Using Features and Types

Weiwei Sun

Institute of Computer Science and Technology
Peking University

August 26, 2019
Outline

Typed feature structure

Phrase-structure rules with features

Representing CCG using TFS
Motivation

Weakness of CFG

CFG treats each grammatical category symbol as atomic without internal structure.

⇒ Two categories are either identical or different.
⇒ There is no mechanism for saying that two categories are alike in some ways, but different in others.

Cross-cutting grammatical properties

<table>
<thead>
<tr>
<th></th>
<th>3rd singular subject</th>
<th>plural subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>direct object NP</td>
<td>denies</td>
<td>deny</td>
</tr>
<tr>
<td>No direct object NP</td>
<td>disappears</td>
<td>disappear</td>
</tr>
</tbody>
</table>
Using features

Observation

Words and phrases in natural languages typically behave alike in certain respects, but not others.

Key idea: Using features

- The elements associated to linguistic expressions, such as words, can be broken down.
- Complex categories can be decomposed to features that are the atomic units.
- Linguistic feature: a property-like element that indicates the grammatical behavior of syntactic constituents.

- The VP has the feature value past tense.
- The verb is a past tense verb.
Linguistic features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Example</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>I go, you go, he goes</td>
<td>1st, 2nd, 3rd</td>
</tr>
<tr>
<td>number</td>
<td>he dances, they dance</td>
<td>singular, plural</td>
</tr>
<tr>
<td>case</td>
<td>he brings Bob, Bob brings him</td>
<td>nominative accusative</td>
</tr>
<tr>
<td>tense</td>
<td>go, went, gone</td>
<td>past, present, future</td>
</tr>
<tr>
<td>modality</td>
<td>may, can,</td>
<td>conditional, subjunctive</td>
</tr>
</tbody>
</table>

A nice summary of linguistic features

http://www.grammaticalfeatures.net
Feature structure

Description

Use a feature structure to specify grammatical information.

- A feature structure is a specification of a set of features, each of which is paired with a particular value.
- A feature structure can be represented by an AVM.

\[
\begin{bmatrix}
\text{FEATURE}_1 & \text{VALUE}_1 \\
\text{FEATURE}_2 & \text{VALUE}_2 \\
\vdots \\
\text{FEATURE}_n & \text{VALUE}_n \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{FORM} & \text{dog} \\
\text{NUMBER} & \text{singular} \\
\text{ANIMACY} & \text{animate} \\
\end{bmatrix}
\]
More on feature values

**Atomic value**
An unstructured value, one with only one part

\[
\begin{bmatrix}
\text{TENSE} & \text{past} \\
\text{PERSON} & 2
\end{bmatrix}
\]

**Complex value**
A structured value, itself a feature structure

\[
\begin{bmatrix}
\text{TENSE} & \text{past} \\
\text{AGREEMENT} & \begin{bmatrix}
\text{PERSON} & 2 \\
\text{NUMBER} & \text{singular}
\end{bmatrix}
\end{bmatrix}
\]
# Outside the linguistic world

<table>
<thead>
<tr>
<th>中文名称</th>
<th>英文名称</th>
<th>结果</th>
<th>单位</th>
<th>参考值</th>
</tr>
</thead>
<tbody>
<tr>
<td>1   水尿干化学检查</td>
<td>干化学</td>
<td>:</td>
<td></td>
<td>1.015-1.025</td>
</tr>
<tr>
<td>2   浓度度</td>
<td>TURB</td>
<td>清亮</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3   颜色</td>
<td>COLOUR</td>
<td>浅黄色</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4   尿糖</td>
<td>G LU(U)</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>5   尿酮体</td>
<td>KET</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>6   尿潜血</td>
<td>BLD</td>
<td>2+</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>7   尿蛋白</td>
<td>PRO</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>8   亚硝酸盐</td>
<td>N IT</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>9   尿胆红素</td>
<td>BIL</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>10  尿胆原</td>
<td>U BG</td>
<td>±</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11  尿比重</td>
<td>SG</td>
<td>1.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12  尿pH</td>
<td>pH</td>
<td>6</td>
<td></td>
<td>5.5-6.5</td>
</tr>
<tr>
<td>13  尿白细胞</td>
<td>LEU</td>
<td>-</td>
<td></td>
<td>阴性</td>
</tr>
<tr>
<td>14  尿流式沉渣检查</td>
<td>尿流式</td>
<td>:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15  尿红细胞数</td>
<td>U-RBC</td>
<td>10</td>
<td>个/ul</td>
<td>0-25</td>
</tr>
<tr>
<td>16  尿白细胞数</td>
<td>U-W BC</td>
<td>6</td>
<td>个/ul</td>
<td>0-30</td>
</tr>
<tr>
<td>17  尿上皮细胞数</td>
<td>U-EC</td>
<td>7</td>
<td>个/ul</td>
<td>0-15</td>
</tr>
<tr>
<td>18  尿管型数</td>
<td>U-CAST</td>
<td>0.4</td>
<td>个/ul</td>
<td>0-2.5</td>
</tr>
<tr>
<td>19  尿细菌数</td>
<td>U-BACT</td>
<td>105</td>
<td>个/ul</td>
<td>≤6000</td>
</tr>
</tbody>
</table>
Wanna a family

The image shows a phylogenetic tree illustrating the evolutionary relationships among different species of primates. The tree is color-coded and includes various primates and their estimated time intervals for divergent evolution. The species and their respective taxonomic classifications are indicated along the branches of the tree.
Typed feature structure

- Entities belonging to a particular type have their own special properties.

⇒ Each type of entity has its own constellation of features
  - Some features are declared appropriate for entities of the indicated type
  - Other features are sanctioned by one of the supertypes

- Type has subtype and supertype ⇒ Hierarchical organization

Example

```
feature structure

expression

word phrase noun verb det prep adj conj

pos
```
## Outside the linguistic world

<table>
<thead>
<tr>
<th>TYPE</th>
<th>FEATURES/VALUES</th>
<th>IMMEDIATE ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>[NAME string, TEL number]</td>
<td>entity</td>
</tr>
<tr>
<td>individual</td>
<td>[BIRTHDAY date]</td>
<td>entity</td>
</tr>
<tr>
<td>organization</td>
<td>[FOUNDERS list(individual)]</td>
<td>entity</td>
</tr>
<tr>
<td>university</td>
<td>[PRESIDENT individual]</td>
<td>organization</td>
</tr>
<tr>
<td>department</td>
<td>[CHAIR individual]</td>
<td>organization</td>
</tr>
</tbody>
</table>

- [NAME Weiwei Sun, TEL 18****5]
- [NAME ICST.PKU, TEL 010-82529922]
## Outside the linguistic world

<table>
<thead>
<tr>
<th>TYPE</th>
<th>FEATURES/VALUES</th>
<th>IMMEDIATE ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>[NAME string]</td>
<td>entity</td>
</tr>
<tr>
<td></td>
<td>[TEL number]</td>
<td></td>
</tr>
<tr>
<td>individual</td>
<td>[BIRTHDAY date]</td>
<td>entity</td>
</tr>
<tr>
<td>organization</td>
<td>[FOUNDERS list(individual)]</td>
<td>entity</td>
</tr>
<tr>
<td>university</td>
<td>[PRESIDENT individual]</td>
<td>organization</td>
</tr>
<tr>
<td>department</td>
<td>[CHAIR individual]</td>
<td>organization</td>
</tr>
</tbody>
</table>

- **Weiwei Sun**
  - NAME: Weiwei Sun
  - TEL: 18****5

- **ICST.PKU**
  - NAME: ICST.PKU
  - TEL: 010-82529922
### Outside the linguistic world

<table>
<thead>
<tr>
<th>TYPE</th>
<th>FEATURES/VALUES</th>
<th>IMMEDIATE ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>[NAME string, TEL number]</td>
<td>entity</td>
</tr>
<tr>
<td>individual</td>
<td>[BIRTHDAY date]</td>
<td>entity</td>
</tr>
<tr>
<td>organization</td>
<td>[FOUNDERS list(individual)]</td>
<td>entity</td>
</tr>
<tr>
<td>university</td>
<td>[PRESIDENT individual]</td>
<td>organization</td>
</tr>
<tr>
<td>department</td>
<td>[CHAIR individual]</td>
<td>organization</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{individual} & : [\text{Name} Weiwei Sun, \text{Tel} 18****5] \\
\text{department} & : [\text{Name} ICST.PKU, \text{Tel} 010-82529922]
\end{align*}
\]
Outside the linguistic world

<table>
<thead>
<tr>
<th>TYPE</th>
<th>FEATURES/VALUES</th>
<th>IMMEDIATE ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>[NAME string, TEL number]</td>
<td>entity</td>
</tr>
<tr>
<td>individual</td>
<td>[BIRTHDAY date]</td>
<td>entity</td>
</tr>
<tr>
<td>organization</td>
<td>[FOUNDERS list(individual)]</td>
<td>entity</td>
</tr>
<tr>
<td>university</td>
<td>[PRESIDENT individual]</td>
<td>organization</td>
</tr>
<tr>
<td>department</td>
<td>[CHAIR individual]</td>
<td>organization</td>
</tr>
</tbody>
</table>

---

**Example:**

- **Individual:** Weiwei Sun
  - **NAME:** Weiwei Sun
  - **BIRTHDAY:** **-**-198*
  - **TEL:** 18****5

- **Department:** ICST.PKU
  - **NAME:** ICST.PKU
  - **FOUNDER:** Xuan Wang
  - **CHAIR:** Zongming Guo
  - **TEL:** 010-82529922
Typed feature structure

Linguistic features

Part-of-speech

feature structure

expression

[HEAD]

word phrase

pos

agr-pos

[AGR]

noun verb det

V: N: NP:

[head word verb] [head word noun] [phrase HEAD noun]
Linguistic features

Valence

Feature: VAL
Value of VAL: val-cat

Feature of val-cat: COMPS
Value of COMPS: itr, str, dtr

Feature of val-cat: SPR
Value of SPR: +/-

Abbreviations

IV:  

TV:  

DTV:  

word
HEAD
verb

word
HEAD
verb

word
HEAD
verb

VAL
[ val-cat
COMPS
itr ]

VAL
[ val-cat
COMPS
str ]

...
Linguistic features

(1)  
a. We created a monster.  
b. our creation of a monster
Mini type hierarchy

```
feature-structure
  expression
    word
    phrase
  val-cat
    [HEAD]
    [SPR, COMPS]
  pos
    agr-pos
      [AGR]
    verb
    det
    noun
    prep
    adj
    conj
```
Outline

Typed feature structure

Phrase-structure rules with features

Representing CCG using TFS
Reformulating the grammar rules (1)

\[ VP \rightarrow V \ NP^* \]
Reformulating the grammar rules (2)

- $S \rightarrow NP\ VP$
- $NP \rightarrow (D)\ NOM$

```
[phrase]
HEAD 1 verb
VAL
       COMPS itr
       SPR +

[phrase]
HEAD 1 noun
VAL
       COMPS itr
       SPR +

[phrase]
HEAD |
VAL
       COMPS itr
       SPR +

[word]
HEAD det
VAL
       COMPS itr
       SPR +

[phrase]
HEAD |
VAL
       COMPS itr
       SPR +
```
**Common and proper nouns**

\[
\langle \text{cat}, \left[ \begin{array}{l}
\text{word} \\
\text{HEAD} \\
\text{VAL}
\end{array} \right] \right]
\quad
\langle \text{David}, \left[ \begin{array}{l}
\text{word} \\
\text{HEAD} \\
\text{VAL}
\end{array} \right] \right]
\]
Phrase-structure rules with features

Representing CCG using TFS

Tree

[phrase
  HEAD verb
  VAL [COMPS itr
       SPR +]]

[phrase
  HEAD noun
  VAL [COMPS itr
       SPR +]]

[phrase
  HEAD verb
  VAL [COMPS itr
       SPR -]]

[phrase
  HEAD noun
  VAL [COMPS itr
       SPR +]]

Alex

[word
  HEAD noun
  VAL [COMPS itr
       SPR +]]

[word
  HEAD verb
  VAL [COMPS str
       SPR -]]

[phrase
  HEAD noun
  VAL [COMPS itr
       SPR -]]

[phrase
  HEAD noun
  VAL [COMPS itr
       SPR -]]

[word
  HEAD det
  VAL [COMPS itr
       SPR +]]

the

[word
  HEAD noun
  VAL [COMPS itr
       SPR -]]

allegation

Using Features and Types
Generalizing grammar rules

PP attachment

- $\text{VP} \rightarrow \text{VP} \ \text{PP}$
- $\text{NOM} \rightarrow \text{NOM} \ \text{PP}$

Combining them

$$
\begin{array}{c}
\text{phrase} \\
\text{HEAD} \ [1] \\
\text{VAL} \ [\text{COMPS} \ \text{itr}] \\
\end{array} \\
\rightarrow \\
\begin{array}{c}
\text{phrase} \\
\text{HEAD} \ [1] \\
\text{VAL} \ [\text{SPR} \ -] \\
\end{array} \\
\text{PP}
$$

Generalization

Only one $\textit{rule}$ is needed.
Agreement

Two features

Two features

Lexical entry & Grammar rule

David,

\[ \langle \text{David,} \rangle \]

\[
\begin{align*}
\text{HEAD} & : \text{noun} \\
\text{AGR} & : \text{agr-cat} \\
\text{PER} & : 3rd \\
\text{NUM} & : \text{sg} \\
\text{agr-cat} & : \text{PER} 3rd, \text{NUM sg} \\
\text{SPR} & : +
\end{align*}
\]

\[
\begin{align*}
\text{HEAD} & : \text{verb} \\
\text{COMPS} & : \text{itr} \\
\text{VAL} & : \text{spr} + \\
\text{phrase} & : \text{valuation}
\end{align*}
\]

\[
\begin{align*}
\text{HEAD} & : \text{AGR} 2 \\
\text{VAL} & : \text{spr} - \\
\text{phrase} & : \text{structure}
\end{align*}
\]
Head feature principle

Head Feature Principle (HFP)

In any headed phrase, the HEAD value of the mother and the HEAD value of the head daughter must be identical.

\[
\begin{align*}
\text{phrase} & \quad \text{VAL} \begin{bmatrix} \text{COMPS} & \text{itr} \end{bmatrix} \rightarrow \text{word} \quad \text{VAL} \begin{bmatrix} \text{COMPS} & \text{itr} \end{bmatrix} \\
\text{phrase} & \quad \text{VAL} \begin{bmatrix} \text{SPR} & - \end{bmatrix} \rightarrow \text{word} \quad \text{VAL} \begin{bmatrix} \text{SPR} & - \end{bmatrix} \\
\text{phrase} & \quad \text{VAL} \begin{bmatrix} \text{COMPS} & \text{itr} \end{bmatrix} \rightarrow \text{phrase} \quad \text{HEAD} \begin{bmatrix} \text{AGR} & 2 \end{bmatrix} \\
\text{phrase} & \quad \text{VAL} \begin{bmatrix} \text{SPR} & + \end{bmatrix} \rightarrow \text{phrase} \quad \text{VAL} \begin{bmatrix} \text{SPR} & - \end{bmatrix}
\end{align*}
\]
Outline

Typed feature structure

Phrase-structure rules with features

Representing CCG using TFS
Representing a cat

- Define a type `word-or-phrase` to represent a word or a constituent. It has two features, viz. `PHON` and `CAT`.
- A `cat` has two features, viz. `RESULT (RES)` and `ARGUMENT (ARG)`, and has two subtypes, viz. `forward("/")` and `backward("\")`.
- Basic categories, such as “np” are also subtypes of `cat`.

**Example:** `likes`

```
word-or-phrase

PHON <LIKES>

forward

CAT RES RES s

ARG np

ARG np
```
Representing a rule (1)

**Forward application**

\[
\begin{align*}
\text{word-or-phrase} & \quad \text{PHON} \quad \frac{1}{2} \\
\text{PHON} & \quad x \\
\text{CAT} & \quad y
\end{align*}
\]

\[
\rightarrow
\begin{align*}
\text{word-or-phrase} & \quad \text{PHON} \quad 1 \\
\text{forward} & \quad \text{PHON} \quad 2 \\
\text{CAT} & \quad \text{RES} \quad x \\
\text{ARG} & \quad \text{y}
\end{align*}
\]

**Forward composition**

\[
\begin{align*}
\text{word-or-phrase} & \quad \text{PHON} \quad \frac{1}{2} \\
\text{forward} & \quad \text{PHON} \quad 1 \\
\text{CAT} & \quad \text{RES} \quad x \\
\text{ARG} & \quad \text{y}
\end{align*}
\]

\[
\rightarrow
\begin{align*}
\text{word-or-phrase} & \quad \text{PHON} \quad 1 \\
\text{forward} & \quad \text{PHON} \quad 2 \\
\text{CAT} & \quad \text{RES} \quad x \\
\text{ARG} & \quad \text{z}
\end{align*}
\]

\[
\begin{align*}
\text{word-or-phrase} & \quad \text{PHON} \quad 2 \\
\text{ARG} & \quad \text{y}
\end{align*}
\]
Representing a rule (2)

Forward type raising

\[
\begin{align*}
\text{word-or-phrase} & \\
\text{PHON} & \begin{cases} 1 \end{cases} \\
\text{CAT} & \begin{cases}
\begin{align*}
\text{forward} & \\
\text{RES} & y \\
\text{ARG} & \begin{cases}
\text{backward} & \\
\text{RES} & y \\
\text{ARG} & x
\end{cases}
\end{align*}
\end{cases} \\
\rightarrow & \\
\text{PHON} & \begin{cases} 1 \end{cases} \\
\text{CAT} & x
\end{align*}
\]
Reading

- §3, *Syntactic Theory: A Formal Introduction*
- §2.3, *Aspects of the Theory of Syntax*