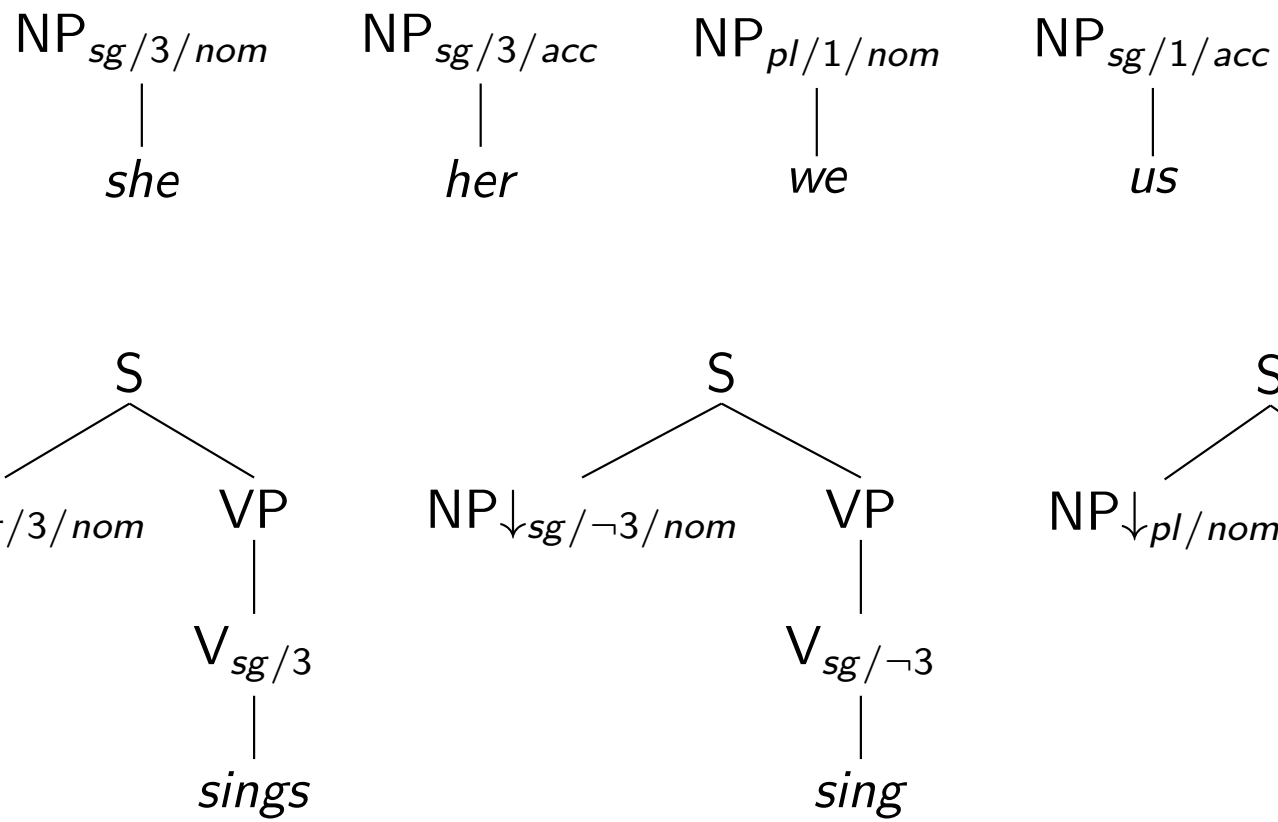
(5) a. Bill$_i$ John claims Mary likes t_i

b. Bill$_i$ John claims Susan thinks Mary likes t_i



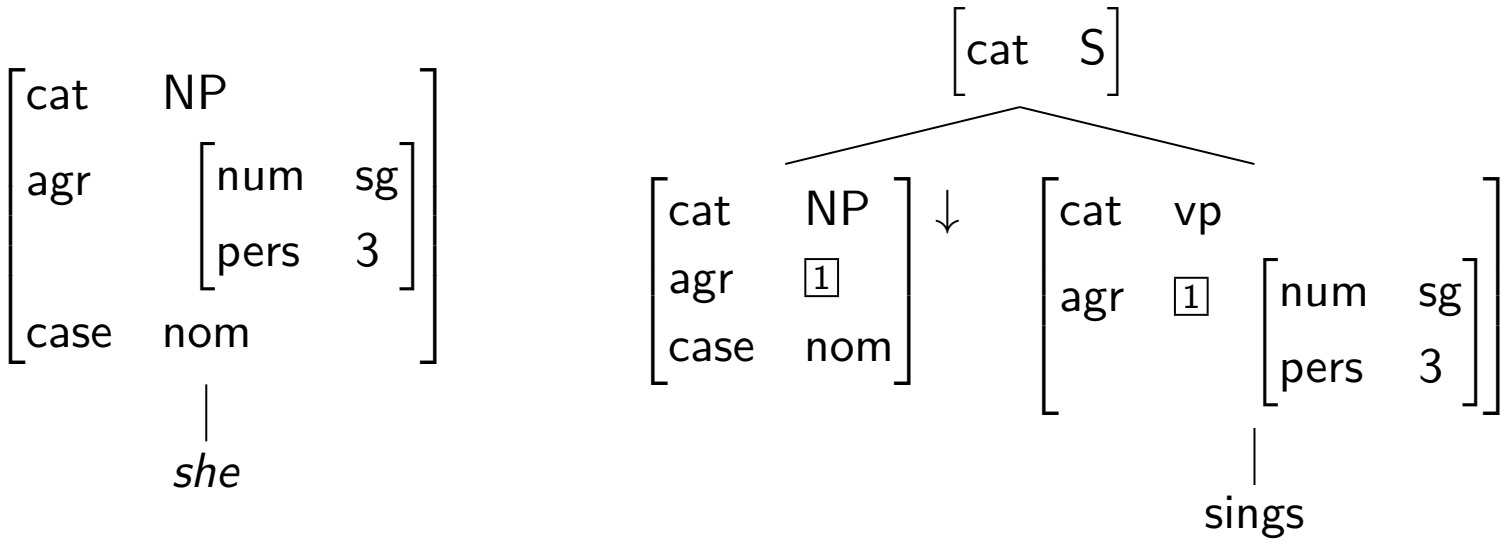
Feature Structure Based TAG (FTAG) (1)

- agreement without feature structures:
  - generalization of agreement is not captured
  - every morphological alternative gives rise to a new elementary tree

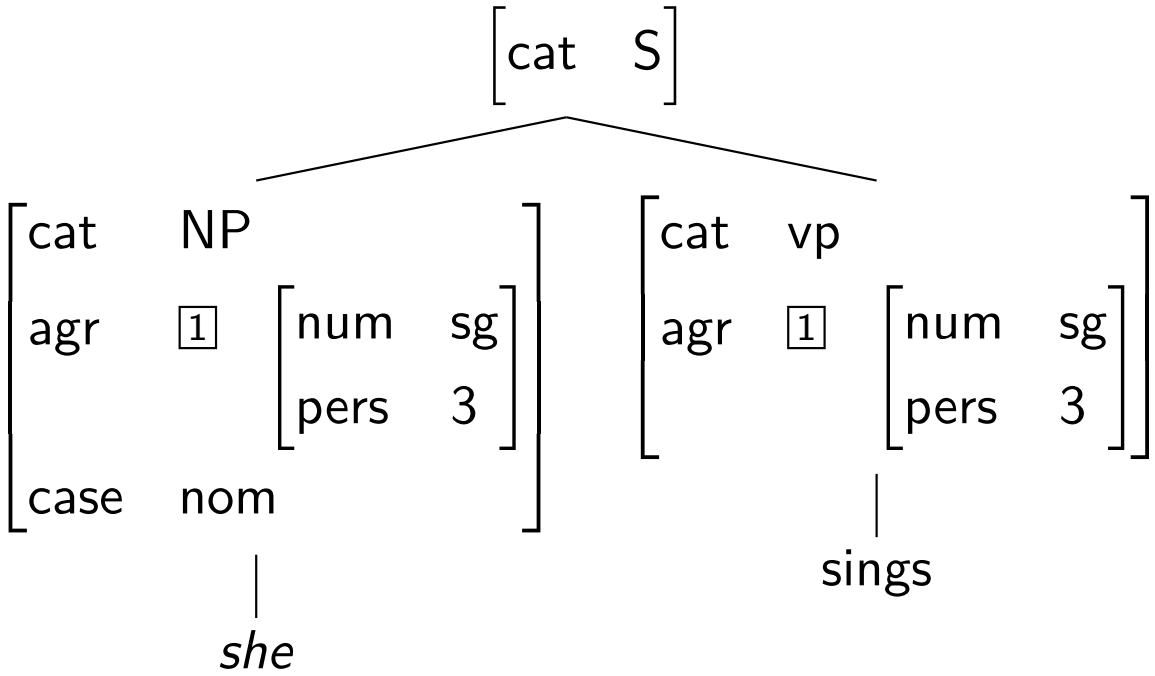


FTAG (2)

Solution: Feature structures as node labels



FTAG (3)


cat S
cat NP
agr 1
num sg
pers 3
case nom
she
cat vp
agr 1
num sg
pers 3
sings

FTAG (4)

- TAG + feature structures = Feature-structure based TAG (FTAG), [VSJ88]
- Idea: feature structures used as non-terminal nodes
- Advantage: smaller grammars that are easier to maintain
- Capture agreement
  - She sings. *She sing.
    → the finite verb and its subject agree in number and person
- Capture case marking
  - Pim likes her. *Pim likes she.
    → the object of a transitive verb must be in accusative case
- Capture selective/obligatory adjunction via specific top and bottom features (see next slides).

FTAG (5)

Each node has a top and a bottom feature structure. Nodes in the same elementary tree can share features (extended domain of locality).

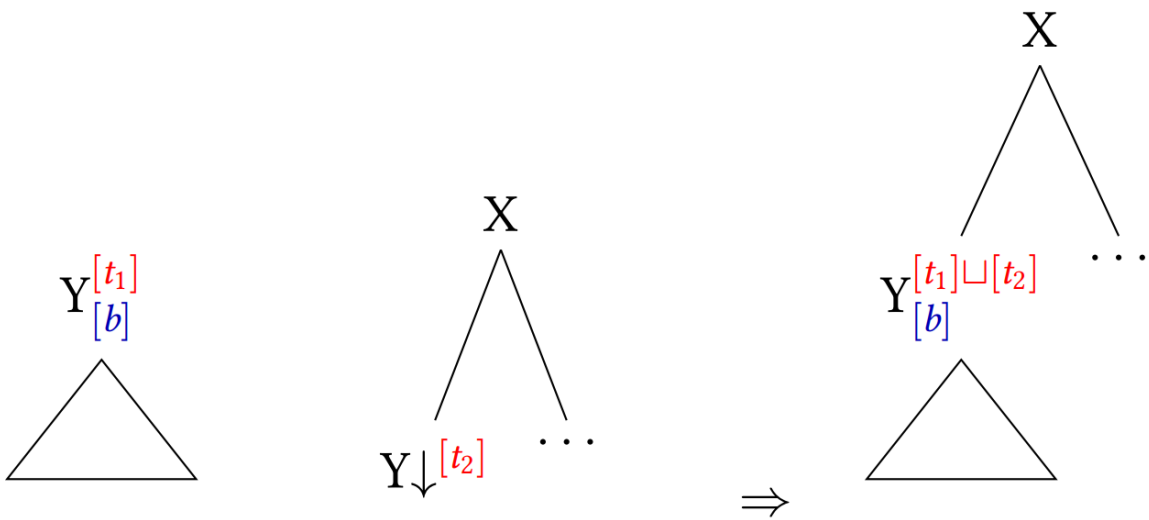
Intuition:

- The top feature structure tells us something about what the node presents within the surrounding structure, and
- the bottom feature structure tells us something about what the tree below the node represents.

In the final derived tree, both must be the same.

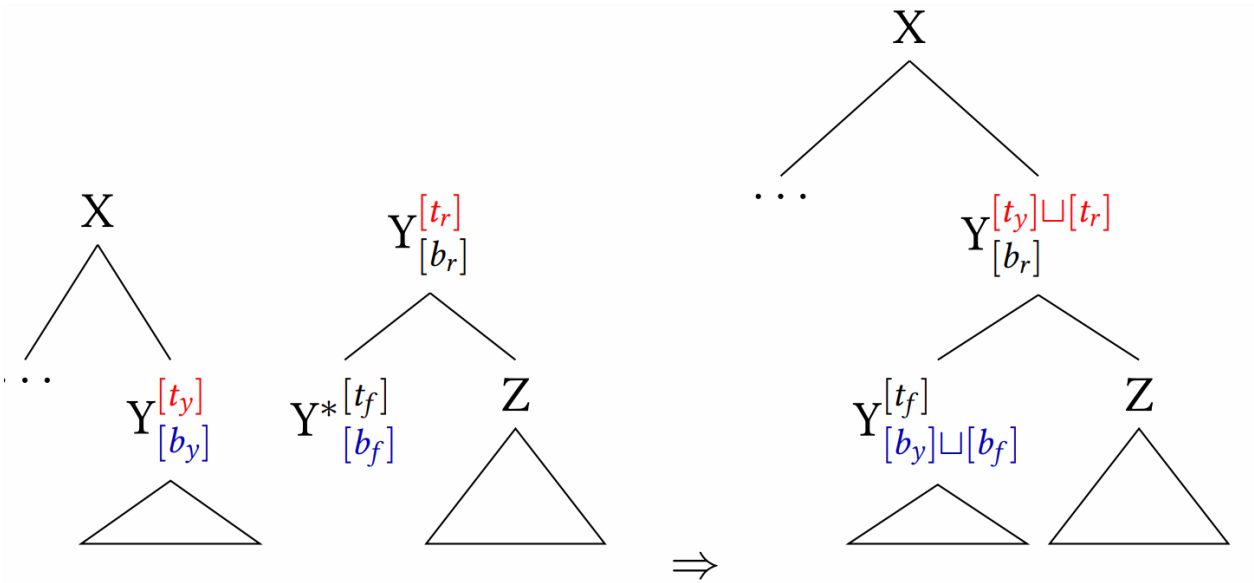
FTAG (6)

- The top features of the root of the tree to substitute unify with the top features of the substitution node



FTAG (7)

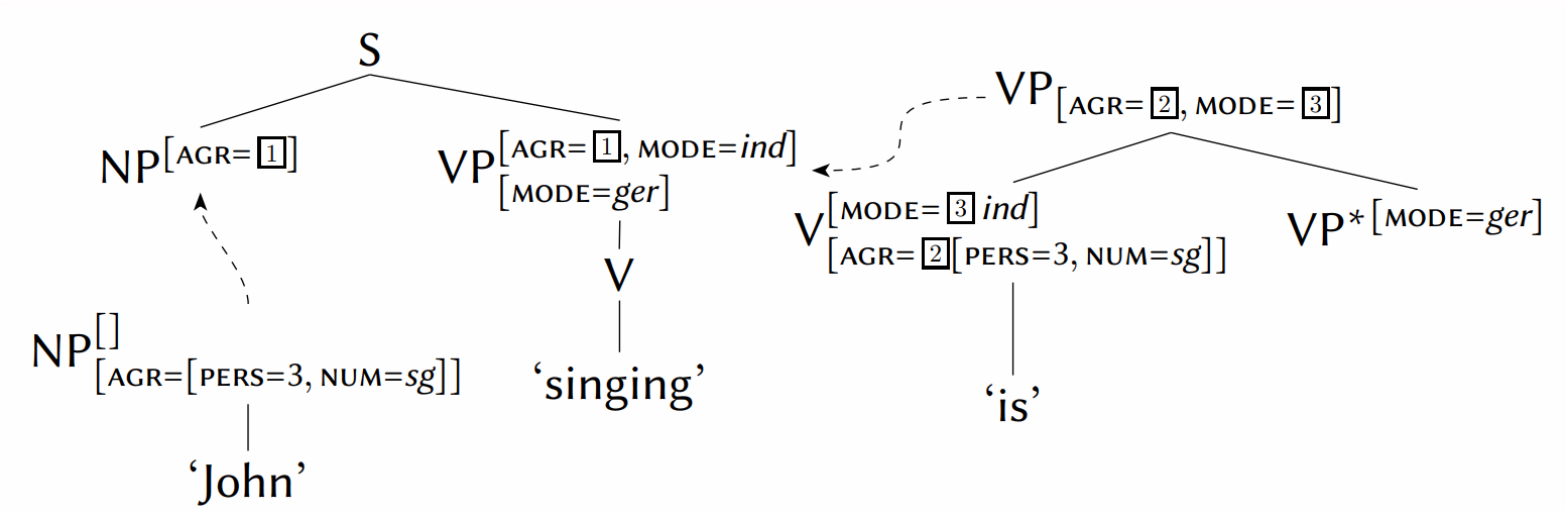
- The top features of the root of the auxiliary tree unify with the top features of the adjunction node, and the bottom features of the footnode of the auxiliary tree unify with the bottom features of the adjunction node.



FTAG (8)

Sample use of features for agreement and for obligatory adjunction:

(6) John is singing



References I

[Abe88] Anne Abeillé.
Parsing French with tree adjoining grammar: some linguistic accounts.
In *Proceedings of COLING*, pages 7–12, Budapest, 1988.

[Abe02] Anne Abeillé.
Une Grammaire Électronique du Français.
CNRS Editions, Paris, 2002.

[Fra02] Robert Frank.
Phrase Structure Composition and Syntactic Dependencies.
MIT Press, Cambridge, Mass, 2002.

[JLT75] Aravind K. Joshi, Leon S. Levy, and Masako Takahashi.
Tree Adjunct Grammars.
Journal of Computer and System Science, 10:136–163, 1975.

[JS97] Aravind K. Joshi and Yves Schabes.
Tree-Adjoining Grammars.
In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, pages 69–123. Springer, Berlin, 1997.

[Kal10] Laura Kallmeyer.
Parsing Beyond Context-Free Grammars.
Cognitive Technologies. Springer, Heidelberg, 2010.

[Kro87] Anthony S. Kroch.
Unbounded dependencies and subjacency in a Tree Adjoining Grammar.
In A. Manaster-Ramer, editor, *Mathematics of Language*, pages 143–172. John Benjamins, Amsterdam, 1987.

[VSJ88] K. Vijay-Shanker and Aravind K. Joshi.
Feature structures based tree adjoining grammar.
In *Proceedings of COLING*, pages 714–719, Budapest, 1988.

References II

[XTA01] XTAG Research Group.
A Lexicalized Tree Adjoining Grammar for English.
Technical report, Institute for Research in Cognitive Science, Philadelphia, 2001.
Available from ftp://ftp.cis.upenn.edu/pub/xtag/release-2.24.2001/tech-report.pdf.