Scaling up with Lexical Resource Semantics

Gerald Penn & Frank Richter

University of Toronto and Universität Tübingen

Alumni Meeting
Freudenstadt
June 21st, 2007

LRS: Background

- Integrated description of syntactic structures and their logical forms
- Strictly model-theoretic view of grammars
- Semantic composition is achieved by accumulating constraints which restrict logical forms
- LRS is designed for the description of
  - lexical idiosyncrasies
  - collocations
  - LF constraints
  - concord phenomena

Ty2 Signature in HPSG

```
ty2
   me type:type
      variable index:integer
      constant index:integer
      application functor:me arg:me
      abstraction var:variable body:me
      equation arg1:me arg2:me
      negation arg:me
      logical-constant arg1:me arg2:me
         conjunction <*...*>
      gen-quant var:variable restr:me scope:me
         every
         three <*...*>
  type
     atomic-type
        entity
        truth
        w-index
     complex-type in:type out:type
  integer
     zero
     n-zero pre:integer
```

Ty2 Principles in HPSG

(1)a. The Natural Numbers Principle:

$$integer \rightarrow \exists x \; {}^{x}[zero]$$

b. The Complex Term Principles:

$$application \rightarrow \begin{bmatrix} \text{TYPE} & \boxed{2} \\ \text{FUNCTOR TYPE} & \begin{bmatrix} \text{IN} & \boxed{1} \\ \text{OUT} & \boxed{2} \end{bmatrix} \\ \text{ARG TYPE} & \boxed{1} \end{bmatrix}$$

$$abstraction \rightarrow \begin{bmatrix} \text{TYPE} & \begin{bmatrix} \text{IN} & \boxed{1} \\ \text{OUT} & \boxed{2} \end{bmatrix} \\ \text{VAR TYPE} & \boxed{1} \\ \text{BODY TYPE} & \boxed{2} \end{bmatrix}$$

$$equation \rightarrow \begin{bmatrix} \text{TYPE} & truth \\ \text{ARG1 TYPE} & \boxed{1} \\ \text{ARG2 TYPE} & \boxed{1} \end{bmatrix}$$

$$negation \rightarrow \begin{bmatrix} \text{TYPE} & truth \\ \text{ARG TYPE} & truth \end{bmatrix}$$

$$logical\text{-}constant \rightarrow \begin{bmatrix} \text{TYPE} & truth \\ \text{ARG1 TYPE} & truth \\ \text{ARG2 TYPE} & truth \end{bmatrix}$$

$$gen\text{-}quant \rightarrow \begin{bmatrix} \text{TYPE} & truth \\ \text{RESTR TYPE} & truth \\ \text{SCOPE TYPE} & truth \end{bmatrix}$$

c. The Ty2 Non-Cyclicity Principle:

$$ty2 \rightarrow \forall_{\boxed{1}} \left(\begin{array}{l} \left(\bigvee \left\{ [\alpha \ \boxed{1}] \,\middle|\, \alpha \in \mathcal{A}_{Ty2} \right\} \right) \rightarrow \\ \neg \, \texttt{ty2-component}(:, \boxed{1}) \end{array} \right)$$

d. The Ty2 Finiteness Principle:

$$ty2 \rightarrow \exists_{\boxed{1}} \forall_{\boxed{2}} \left(\begin{array}{l} \texttt{ty2-component}(\boxed{2}, :) \ \rightarrow \\ \texttt{member}(\boxed{2}, \boxed{1}[chain]) \end{array} \right)$$

e. The Ty2 Identity Principle:

$$ty2 \rightarrow \forall_{\boxed{1}} \forall_{\boxed{2}} (\texttt{copy}(\boxed{1}, \boxed{2}) \ \rightarrow \ \boxed{1} = \boxed{2})$$

f. The Ty2-component Principle:

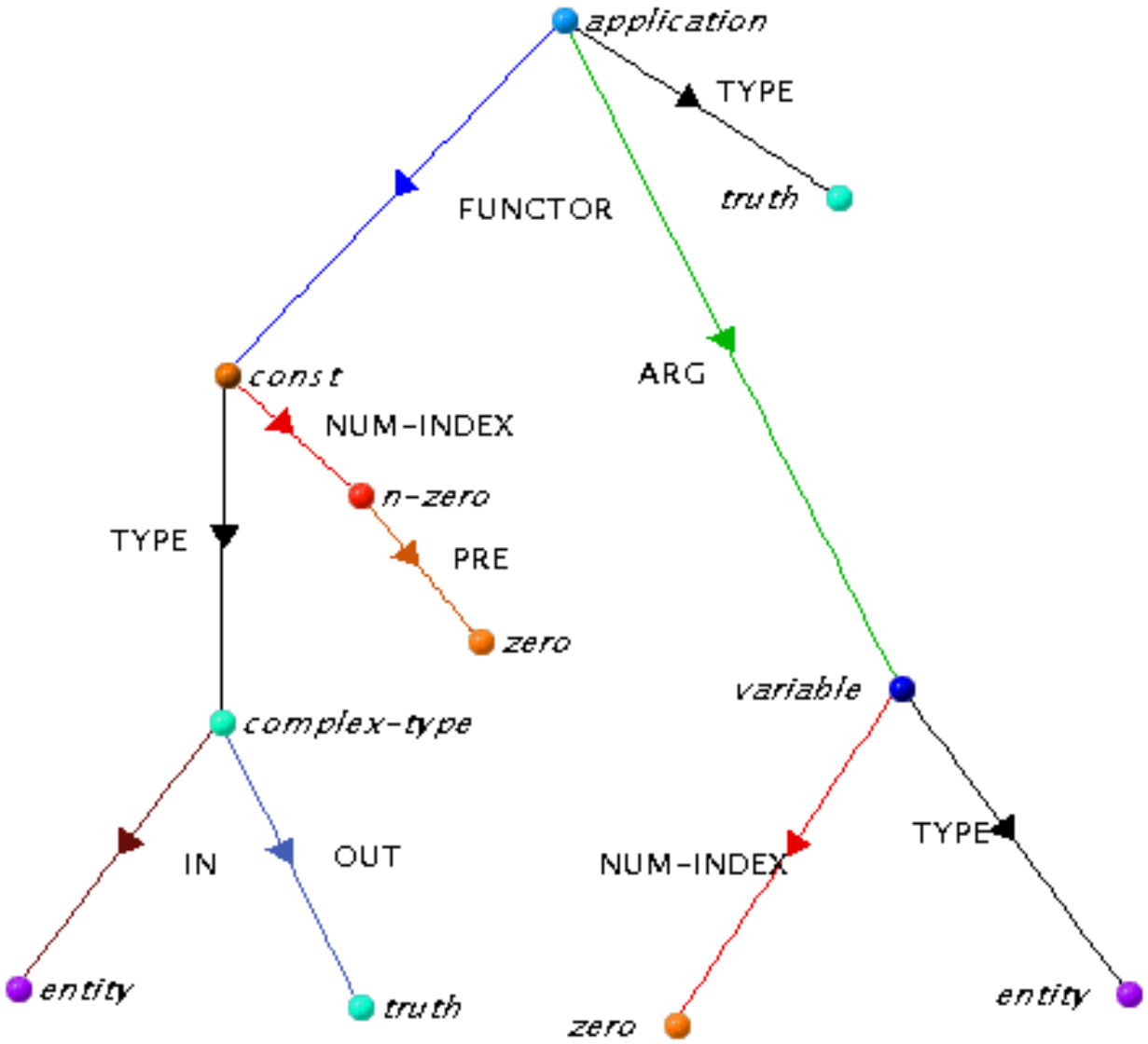
$$\forall x \forall y \left(\texttt{ty2-component}(x,y) \leftrightarrow \left(\begin{array}{l} x = y \vee \\ \bigvee \left\{ \exists_{\boxed{1}} \left(\begin{array}{l} {}^{y}[\alpha\ \boxed{1}] \wedge \\ \texttt{ty2-component}(x, \boxed{1}) \end{array} \right) \middle| \alpha \in \mathcal{A}_{Ty2} \right\} \end{array} \right) \right)$$

g. The Copy Principle:

$$\forall x \forall y \left(\texttt{copy}(x,y) \leftrightarrow \left(\begin{array}{l} \bigvee \left\{ {}^{x}[\sigma] \wedge {}^{y}[\sigma] \middle| \sigma \in \mathcal{S}_{Ty2} \right\} \wedge \\ \bigwedge \left\{ \forall_{\boxed{1}} \left(\begin{array}{l} {}^{x}[\alpha\ \boxed{1}] \rightarrow \\ \exists_{\boxed{2}} \left(\begin{array}{l} {}^{y}[\alpha\ \boxed{2}] \wedge \\ \texttt{copy}(\boxed{1}, \boxed{2}) \end{array} \right) \end{array} \right) \middle| \alpha \in \mathcal{A}_{Ty2} \right\} \end{array} \right) \right)$$

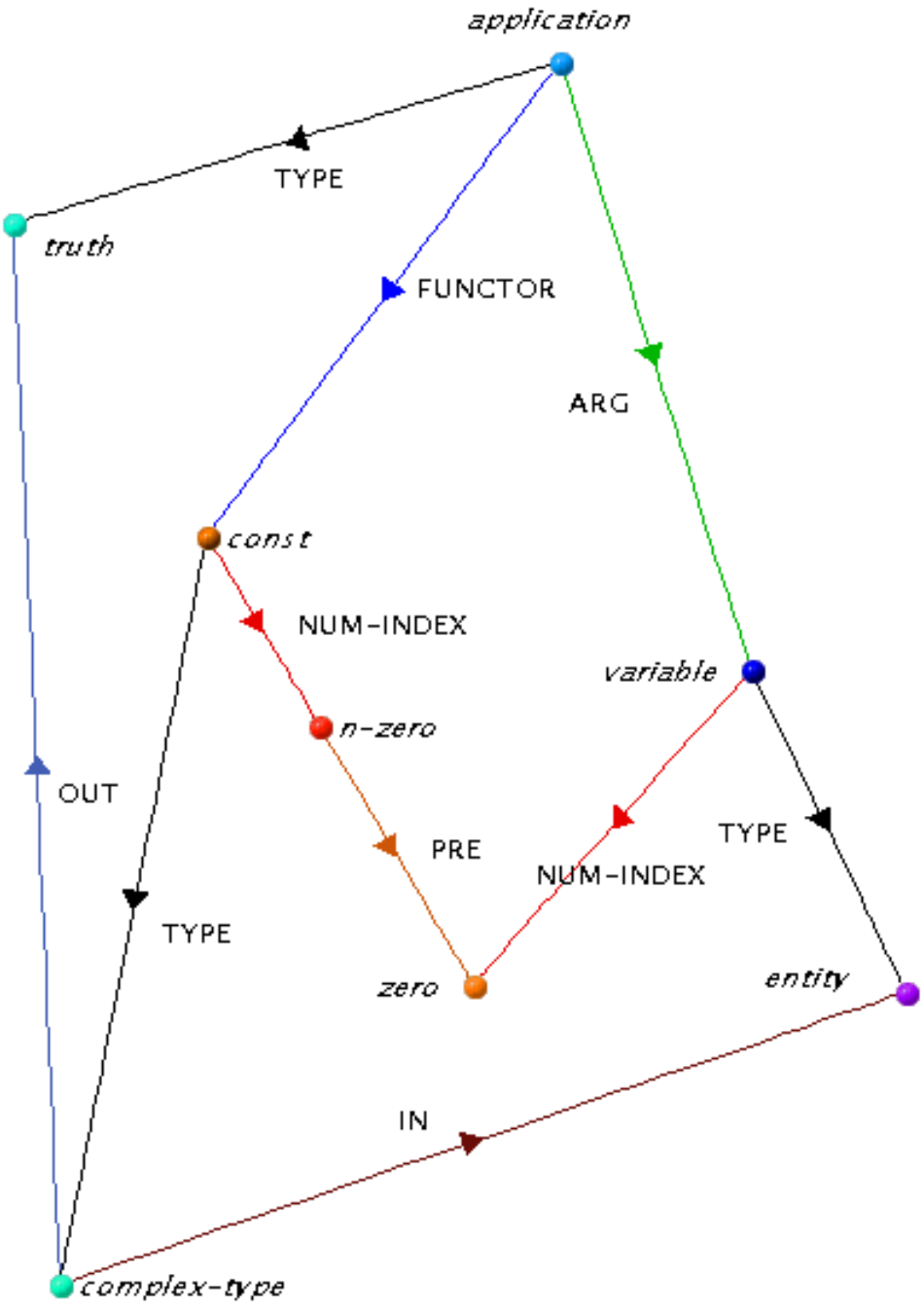
Ty2 Expressions in TFL (1)

A tree-shaped representation of $c_{et,1}(v_{e,0})$:



Ty2 Expressions in TFL (2)

The final representation of $c_{et,1}(v_{e,0})$:



Semantics of Signs in LRS

Different linguistic sources of restrictions on terms:

$$\begin{bmatrix} sign & & \\ \text{SYNSEM} & \begin{bmatrix} synsem & \\ \text{LOC CONTENT} & \begin{bmatrix} \text{VAR} & me \\ \text{MAIN} & me \end{bmatrix} \end{bmatrix} \\ \text{LF} & \begin{bmatrix} \text{EXCONT} & me \\ \text{INCONT} & me \\ \text{PARTS} & list(me) \end{bmatrix} \end{bmatrix}$$

- Combinatorial Semantics
  - External Content:
    Meaning contribution of the maximal projection of the semantic head to an utterance
  - Internal Content:
    Scopally lowest contribution of the semantic head of a sign
  - Parts (contribution constraints):
    Record of term contributions

- Local Semantics
  - Referential variable of a sign
  - Main content of a (lexical) sign

Why Constraint-based Composition

(2) *kein-* $\Rightarrow \lambda P_{s((se)t)} \lambda Q_{s((se)t)} . \neg \exists x_{se} [P_@(x) \wedge Q_@(x)]$

(3) Hans muss keine Krawatte tragen.

a. 'What Hans must do is not wear a tie.'
$\mathsf{must}'_@(\mathsf{hans}_e, \lambda @. \neg \exists x_{se} [\mathsf{tie}'_@(x) \wedge \mathsf{wear}'_@(\mathsf{hans}_e, x)])$

b. 'There is no tie such that Hans must wear that tie.'
$\neg \exists x_{se} [\mathsf{tie}'_@(x) \wedge \mathsf{must}'_@(\mathsf{hans}_e, \lambda @. \mathsf{wear}'_@(\mathsf{hans}, x))]$

c. 'It is not the case that Hans must wear a tie.'
$\neg \mathsf{must}'_@(\mathsf{hans}_e, \lambda @. \exists x_{se} [\mathsf{tie}'_@(x) \wedge \mathsf{wear}'_@(\mathsf{hans}_e, x)])$

(4)a. Chris sucht keine Wohnung.

b. *de re*:
$\neg \exists x_{se} [\mathsf{apartment}'_@(x) \wedge \mathsf{seek}'_@(\mathsf{chris}_e, \lambda @ \lambda P. P_@(x))]$
(there is no apartment x such that Chris seeks x)

c. *de dicto*:
$\neg [\mathsf{seek}'_@(\mathsf{chris}_e, \lambda @ \lambda P_{s((se)t)} . \exists x_{se} [\mathsf{apartment}'_@(x) \wedge P_@(x)])]$
(it is not the case that Chris seeks an apartment)

Why Constraint-based Composition

(5) Janek *(nie) pomaga nikomu.
Janek NM helped nobody
'Janek didn't help anybody.'
$ 'Janek didn't help nobody.'

(6) Nikt *(nie) pomaga nikomu.
Nobody NM helped nobody
'Nobody helped anybody.'

Three Lexical Entries

(7)a. John:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{john} \rangle \\ \textsc{ss loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \textit{noun} \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{index var} & \boxed{1}\ \mathsf{john}' \\ \textsc{main} & \boxed{1}\ \mathsf{john}' \end{bmatrix} \end{bmatrix} \\ \textsc{lf} & \begin{bmatrix} \textsc{excont} & \textit{me} \\ \textsc{incont} & \boxed{1}\ \mathsf{john}' \\ \textsc{parts} & \langle \boxed{1}\ \mathsf{john}' \rangle \end{bmatrix} \end{bmatrix}$$

b. laughs:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{laughs} \rangle \\ \textsc{ss loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \textit{verb} \\ \textsc{subcat} & \langle \mathrm{NP}_{\boxed{1}} \rangle \end{bmatrix} \\ \textsc{cont main} & \boxed{2a}\ \mathsf{laugh}' \end{bmatrix} \\ \textsc{lf} & \begin{bmatrix} \textsc{excont} & \textit{me} \\ \textsc{incont} & \boxed{2}\ \mathsf{laugh}'(\boxed{1}) \\ \textsc{parts} & \langle \boxed{2}\ \mathsf{laugh}'(\boxed{1}),\ \boxed{2a}\ \mathsf{laugh}' \rangle \end{bmatrix} \end{bmatrix}$$

c. always:

$$\left[\begin{array}{ll}
\textsc{phon} & \langle \textit{always} \rangle \\
\textsc{ss loc} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{head} & \left[\begin{array}{l} \textit{adv} \\ \textsc{mod}\ \ \mathrm{V}\left[\textsc{loc cont main}\ \boxed{2a}\ \right] \end{array}\right] \\
\textsc{subcat} & \textit{elist}
\end{array}\right] \\
\textsc{cont main} & \boxed{5a}\ \mathsf{always}'
\end{array}\right] \\
\textsc{lf} & \left[\begin{array}{ll}
\textsc{excont} & \textit{me} \\
\textsc{incont} & \boxed{5}\ \mathsf{always}'(\boxed{3}) \\
\textsc{parts} & \left\langle \boxed{5}\ \mathsf{always}'(\boxed{3}),\ \boxed{5a}\ \mathsf{always}' \right\rangle
\end{array}\right]
\end{array}\right]$$

$\&\ \boxed{2a} \lhd \boxed{3}$

Basic Principles

(8)a. The INCONT PRINCIPLE:

In each *lrs*, the INCONT value is an element of the PARTS list and a component of the EXCONT value.

$$lrs \rightarrow \left(\begin{bmatrix} \text{EXCONT} & \boxed{1} \\ \text{INCONT} & \boxed{2} \\ \text{PARTS} & \boxed{3} \end{bmatrix} \wedge \ \texttt{member}(\boxed{2}, \boxed{3}) \wedge \ \boxed{2} \lhd \boxed{1} \right)$$

b. The EXCONT PRINCIPLE:

Clause (a):
In every phrase, the EXCONT value of the non-head daughter is an element of the non-head daughter's PARTS list.

$$phrase \rightarrow \left(\left[\text{NH-DTR LF} \begin{bmatrix} \text{EXCONT} & \boxed{1} \\ \text{PARTS} & \boxed{2} \end{bmatrix} \right] \wedge \ \texttt{member}(\boxed{1}, \boxed{2}) \right)$$

Clause (b):
In every utterance, every subexpression of the EXCONT value of the utterance is an element of its PARTS list, and every element of the utterance's PARTS list is a subexpression of the EXCONT value.

$$u\text{-}sign \rightarrow$$

$$\forall\boxed{1}\ \forall\boxed{2}\ \forall\boxed{3}\ \forall\boxed{4}\left(\begin{array}{l}\left(\left[\text{LF}\left[\begin{array}{ll}\text{EXCONT} & \boxed{1}\\ \text{PARTS} & \boxed{2}\end{array}\right]\right]\ \wedge\ \boxed{3}\lhd\boxed{1}\ \wedge\ \texttt{member}(\boxed{4},\boxed{2})\right)\rightarrow\\ (\texttt{member}(\boxed{3},\boxed{2})\ \wedge\ \boxed{4}\lhd\boxed{1})\end{array}\right)$$

c. LRS Projection Principle:

In each *phrase*,
1. the EXCONT values of the head and the mother are identical,

$$phrase \rightarrow \left[\begin{array}{ll}\text{LF EXCONT} & \boxed{1}\\ \text{H-DTR LF EXCONT} & \boxed{1}\end{array}\right]$$

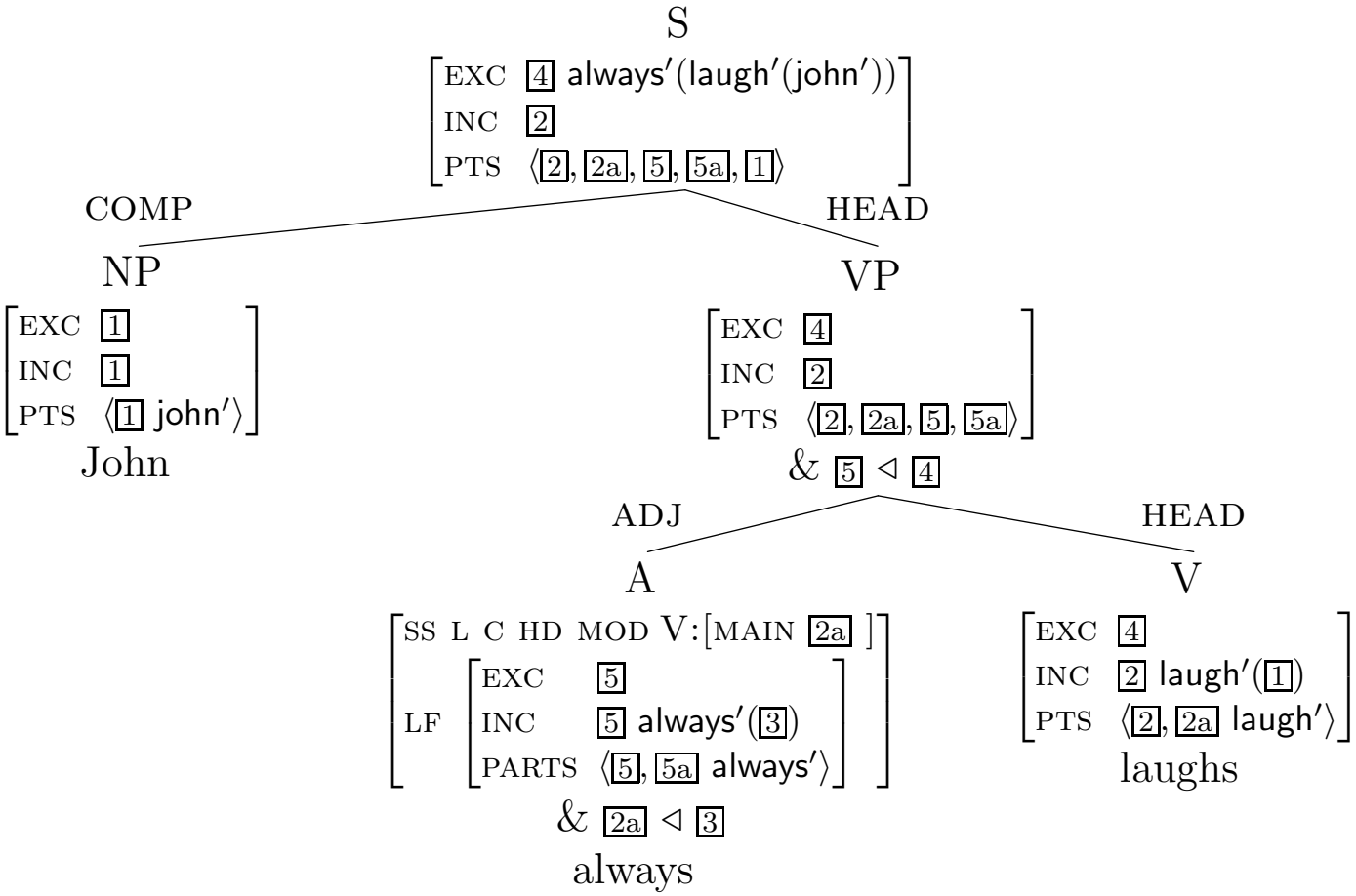
2. the INCONT values of the head and the mother are identical,

$$phrase \rightarrow \left[\begin{array}{ll}\text{LF INCONT} & \boxed{1}\\ \text{H-DTR LF INCONT} & \boxed{1}\end{array}\right]$$

3. the PARTS value contains all and only the elements of the PARTS values of the daughters.

$$phrase \rightarrow \left(\left[\begin{array}{ll}\text{LF PARTS} & \boxed{1}\\ \text{H-DTR LF PARTS} & \boxed{2}\\ \text{NH-DTR LF PARTS} & \boxed{3}\end{array}\right]\ \wedge\ \texttt{append}(\boxed{2},\boxed{3},\boxed{1})\right)$$

Analysis of a Sentence



The empirical domain

- Quantifier scope ambiguities

  (9)a. Every student reads a book.
  b. Three girls are likely to come.

- Concord phenomena (negative, interrogative, temporal)

  (10)a. Personne n'a rien vu.
  Nobody saw anything.

  b. Nikt nie pomaga nikomu.
  Nobody helps anybody.

  c. Wer hat gestern wen getroffen?
  Who met whom yesterday?

  d. Hy wou die boek gelees het.
  He wanted to read the book.

- LF discontinuities (split readings)

  (11) Hans braucht keine Krawatte zu tragen.
  It is not necessary that Hans wears a tie.

- Reconstruction

  (12) Ein Kennzeichen muss jedes Auto in Deutschland haben.
  Every car in Germany must have a license plate.

- Local and nonlocal semantics

  (13)a. Kim pflückt eine Blume/zwei Blumen/die meisten Blumen.
  Kim is picking a flower/two flowers/most flowers.
  b. # Kim pflückt ein Buch/zwei Bücher/die meisten Bücher.
  # Kim is picking a book/two books/most books.
  c. [Das Institut]$_i$ steht in der Wilhelmstraße. # Es$_i$ trifft sich jede Woche einmal zum Mittagessen.
  [The department]$_i$ is in Wilhelmstraße. # It$_i$ meets once every week for lunch.