

Evaluation of Universal Dependency parsers for Hungarian

Evelin Ács¹, Gábor Recski²

¹ Department of Theoretical Linguistics

Eötvös Loránd University

² Department of Automation and Applied Informatics

Budapest University of Technology and Economics

Summary

- We present the results of manual error analysis performed on sample outputs of three recent dependency parser systems.
- We separate error classes that are sensitive to the technicalities of UD annotation and evaluation principles.
- Main error types are root elements, coordination, modifiers and structural ambiguity.

Background

- Universal Dependencies projects: over 100 UD treebanks in 60 languages
- CoNLL Shared Task: 'Multilingual Parsing from Raw Text to Universal Dependencies'
- The top 3 parsers are Stanford's UnstableParser, Cornell's C2L2, and Stuttgart's IMS.
- The Hungarian section of the UD dataset has been created using the Szeged Dependency Treebank.

LAS, CLAS and UAS scores of all three parsers

	Overall			Hungarian		
	LAS	CLAS	UAS	LAS	CLAS	UAS
UnstableParser (Stanford)	76.30	72.57	81.30	77.56	76.08	82.35
C2L2 (Cornell)	75.00	70.90	80.32	76.55	74.36	82.07
IMS (Stuttgart)	74.42	70.18	79.90	73.55	70.87	79.90

Evaluation

- We inspected manually the analyses given by each of the three parsers on the first 50 sentences of the Hungarian test data.
- We grouped errors both by the types of dependency relations they involved and by the types of errors.

Types of erroneous edges

	UnstableParser	C2L2	IMS
punct	43	punct 46	punct 44
cc	13	cc 17	cc 17
det	11	det 16	det 11
advmod	9	conj 6	advmod 11
amod	7	conj-nmod 5	amod 7
conj	7	cc-advmod 5	amod-conj 6

Error types

- **ROOT ELEMENTS:** In nearly a fifth of all sentences observed, parsers assigned the root dependency to the wrong word, i.e. they failed to identify the main predicate of the sentence.
 - These errors are worthy of attention not only because of their frequency but because they are usually responsible for several further erroneous edges – if the parser misses the main predicate, it is likely to miss relations of each of its dependents.
 - **COORDINATION:** In UD, conjunctions are treated asymmetrically: one of the coordinated elements is considered the head of the conjunction and others are connected only to this element (via the conj relation) but not to any other word in the sentence.
- (1) *Ezek is leginkább csak a januári napokban, amikor a fa már kiszáradt és egy csillagszóró is lángba boríthatja – mondta az alezredes.*
 already dry-out-PAST and a sparkler too flame-INE cover-DEF – say-PAST the colonel.

But only in the days of January, when the tree is dry and a sparkler might burn it down – said the colonel.

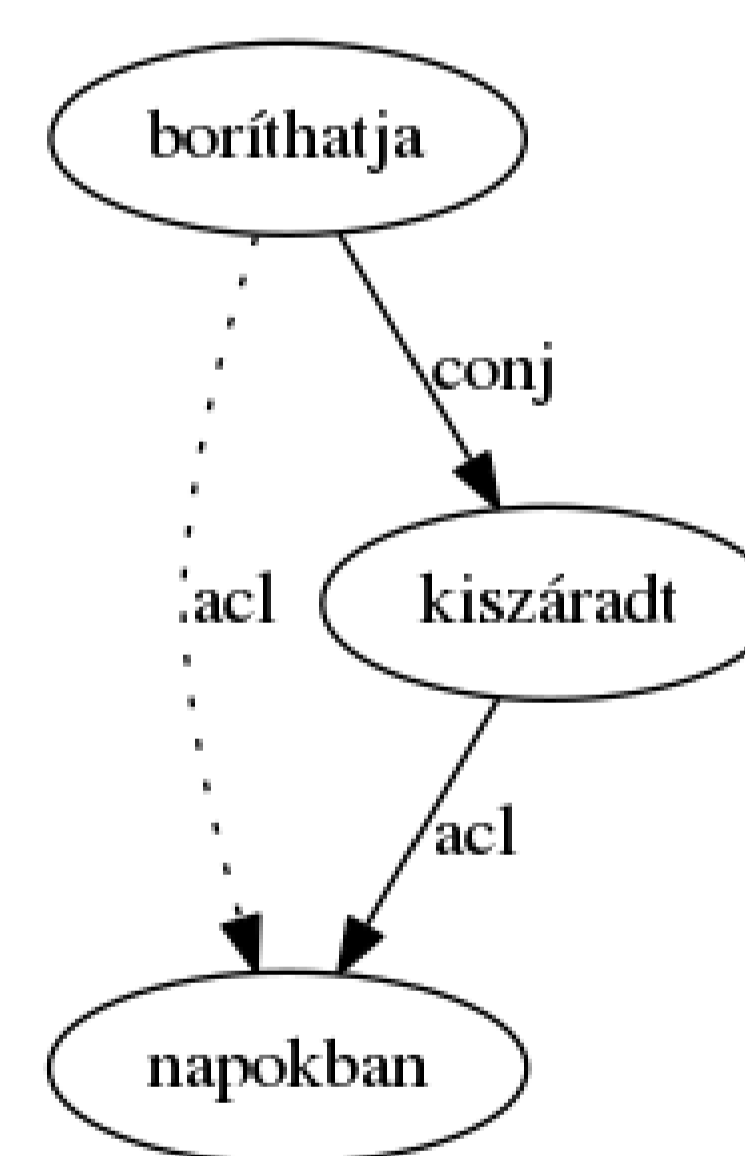


Figure 1: Partial analysis of (1). Dotted edges represent the solution of both UnstableParser and C2L2.

- **MODIFIERS:** The UD relations nmod and amod represent the dependencies between a noun and its nominal or adjectival modifier, respectively. Similarly, the advmod relation connects adverbs to predicates or modifiers.
- A large portion of errors were caused by parsers mixing the above three labels on edges that were otherwise correctly identified, i.e. they connected the modifiers to the right word.
- In most cases either the output contains an incorrect POS-tag and the label reflects the same mistake, or labels were assigned incorrectly despite a correct POS-tag, but in 4 cases the parser's solution can be justified.

Gold and IMS analyses of a noun phrase

	az	Y2K	problémát
	the	Y2K	problem-ACC
gold POS	DET	NOUN	NOUN
gold dependency	det	nmod	
IMS POS	DET	ADJ	NOUN
IMS dependency	det	amod	

- **STRUCTURAL AMBIGUITY:** The final group involves sentences that are structurally ambiguous and whose parses are consistent with a different constituent structure than the one reflected by the gold dependency annotation.
- (2) *a Péterfy kórház sürgősségi belgyógyászati és klinikai toxikológiai osztálya*
 department-POSS

The department of emergency internal medicine and clinical toxicology
 The emergency department of internal medicine and clinical toxicology

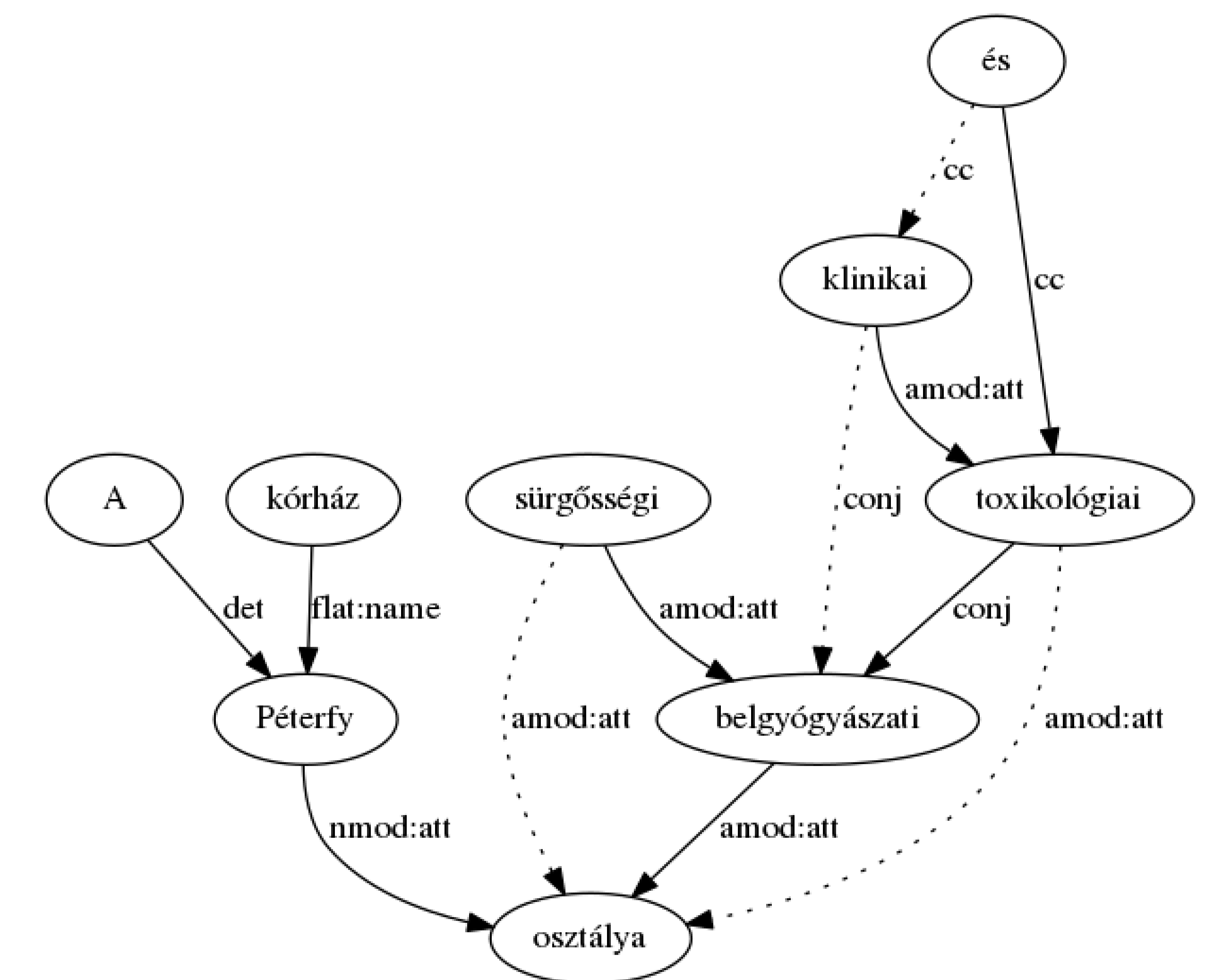


Figure 2: Analysis of (2). Dotted edges represent the solution of UnstableParser.