Harmonization of Dependency Trees

Daniel Zeman, Rudolf Rosa

May 9, 2024
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...
Danish Dependency Treebank

Two known Russian historians
In an authoritarian regime
Na oslavách odsoudil During celebrations criticized

Adpositions
Adpositions

in diesem Jahr kommen in this year come

APP R PDAT NN VV INF
ADP DET NOUN VERB

NK DET obl case
MO

5/1
Lahore in closed

Adpositions
Ельцин считал что стране нужен экономист.

Yeltsin realized that by-country needed economist.
hasta que se fijen los comicios

until that themselves fix the elections
Una sala ha dovuto essere sgomberata a room has had to-be evacuated.
Verb Groups

Based-on can for-example be determined whether

Op_grond_van kan bijvoorbeeld worden vastgesteld of
apples, oranges and lemons
Coordination: Mel’čuk [de]

Äpfel, Orangen und Zitronen
äpplen, apelsiner och citroner
Coordination: Mel’čuk [ja]

リンゴ と オレンジ と レモン
ringe to orenji to remon
Coordination: Mel’čuk [tr]
Coordination: Prague [cs]

Diagram:

- **jablka**
- **pomeranče**
- **citróny**

Connections:
- Coordinating conjunction (conj)
- Auxiliar conjunction (AuxX)
- Punctuator (punct)
- Coordinating conjunction (cc)
Coordination: Prague [te]
Coordination: Half-Prague [ro]

mere  portocale  și  lămâi

conj  conj  conj  rel.conj.  rel.conj.

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

蘋果，桔子和檸檬

pingguo, juzi he ningmeng
ябълки, портокали и лимони
jabălki, portokali i limoni
Coordination: Half-Stanford [da]

æbler, appelsiner og citroner
Coordination: Tesnière [hu]
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

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

Graphical representation of sentence structure.
Interaction of Structures in Prague

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

Coord
AuxP_M
Atr
conj
cc
conj
case
nmod
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 liberce

trains to and from prague or from liberce
Interaction of Structures in Prague

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

Trains to and from Prague or from Liberec.
Round-Trip: Shared Dependents

Atr

Coord

XXX_M

XXX_M

čeští studenti a učitelé
Czech students and teachers

Atr

Coord

XXX_M

XXX_M

XXX

conj

cc

amod

čeští studenti a učitelé
Czech students and teachers

čeští studenti a učitelé
Czech students and teachers
Round-Trip: Nested Coordination

- 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

he could have seen her

he could have seen her
Tree-Rewriting Toolkits

- Udapi
  - http://udapi.github.io/ (main page)

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

```
NP
  |     S
  |
  |      VP
  |
  N    V    NP    NP
  Paul gave N    C    N
          Peter two pears

    root
    ↓    ↓    ↓    ↓    ↓
  nsubj subj obj nummod

Paul gave Peter two pears
```

Peter gave two pears to Paul
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, \textit{proximity} of constituents)
Example

S
  NP
  N
  John

VP
  V
  loves

NP
  N
  Mary

S
  VP
  N
  John

V
  loves

NP
  N
  Mary

nsubj

John loves Mary

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

- Actor
- Location
- Patient
- COREF
- COPY

Clause_M

Coord
he will meet Mary today and (he) (will) (meet) Clara tomorrow