How useful are Enhanced Universal Dependencies for semantic interpretation?

Jamie Y. Findlay & Dag T. T. Haug

University of Oslo

March 22, 2022

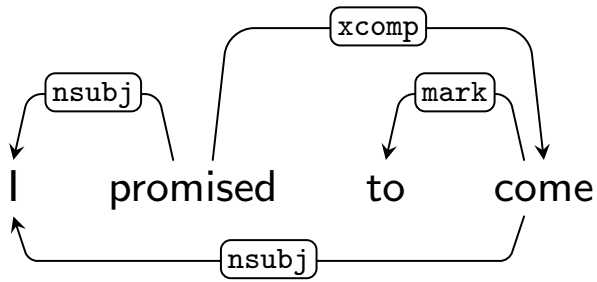
E-UD and semantic interpretation

- Context: **Universal Natural Language Understanding** project – universal semantic parsing for UD using Glue Semantics.
- Basic UD does not provide sufficient information to derive predicate-argument structures
- Enhanced UD (E-UD) helps, but tradeoff with coverage
  - Only 31/213 UD treebanks have useful E-UD edges.
- How far can we recreate E-UD information from basic UD using universal, linguistically-based heuristics?
- Six types of E-UD annotation:
  - empty (null) nodes for elided predicates
  - propagation of incoming dependencies to conjuncts
  - additional subject relations for open complement clauses (`xcomps`)
  - propagation of outgoing dependencies from conjuncts
  - coreference in relative clause constructions
  - modifier labels that contain the preposition or other case-marking information

Additional subject edges for `xcomp`s

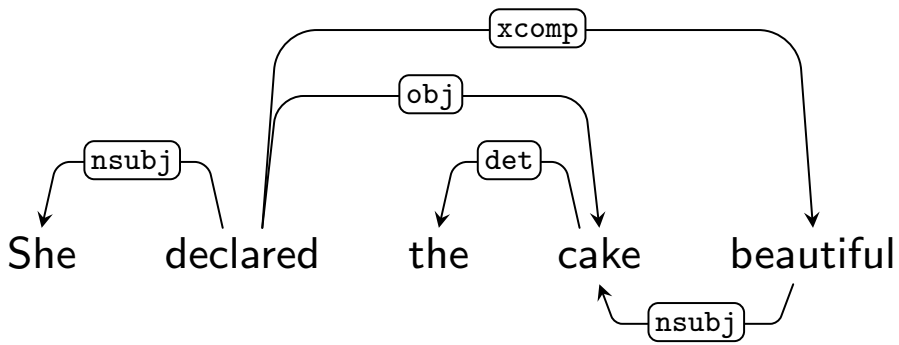
- UD guidelines specify 2 types of `xcomp`:
  - raising/control

(1)



  - secondary predication

(2)

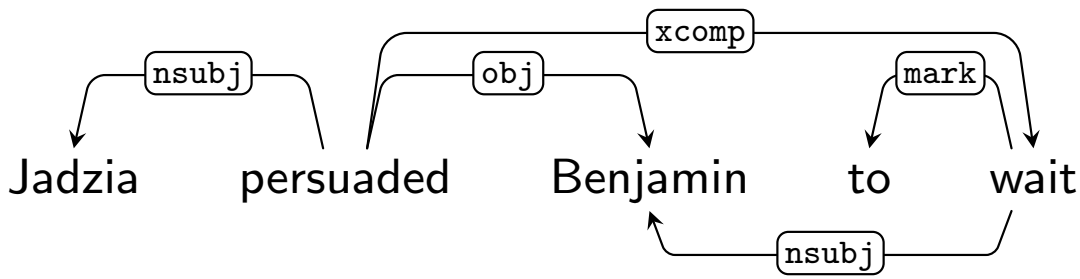


- Excludes lots of cases of 'shared subjects' (e.g. optional or arbitrary control, such as purpose clauses) – focusses on *grammatically determined* instances of control.

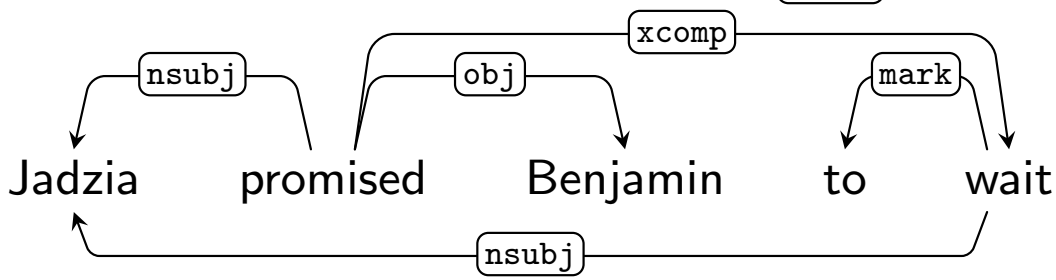
Additional `subj` edges for `xcomps`

- E-UD information required for correct predicate-argument structure.
- In principle, cannot be determined based purely on UD tree:

(3) a.



b.

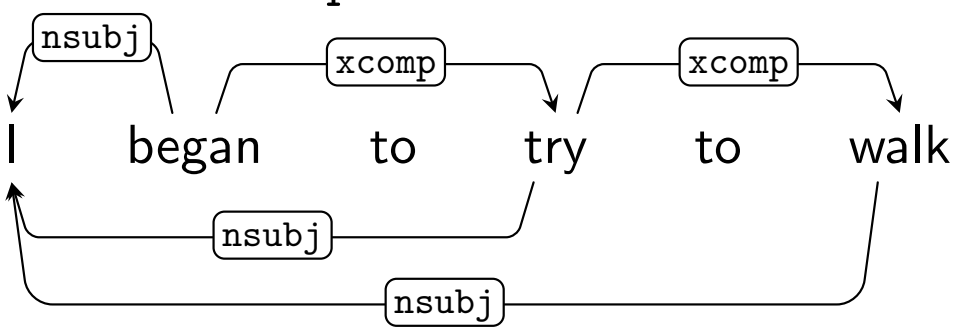


Our heuristic

- But there are cross-linguistic trends which can be exploited:

1. If head of `xcomp` has an object dependent (`obj`, `iobj`, `ccomp`), then that is the controller.
2. Otherwise if head of `xcomp` has a subject dependent (`nsubj`, `csubj`), then that is the controller.
3. If neither, check whether the head is itself an `xcomp` dependent; if so, go back to 1. with the higher `xcomp` as the starting point.
4. Otherwise, don't add a controller edge.

(4) *Recursive* `xcomp`s



Results

Corpus name	Precision	Recall
Albanian-TSA	66.67%	66.67%
Belarusian-HSE	60.24%	73.82%
Bulgarian-BTB	71.87%	75.69%
Czech-CAC	67.00%	78.12%
Czech-FicTree	63.03%	78.83%
Czech-PDT	74.60%	83.53%
Czech-PUD	55.36%	75.61%
Dutch-Alpino	37.00%	90.01%
Dutch-LassySmall	30.92%	83.94%
English-EWT	93.84%	90.76%
English-GUM	92.70%	94.24%
English-GUMReddit	93.23%	89.21%
English-PUD	92.00%	88.09%
Finnish-TDT	56.89%	60.29%
Italian-ISDT	76.94%	80.28%
Latvian-LVTB	69.30%	87.88%
Lithuanian-ALKSNIS	59.78%	78.16%
Polish-LFG	95.26%	94.41%
Slovak-SNK	68.81%	84.03%
Swedish-PUD	87.62%	84.29%
Swedish-Talbanken	86.31%	86.13%
Ukrainian-IU	95.34%	88.39%
AVERAGE	72.49%	82.38%

Table: Performance of the heuristic used for adding external subjects

Error analysis

- Source of 100 errors in Dutch-Alpino and English-GUM (numbers don't sum to 100 because the 2 heuristic errors also involved E-UD errors):

Corpus	Basic UD	E-UD	Heuristic	Not an error
Dutch-Alpino	3	95	1	2
English-GUM	25	74	1	1

Basic UD: *Do you know what it's like to be* **chased** *by the Ghost of Failure while staring through Victory's door?* (`GUM_interview_messina-36`).

- relation: ~~`xcomp`~~ → `ccomp`

E-UD: *[. . .] several municipalities have started asking religious schools to* **pay** *taxes [. . .]* (`GUM_news_taxes-3`)

- controller: ~~*municipalities*~~ → *schools*

Ik wist dat mijn helft van het schema **open** *ligt.* 'I knew my half of the schedule is open' (`WR-P-P-H-0000000006\WR-P-P-H-0000000006.p.1.s.5`)

- No controller marked (should be *helft*)

Results (only marked controllers)

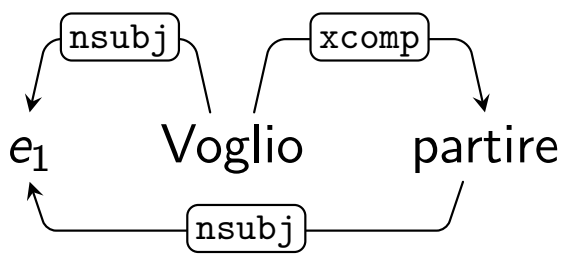
Corpus name	Precision (controllers marked)	Recall
Albanian-TSA	100.00%	66.67%
Belarusian-HSE	76.70%	73.82%
Bulgarian-BTB	98.49%	75.69%
Czech-CAC	85.23%	78.12%
Czech-FicTree	88.58%	78.83%
Czech-PDT	87.89%	83.53%
Czech-PUD	78.81%	75.61%
Dutch-Alpino	92.97%	90.01%
Dutch-LassySmall	94.57%	83.94%
English-EWT	95.65%	90.76%
English-GUM	99.46%	94.24%
English-GUMReddit	99.20%	89.21%
English-PUD	94.52%	88.09%
Finnish-TDT	99.43%	60.29%
Italian-ISDT	82.39%	80.28%
Latvian-LVTB	95.03%	87.88%
Lithuanian-ALKSNIS	93.79%	78.16%
Polish-LFG	98.37%	94.41%
Slovak-SNK	87.98%	84.03%
Swedish-PUD	89.39%	84.29%
Swedish-Talbanken	90.92%	86.13%
Ukrainian-IU	98.50%	88.39%
AVERAGE	92.18%	82.38%

Table: Performance of the heuristic used for adding external subjects when only marked controllers are considered

xcomps without controllers

- Validation check to ensure xcomps are properly controlled?
  - No, because of pro-drop or extraction gaps:
    - *Voglio* **partire** 'I want to leave' (Italian)
    - *the man I told to* **leave** (this would be remedied in E-UD, though – see below)
- Not having controllers is problematic:
  1. Harder to verify/enhance automatically.
  2. We lose linguistic information: the control relationship is not represented in the case of pro-drop, even though the coreference is just as obligatory.
     - A failure in terms of UD's *universal* goals.
- Solution: represent pro-dropped arguments as 'empty' nodes in the basic UD tree:

(5)

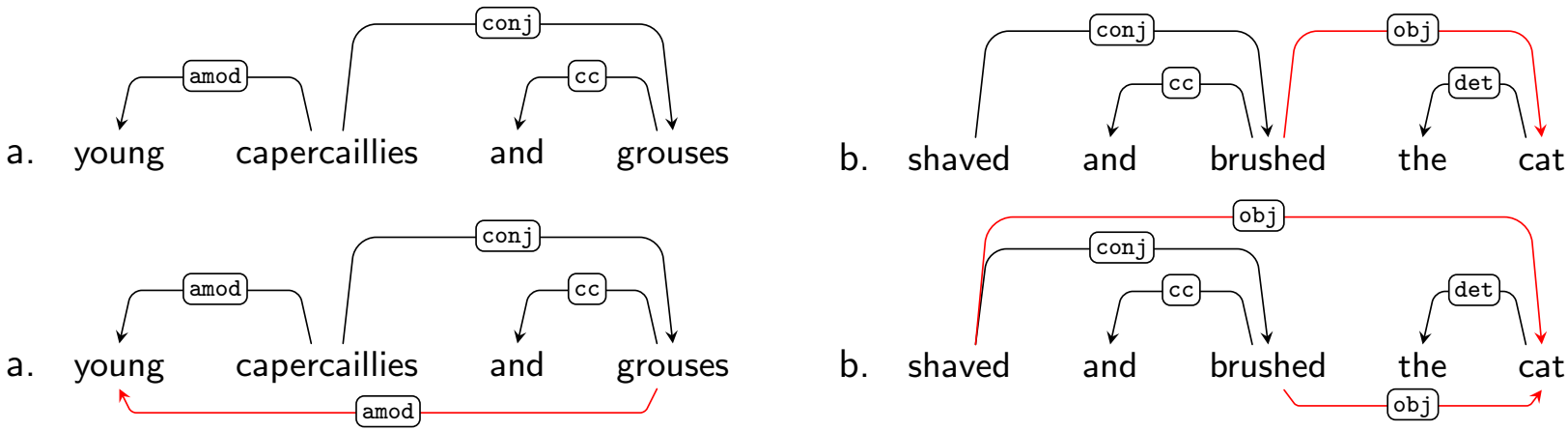


Propagation of dependents

Is *young* and *the cat* shared across conjuncts?

(6) a. young capercaillies and grouses
b. shaved and brushed the cat

Basic UD disambiguates (6b) but not (6a):



Heuristics

- Resolution of such cases depend on a complex interplay of word order, valency, semantic restrictions and world knowledge
- UD lacks most of this information, but there are some cues
  - Most core relations and grammatical relations are functional, so don't propagate if this leads to duplication
  - Most verbs need a subj (`nsubj` or `csubj`)
  - Word order also helps because the basic UD disambiguates cases where the alternative analysis has the dependent be private to the *second* conjunct
- Heuristic 1 (*never* applied if it leads to duplication):
  a. Propagate `nsubj` and `csubj`
  b. Propagate distant dependents (i.e. dependents of the first conjunct that occur linearly to the right of the coordinator)
- Heuristic 2 is as 1 except we only propagate distant objects

Results

	1: subj + dist		1a: all and only subj		1b: dist only		2: subj + dist obj	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
Arabic-PADT	61.8	77.5	27.1	14.4	87.1	64.1	27.7	14.9
Bulgarian-BTB	40.2	100.0	63.6	100.0	0.2	0.2	62.2	100.0
Czech-CAC	81.7	50.0	64.9	18.5	95.3	33.0	66.7	20.2
Czech-FicTree	69.7	46.6	66.6	36.9	82.7	10.1	66.8	37.3
Czech-PDT	69.6	56.0	54.9	25.9	89.9	30.8	56.2	27.5
Dutch-Alpino	50.9	48.3	59.7	32.9	38.5	15.6	59.7	33.4
Dutch-LassySmall	50.8	48.3	51.6	32.7	49.7	16.0	51.6	32.7
English-EWT	61.1	100.0	98.7	93.4	10.9	7.7	98.1	99.1
English-GUM	59.6	83.7	97.9	78.2	10.0	6.1	97.6	83.3
English-GUMReddit	69.8	80.6	100.0	78.3	11.8	4.7	99.0	80.6
English-PUD	63.7	100.0	98.9	93.0	12.5	8.0	99.0	99.0
Finnish-TDT	84.5	38.9	84.3	26.1	85.2	13.2	84.8	27.2
Italian-ISDT	63.5	97.2	92.4	93.5	10.1	5.4	92.6	96.7
Latvian-LVTB	83.2	38.2	79.5	28.1	95.9	10.7	80.5	29.8
Lithuanian-ALKSNIS	59.9	36.1	48.3	19.3	77.4	18.2	51.7	22.1
Polish-LFG	69.5	31.9	67.9	29.6	100.0	2.9	68.2	30.0
Polish-PDB	81.0	34.9	72.8	21.4	98.6	13.8	74.1	22.8
Polish-PUD	87.9	33.4	81.8	20.7	100.0	13.4	82.7	22.0
Slovak-SNK	53.0	58.8	40.6	34.7	92.8	25.8	42.6	37.6
Swedish-PUD	63.6	100.0	100.0	93.8	11.3	7.3	100.0	100.0
Swedish-Talbanken	72.0	100.0	99.1	88.8	24.2	12.2	99.2	96.9
Ukrainian-IU	26.4	48.4	31.5	37.8	17.1	11.2	31.2	38.7

- Some treebanks have used something similar to our heuristic 2
- Ukrainian is manual gold-standard, but only 40% done
- Boldfaced treebanks are gold standard

Results

	1: subj + dist		1a: all and only subj		1b: dist only		2: subj + dist obj	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
Arabic-PADT	61.8	77.5	27.1	14.4	87.1	64.1	27.7	14.9
Czech-CAC	81.7	50.0	64.9	18.5	95.3	33.0	66.7	20.2
Czech-FicTree	69.7	46.6	66.6	36.9	82.7	10.1	66.8	37.3
Czech-PDT	69.6	56.0	54.9	25.9	89.9	30.8	56.2	27.5
Finnish-TDT	84.5	38.9	84.3	26.1	85.2	13.2	84.8	27.2
Latvian-LVTB	83.2	38.2	79.5	28.1	95.9	10.7	80.5	29.8
Lithuanian-ALKSNIS	59.9	36.1	48.3	19.3	77.4	18.2	51.7	22.1
Polish-LFG	69.5	31.9	67.9	29.6	100.0	2.9	68.2	30.0
Polish-PDB	81.0	34.9	72.8	21.4	98.6	13.8	74.1	22.8
Polish-PUD	87.9	33.4	81.8	20.7	100.0	13.4	82.7	22.0
Slovak-SNK	53.0	58.8	40.6	34.7	92.8	25.8	42.6	37.6

- 1a achieves mixed results, but 1b generally achieves good precision
- Manual error analysis on the Lithuanian treebank

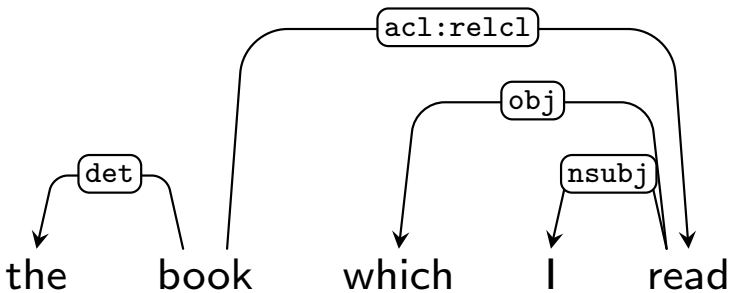
	Impers. verb	Subj. shift	Basic UD	E-UD
1a) Subject	12	6	19	13
1b) Dist. dep.	—	—	48	2

- Some subject shift errors could be captured with feature analysis, but would often need an impersonal verb feature
- In general, there are no good high-coverage heuristics, so this is a case where enhanced dependencies would have been very important

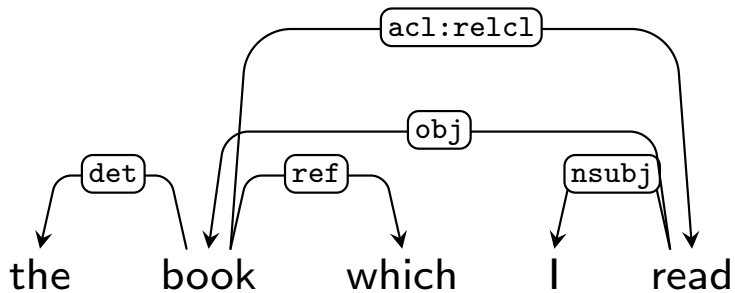
Enhanced dependencies for relative clauses

The basic and E-UD analyses are really different perspectives on relative clauses, but when the relativizer is missing, the enhanced but not the basic version makes it possible to identify the gap.

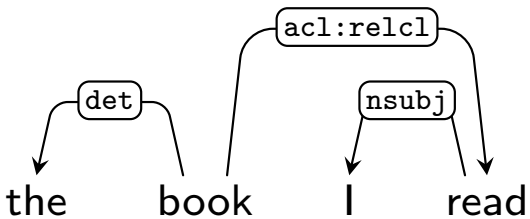
a. *Basic UD graph*



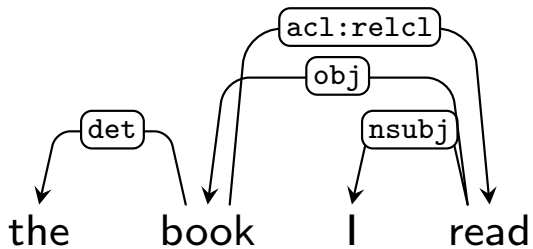
b. *Enhanced UD graph*



a. *Basic UD graph*



b. *Enhanced UD graph*



E-UD for relative clauses

- The enhanced analysis of relative clauses has had poor uptake, perhaps because it looks more like an alternative than an enhancement
- Most treebanks either does not have the E-UD, or only have it when it is predictable from the basic dependencies
- Only Tamil-TTB, Ukrainian-IU and Belarusian-HSE have more than 10 non-predictable E-UD edges for relative clauses
- In Tamil and Ukrainian we can to some extent predict the gap based on the Accessibility Hierarchy (see full paper)
- But overall, it seems like it would be better to revise the guidelines
- One option is to make indication of the gap obligatory already in the basic UD using an empty node

Summary

- A lot E-UD edges are simply added with automatic heuristics with no apprently little post-processing
- Manual E-UD for relative clauses has had particularly little success
- (Small-scale) manual error analysis suggests there are quality issues
- In some cases you could get more quicker by making the basic UD more expressive
- For propagation of dependents, E-UD does seem like the way to go, but doing that properly requires *much* work, especially because E-UD edges are not very useful until you can trust the *absence* of an edge as well