Harmonization of Dependency Trees

Daniel Zeman, Rudolf Rosa

May 12, 2022
Harmonization of Dependency Tree Style

• Universal Dependencies
  • Solved problem?
  • Almost. But not completely.

• Transformations sometimes needed for learnability experiments
• There are still non-UD treebanks that need to be converted
• SUD (Surface-syntactic Universal Dependencies)

• Transformations trivial?
  • Not necessarily...
Two known Russian historians
Danish Dependency Treebank
Na slavách odsoudil podczas czynów krytykował.
Adpositions

in diesem Jahr kommen in this year come

APPR PDAT NN VVINF
ADP DET NOUN VERB

MO OC

NK

NK

NK

det

obl

case

xcomp
Lahore in closed

Adpositions
Ельцин считал, что стране нужен экономист.

Yeltsin realized that by-country needed economist.
hasta que se fijen los comicios
Una sala ha dovuto essere sgomberata.
Verb Groups

Based-on can for-example be determined whether

Op_grond_van kan bijvoorbeeld worden vastgesteld of
apples, oranges and lemons
Äpfel, Orangen und Zitronen
äpplen, apelsiner och citroner
Coordination: Mel’čuk [ja]

リンゴ と オレンジ と レモン
ringo to orenji to remon

conj

punct

cc

conj

HD

MRK

HD

MRK

XXX

conj
Coordination: Mel’čuk [tr]
Coordination: Prague [cs]

```
jablka, pomaranče a citróny
```

Diagram:
- Coordination structure
- Coordination: Prague [cs]
- Words: jablka, pomaranče, citróny
- Conjunctions: a
- Punctuation: ,

Diagram elements:
- Coord
- AuxX
- XXX_M
- punct
- conj
- cc

Diagram interpretation:
- The diagram illustrates the coordination structure of the sentence in Czech.
- The words 'jablka', 'pomaranče', and 'citróny' are coordinated with the conjunction 'a'.
అపిలా, నారింజము మరియు నిమమ్పండు‌

Coordination: Prague [te]
 Coordination: Obligatorily Nested Prague [zh]

蘋果（píngguǒ）, 桔子（júzi）, 和（hé）, 檸檬（níngméng）
ябълки, портокали и лимони
jabălki, portokali i limoni
æbler, appelsiner og citroner
Coordination: Tesnière [hu]
Interaction of Structures in Prague

Trains to and from Prague or from Liberec.

Diagram:
- "vlaky" (trains) -> AuxP_M -> "do" (to) -> Coord_M -> "nebo" (or) -> Coord
- "a" (and) -> AuxP_M -> "od" (from) -> Coord_M -> "od" (from) -> Coord
- "z" (from) -> Coord_M -> "Liberec" (Liberec)
- "Prahy" (Prague) -> AuxP_M -> "Liberce" (Liberec)
- "case" -> "conj"
Interaction of Structures in Prague

trains to and from Prague or from Liberec
vlaky do a z Prahy nebo od Liberce

trains to and from Prague or from Liberec
Interaction of Structures in Prague

Vlaky do a z Prahy nebo od Liberce

Trains to and from Prague or from Liberec
Interactions of Structures in Prague

```
vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
```
Interaction of Structures in Prague

- vlaky (trains) to and from Prague or Liberec
- do a z nebo od Liberec
Interaction of Structures in Prague

Trains to and from Prague or from Liberec.
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
Round-Trip: Shared Dependents

- Czech students and teachers
- Coord
- Atr
- XXX_M
- amod
- conj
- cc

- Czech students and teachers
- Coord
- Atr
- XXX_M
- XXX_M
- XXX_M
Round-Trip: Nested Coordination

- jablka a hrušky nebo citróny
  - apples and pears or lemons

- jablka a hrušky nebo citróny
  - apples and pears or lemons

- jablka a hrušky nebo citróny
  - apples and pears or lemons
Round-Trip: Auxiliary Verbs

he could have seen her
Tree-Rewriting Toolkits

• Udapi
  • http://udapi.github.io/ (main page)
  • https://ufal.mff.cuni.cz/~zeman/vyuka/deptreebanks/NPFL075-working-with-UD.pdf (tutorial)

• Grew
  • https://grew.fr/
Phrase vs. Dependency Trees

[Diagram of a phrase tree and a dependency tree, showing the sentence: "Peter gave two pears to Paul." Each noun phrase (NP) and verb phrase (VP) is labeled with its corresponding parts of speech (POS) and dependencies.]
Phrase vs. Dependency Trees

- **Phrase trees**
  - Usually do not mark the *head*
  - May not mark the *function* of the constituent in the superordinate constituent
Phrase vs. Dependency Trees

• Phrase trees
  • Usually do not mark the head
  • May not mark the function of the constituent in the superordinate constituent

• Dependency trees
  • Do not show nonterminals (phrase types)
    • Nor any other phrase-level features
  • Do not show “how the sentence is generated” (order, recursion, proximity of constituents)
Example

S

NP  VP

N  V  NP

John  loves  N

Mary

S

VP

N  V

John  loves  N

Mary
Dick Darman, call your office.
he will meet Mary today and (he) (will meet) Clara tomorrow
he will meet Mary today and (he) (will) (meet) Clara tomorrow