

Enhanced Universal Dependencies

Daniel Zeman

📅 July 23, 2026



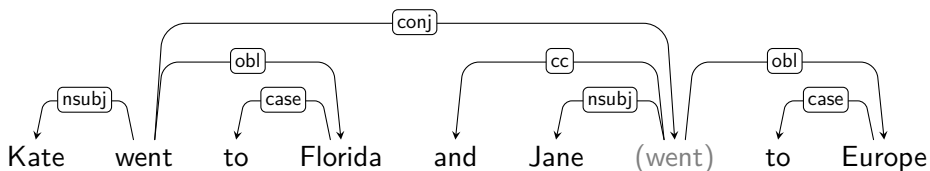
Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics



unless otherwise stated



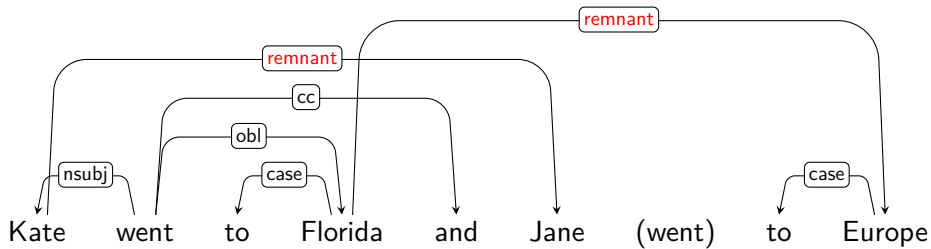
Deleted Predicates in Coordination



- Some treebanks would use an **empty node** to represent the second *went*.
- UD **enhanced representation** allows empty nodes!
- But the basic representation sticks with the overt words.

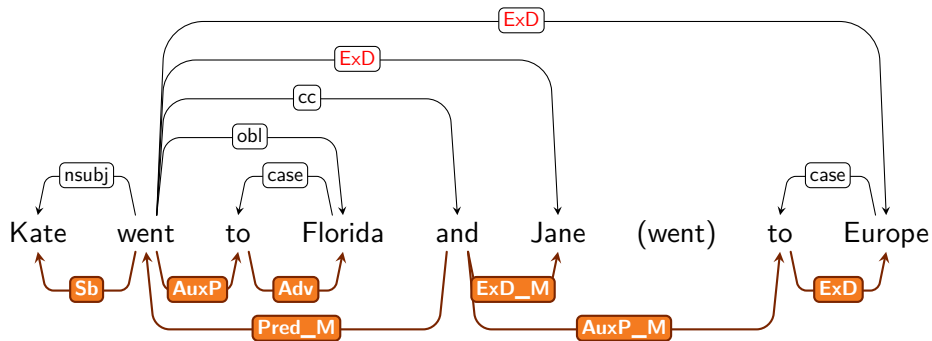


UD V1: The remnant Relation



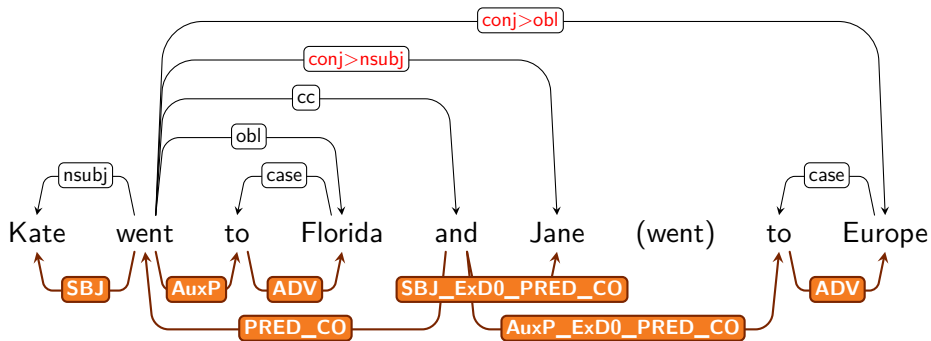


PDT: The ExD Relation



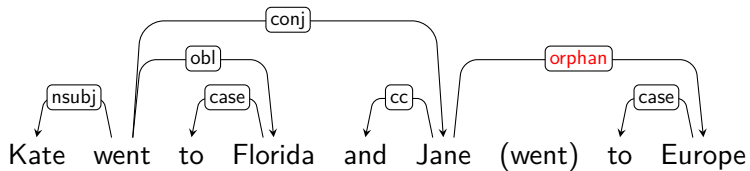


Perseus Treebanks: Chained Relations



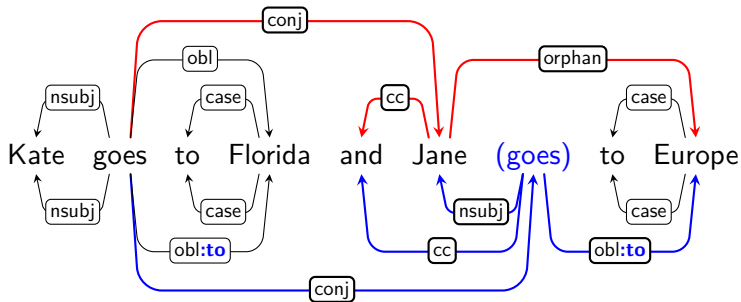


UD V2 Basic Dependencies: The orphan Relation



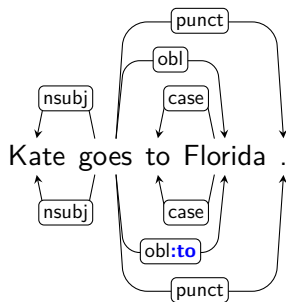


UD Enhanced Dependencies: Gapping and Stripping



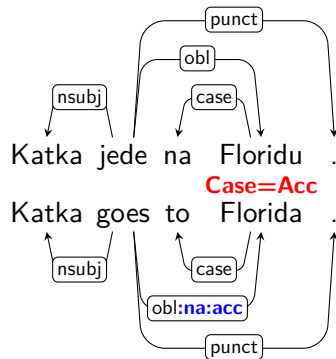
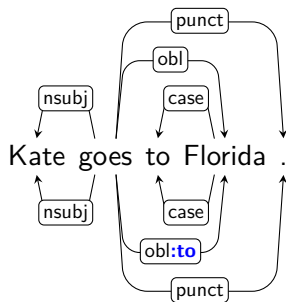


Enhanced UD: Case Information in Dependency Label



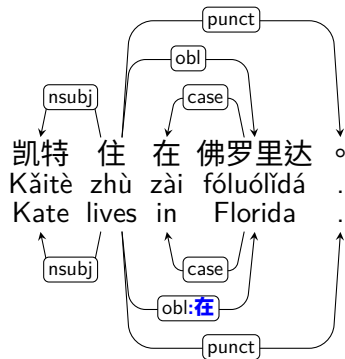
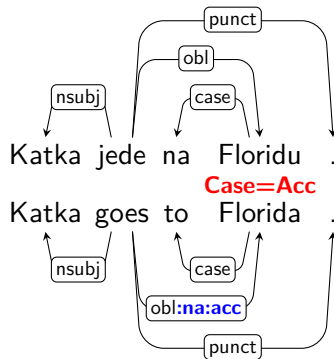
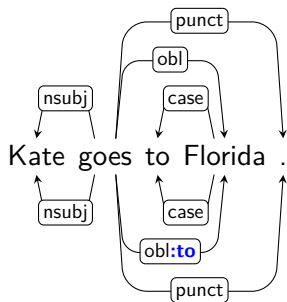


Enhanced UD: Case Information in Dependency Label



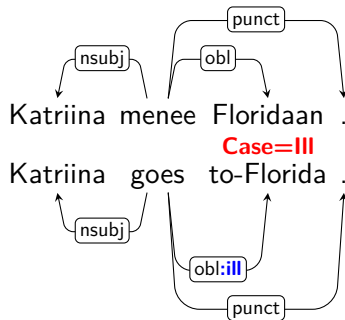
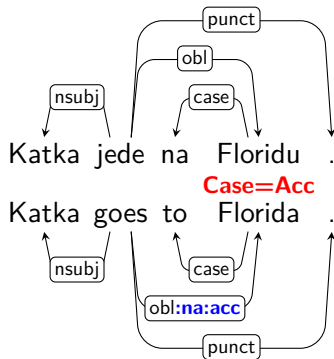
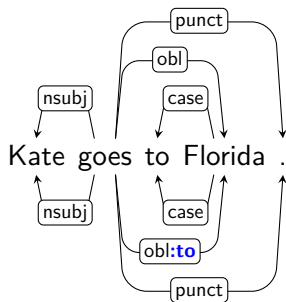


Enhanced UD: Case Information in Dependency Label



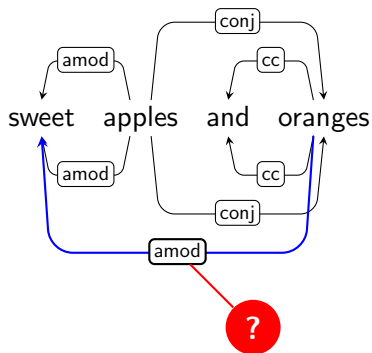


Enhanced UD: Case Information in Dependency Label



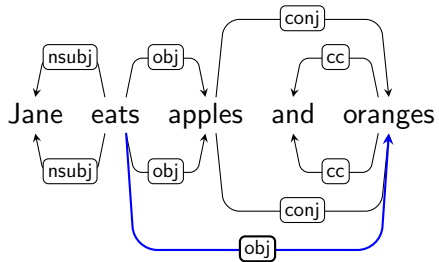


Enhanced UD: Shared Dependent of Coordination



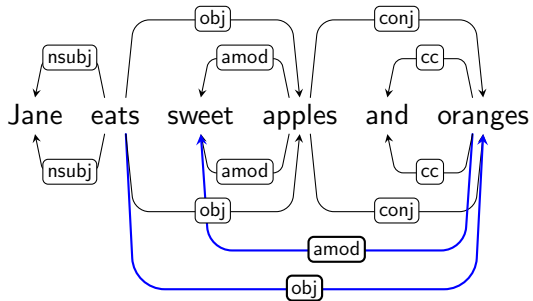


Enhanced UD: Parent of Coordination



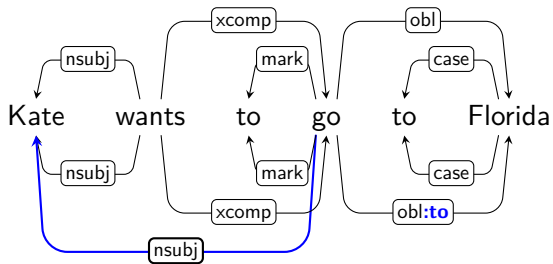


Enhanced UD: Coordination



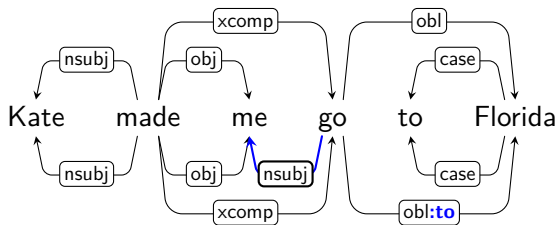


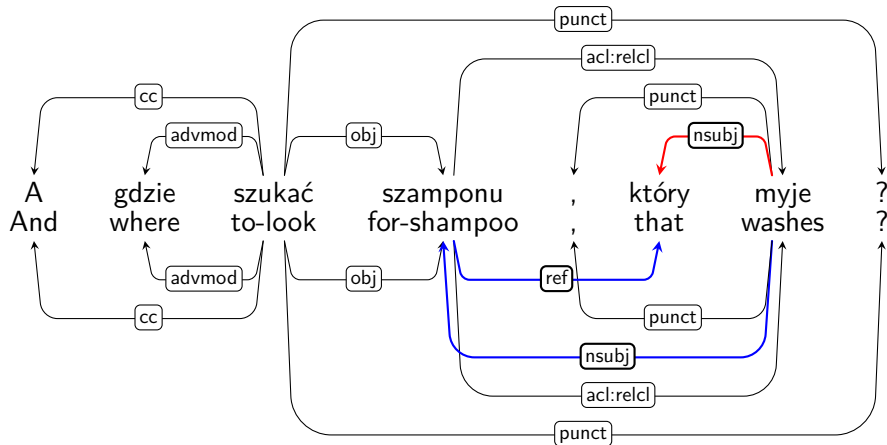
Enhanced UD: External Subject of Controlled Predicate





Enhanced UD: External Subject in Object-Control Construction

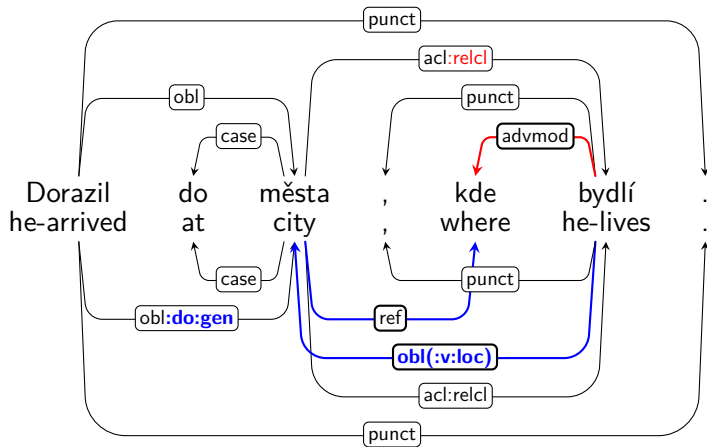




“And where to look for shampoo that works?”

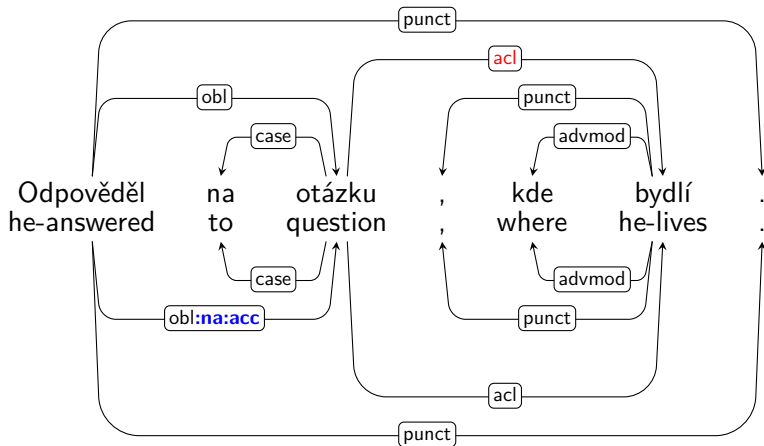


How to Recognize Relative Clauses



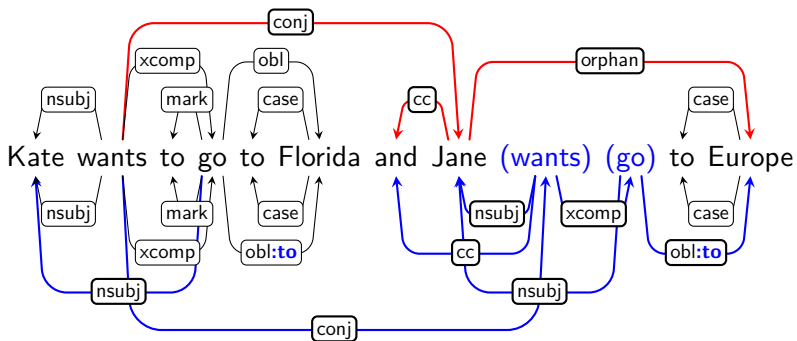


How to Recognize Relative Clauses





Enhanced UD: Gapping + Control Verb Construction



Enhanced UD: Six Enhancements

- Null nodes for **gapping** (34 treebanks in UD 2.18)
 - But in some of them the null nodes are actually not used (only) for gapping
- Coordination: **Common parent** (40 treebanks)
- Coordination: **Shared dependents** (38 treebanks)
- External subjects of **controlled predicates** (34 treebanks)
- Cyclic dependencies to/from **relative clauses** (31 treebanks)
- **Case**-enhanced dependency labels (31 treebanks)

- All 6 types: 19 treebanks
- At least 1 type: 47 treebanks, 29 languages
- Only basic UD: 306 treebanks

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD
- **Gapping**: embeddings for similarity of arguments

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD
- **Gapping**: embeddings for similarity of arguments
- Coordination:
 - **Parent propagation**: deterministic
 - **Shared dependents**: heuristics (human desirable!)

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD
- **Gapping**: embeddings for similarity of arguments
- Coordination:
 - **Parent propagation**: deterministic
 - **Shared dependents**: heuristics (human desirable!)
- **External subjects**: heuristics (subject vs. object control)

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD
- **Gapping**: embeddings for similarity of arguments
- Coordination:
 - **Parent propagation**: deterministic
 - **Shared dependents**: heuristics (human desirable!)
- **External subjects**: heuristics (subject vs. object control)
- **Relative clauses**: need `ac1:relcl` and list of relative pronouns

- Part of Stanford CoreNLP (Java)
- Rules from basic to enhanced UD
- **Gapping**: embeddings for similarity of arguments
- Coordination:
 - **Parent propagation**: deterministic
 - **Shared dependents**: heuristics (human desirable!)
- **External subjects**: heuristics (subject vs. object control)
- **Relative clauses**: need `ac1:relc1` and list of relative pronouns
- **Case-enhanced labels**: deterministic



Conversion from non-UD Data: Extra Information?

- Analytical layer of Prague-style treebanks: **shared dependents of coordination** are known!

