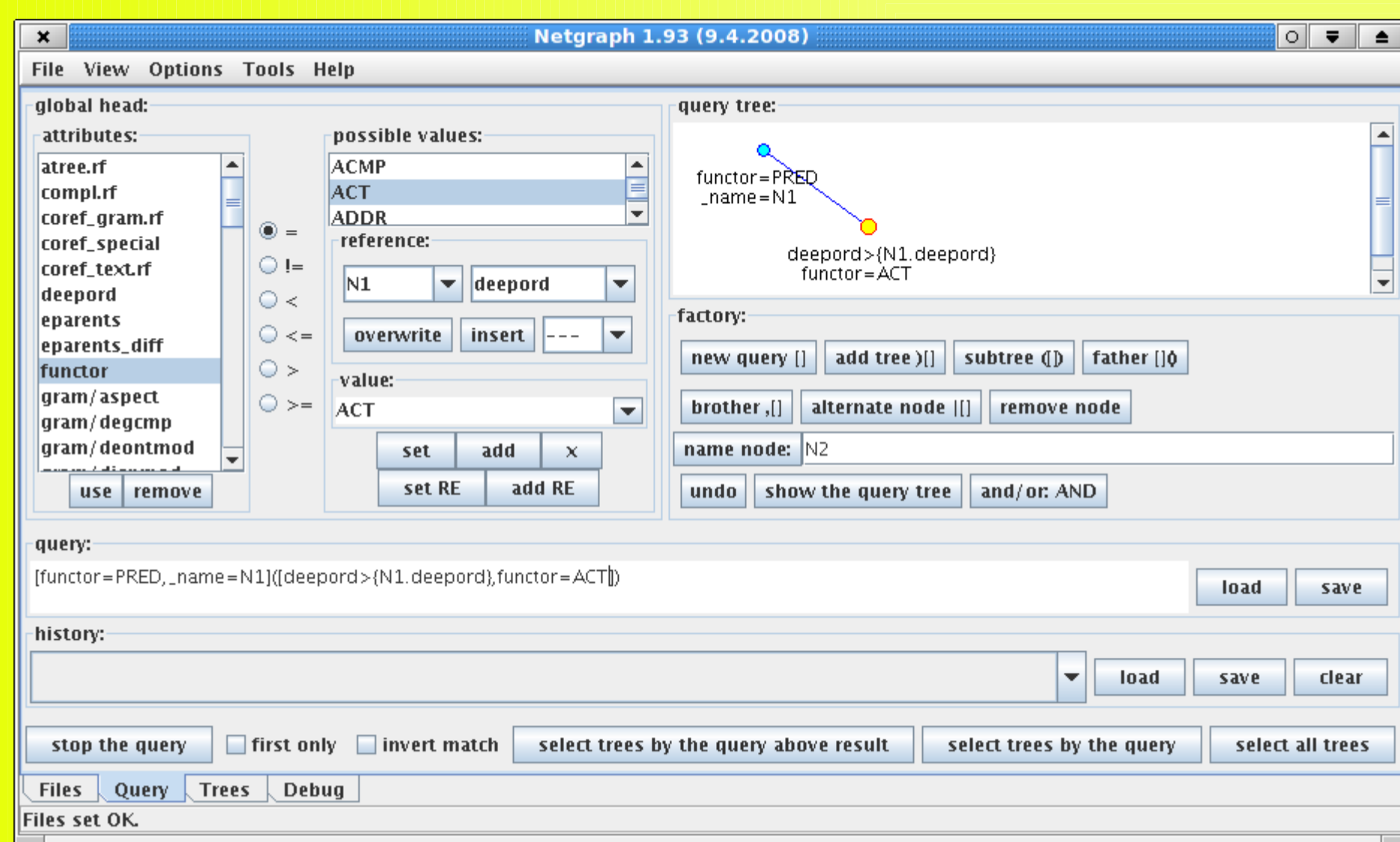


Does Netgraph Fit the Prague Dependency Treebank 2.0?

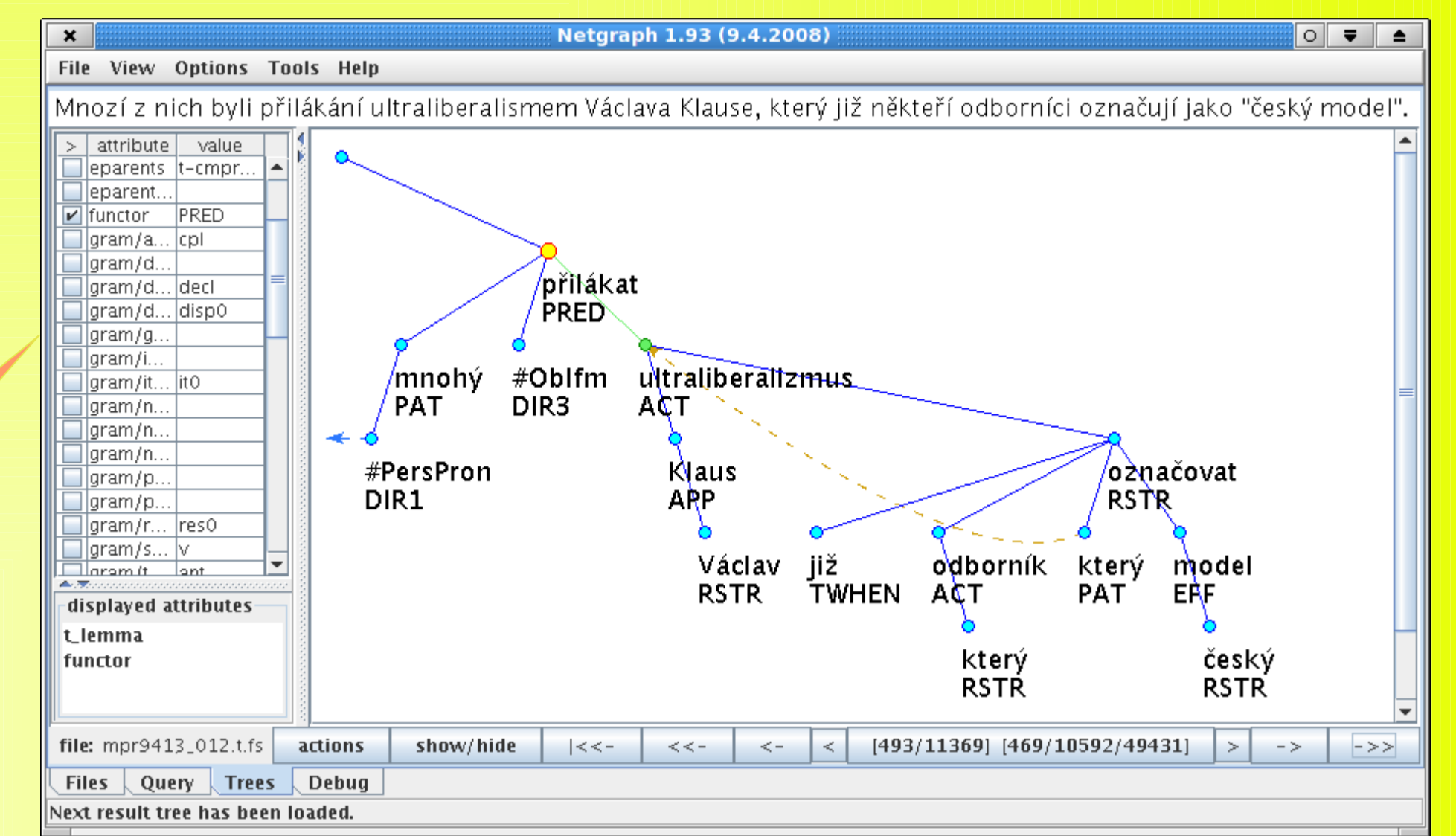
Jiří Mirovský, Charles University in Prague, Czech Republic, mirovsky@ufal.mff.cuni.cz



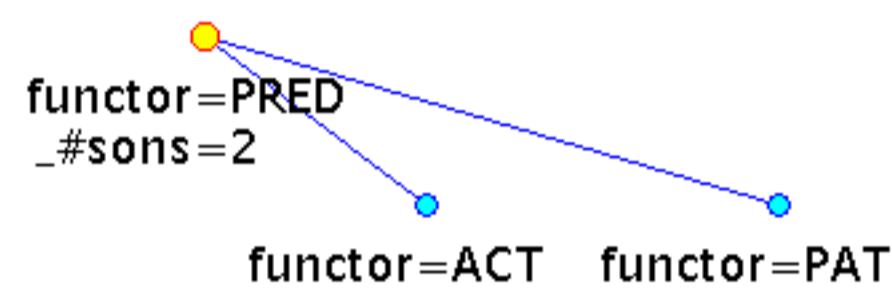
Netgraph is a simple yet powerful tool for searching in linguistic treebanks. Do its features fit PDT 2.0?

A graphically created query:
A **PRED**icate governing a more dynamic **ACT**or

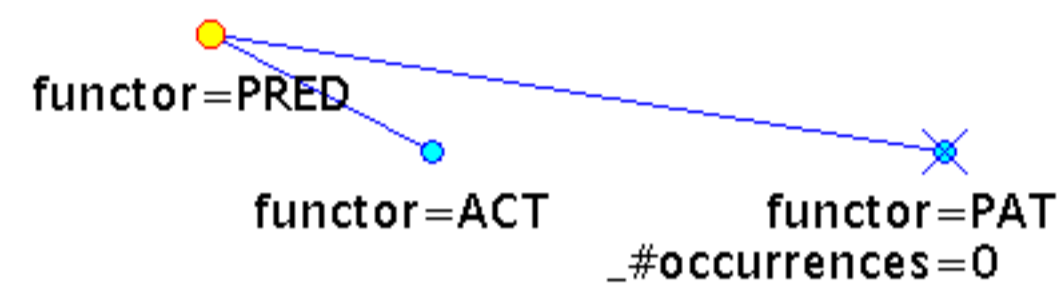
A possible result tree



VALENCY

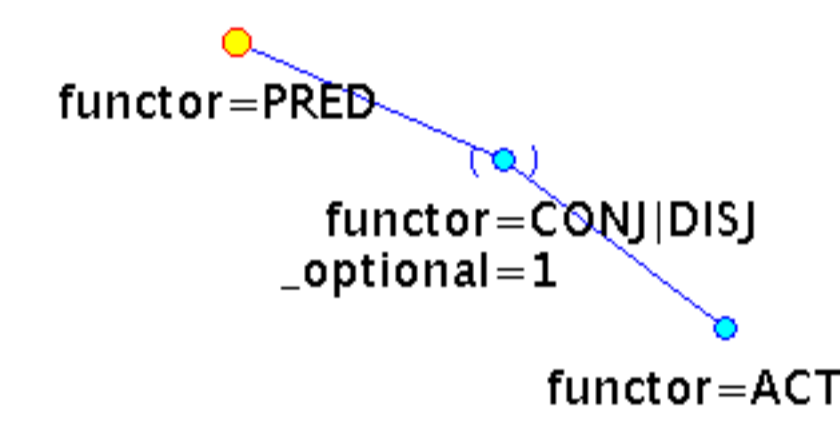


A **PRED**icate that governs an **ACT**or, a **PAT**ient and nothing else (`_#sons=2`)



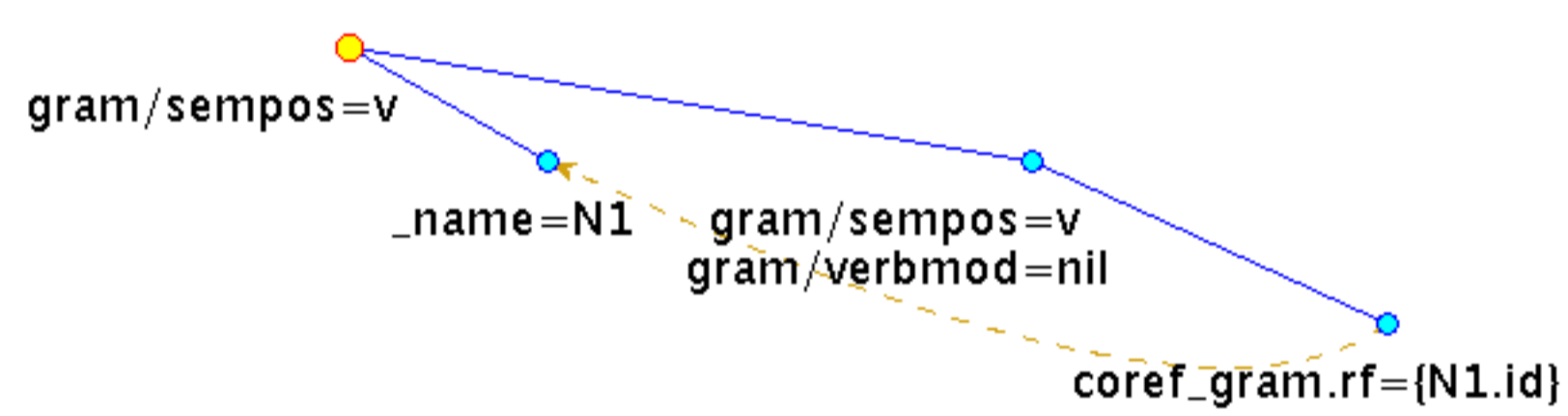
A **PRED**icate that governs an **ACT**or and anything else but a **PAT**ient

COORDINATION etc.



The coordinating node becomes optional and can be skipped.

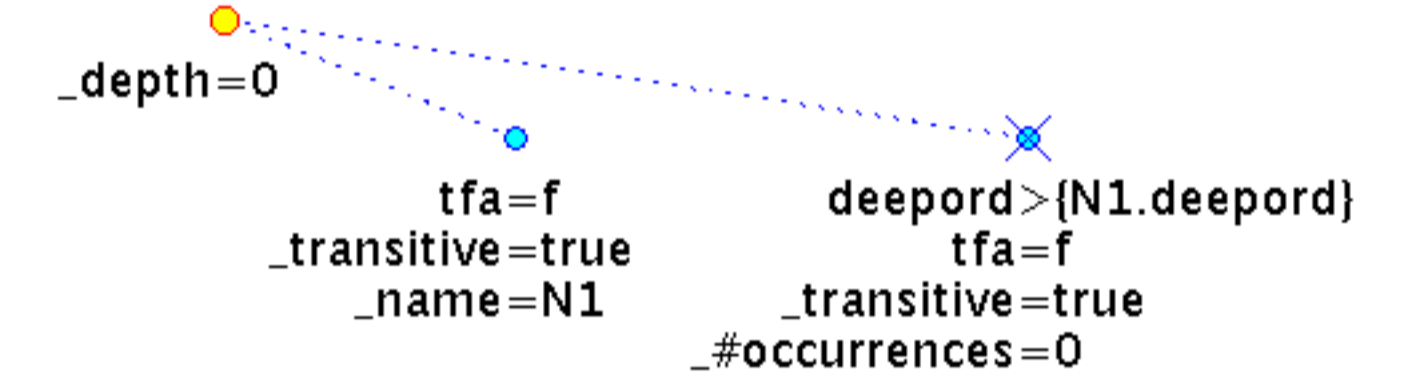
CO-REFERENCE



Type-1 control construction (a type of the grammatical co-reference): an infinitive depends on a control verb

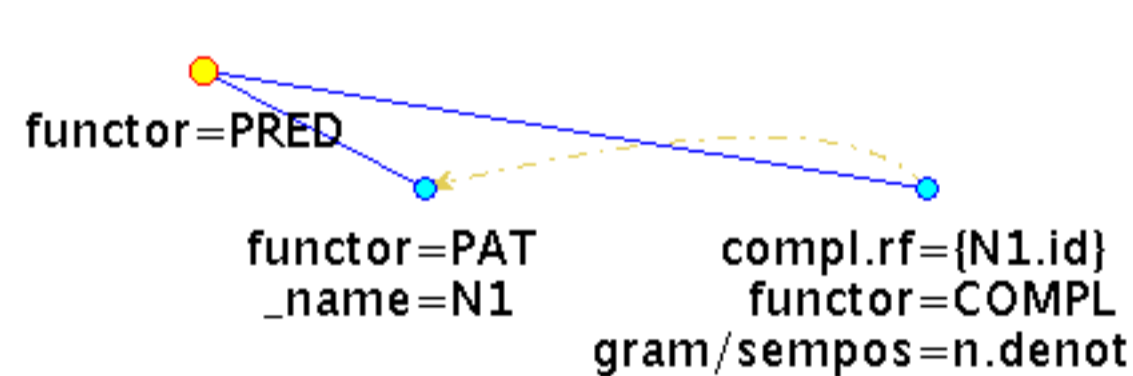
Queries in Netgraph for various linguistic phenomena annotated in PDT 2.0

FOCUS PROPER



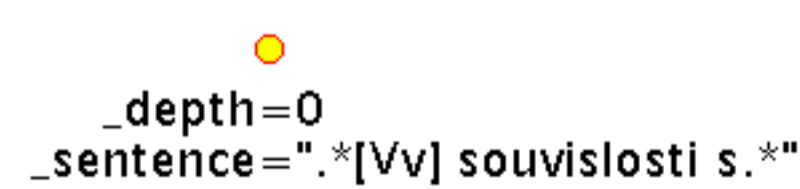
Focus proper is the most dynamic contextually non-bound part of the sentence (rightmost in the tree);

DUAL DEPENDENCY



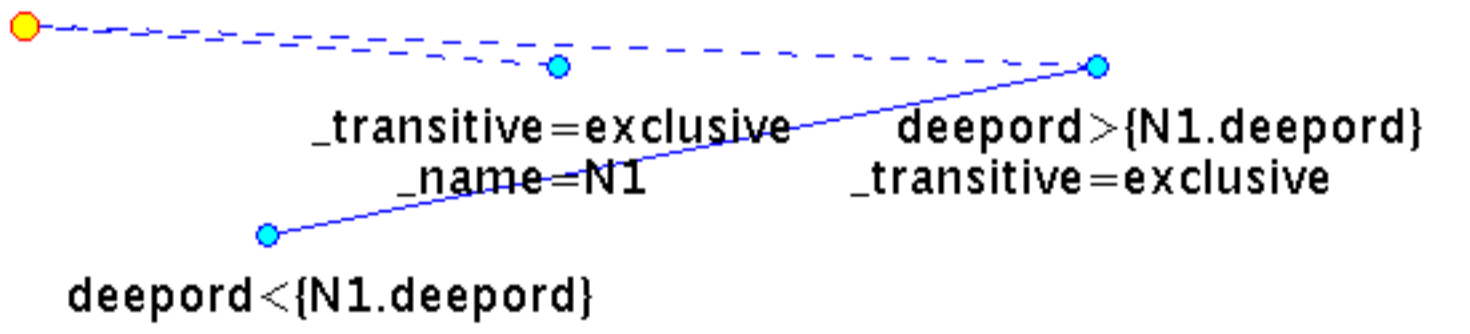
Second dependency of the **COMPL**ement goes to a **PAT**ient

IDIOMS



Searching in the surface form of the sentence with a regular expression for the phrase "v souvislosti s" ~ "in relation to"

(NON-)PROJECTIVITY



One of four possible configurations of a node and an edge causing non-projectivity

other phenomena

Searching for other linguistic phenomena annotated in PDT 2.0 is described in the paper.

Conclusion

All linguistic phenomena annotated in PDT 2.0 can be searched and studied using Netgraph. Therefore, Netgraph fits PDT 2.0 well.

Future Work

- further simplification of the query language
- parallelization / indexing for faster searching
- improvements to the tool

References

- Netgraph home page: <http://quest.ms.mff.cuni.cz/netgraph>
- PDT 2.0: <http://ufal.mff.cuni.cz/pdt2.0>