

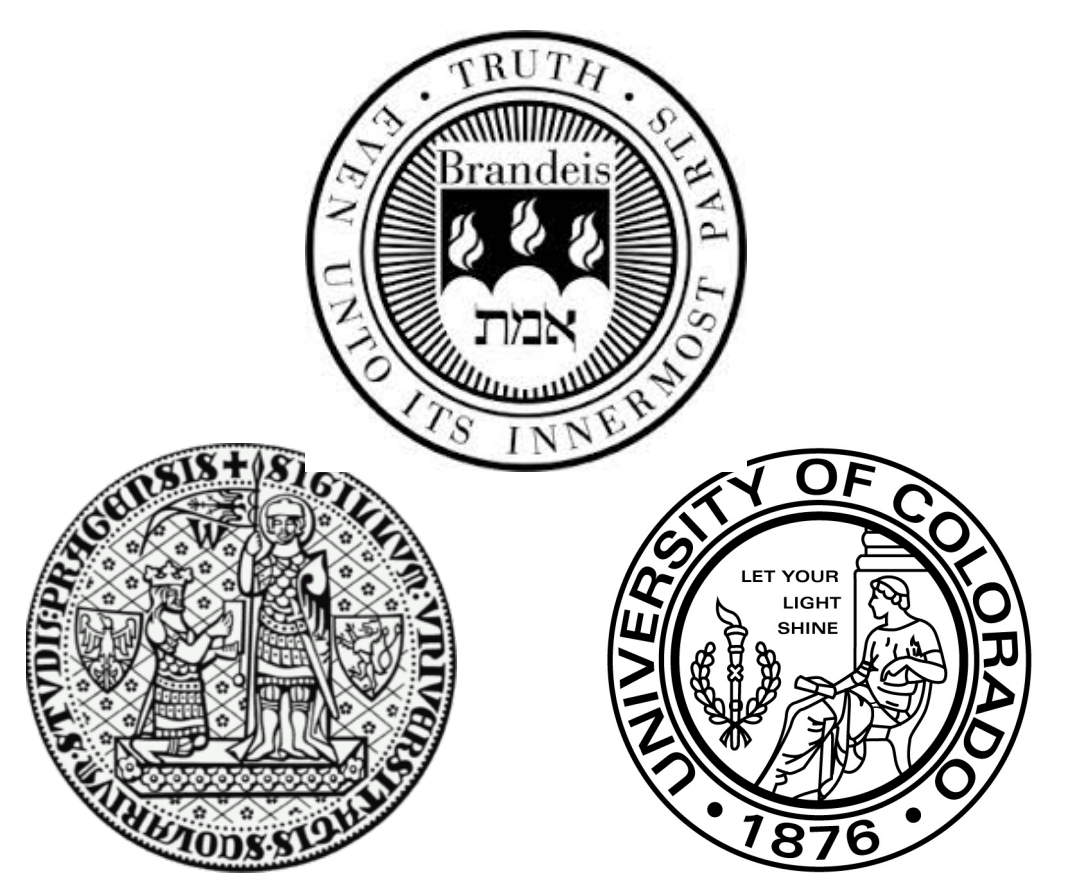
# Not an Interlingua, but Close: Comparison of English AMRs to Chinese and Czech

Nianwen Xue\*, Ondřej Bojar†, Jan Hajič†, Martha Palmer#, Zdeňka Urešová†, Xuhong Zhang\*

\*Computer Science Department, Brandeis University, USA.

†Charles University in Prague, Institute of Formal and Applied Linguistics, Czech Republic

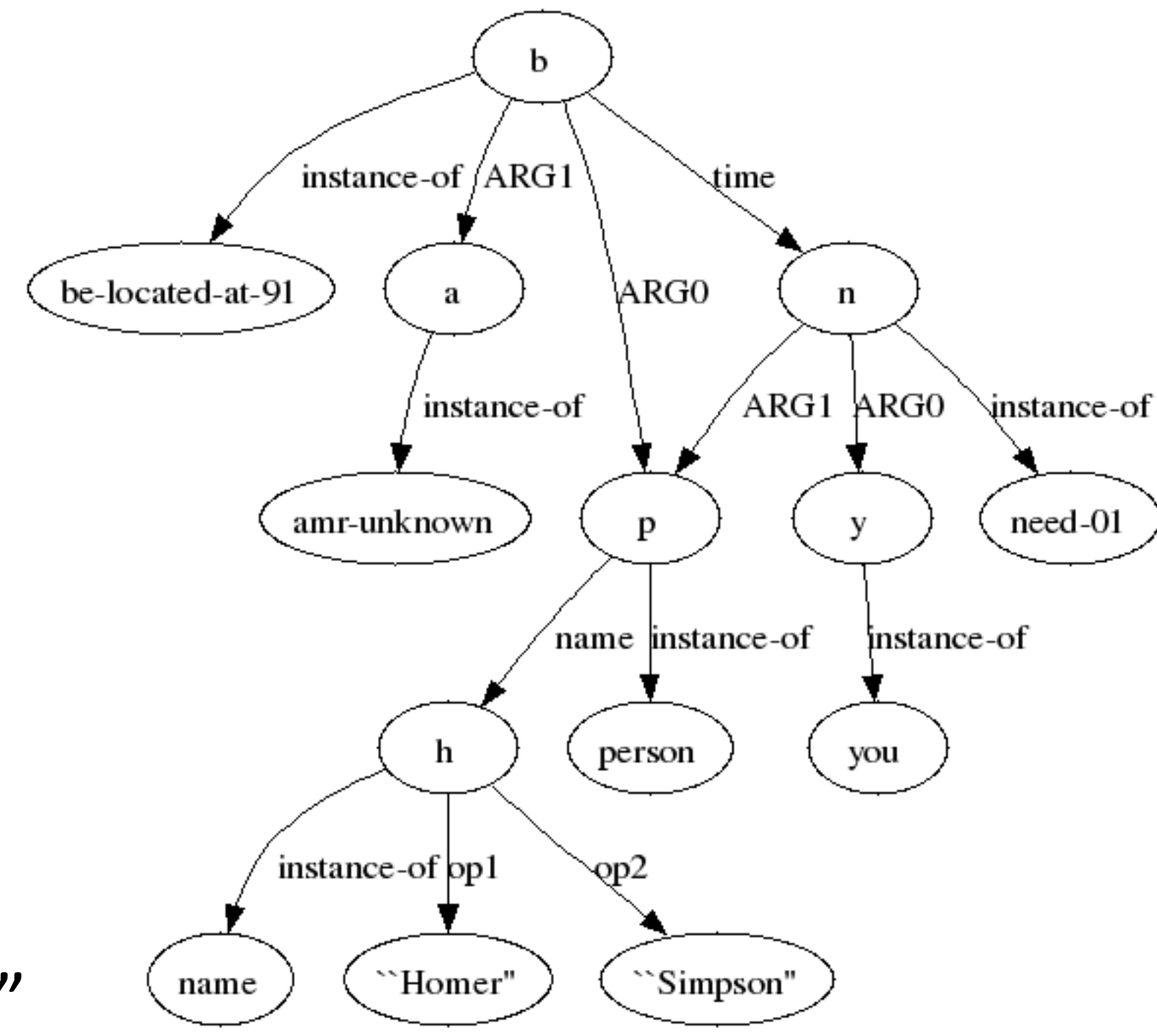
#Department of Linguistics, University of Colorado at Boulder, USA



An Abstract Meaning Representation (AMR) is a rooted, directional and labeled graph that represents the meaning of a sentence. It focuses on *relations* between *concepts*.

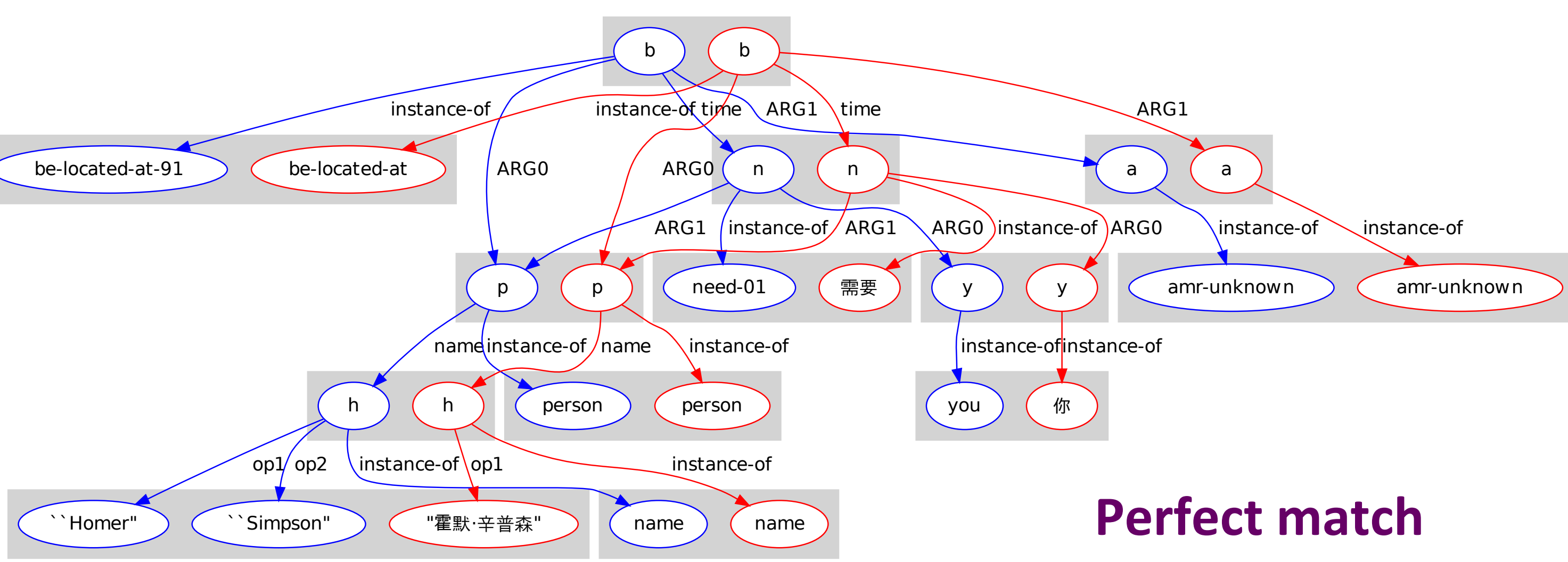
(B / be-located-at-91  
:ARG0 (p / person  
:name (h / name  
:op1 "Homer"  
:op2 "Simpson"))  
:ARG1 (a / amr-unknown)  
:time (n / need-01  
:ARG0 (y / you  
:ARG1 p))

"Where is Homer Simpson when you need him?"



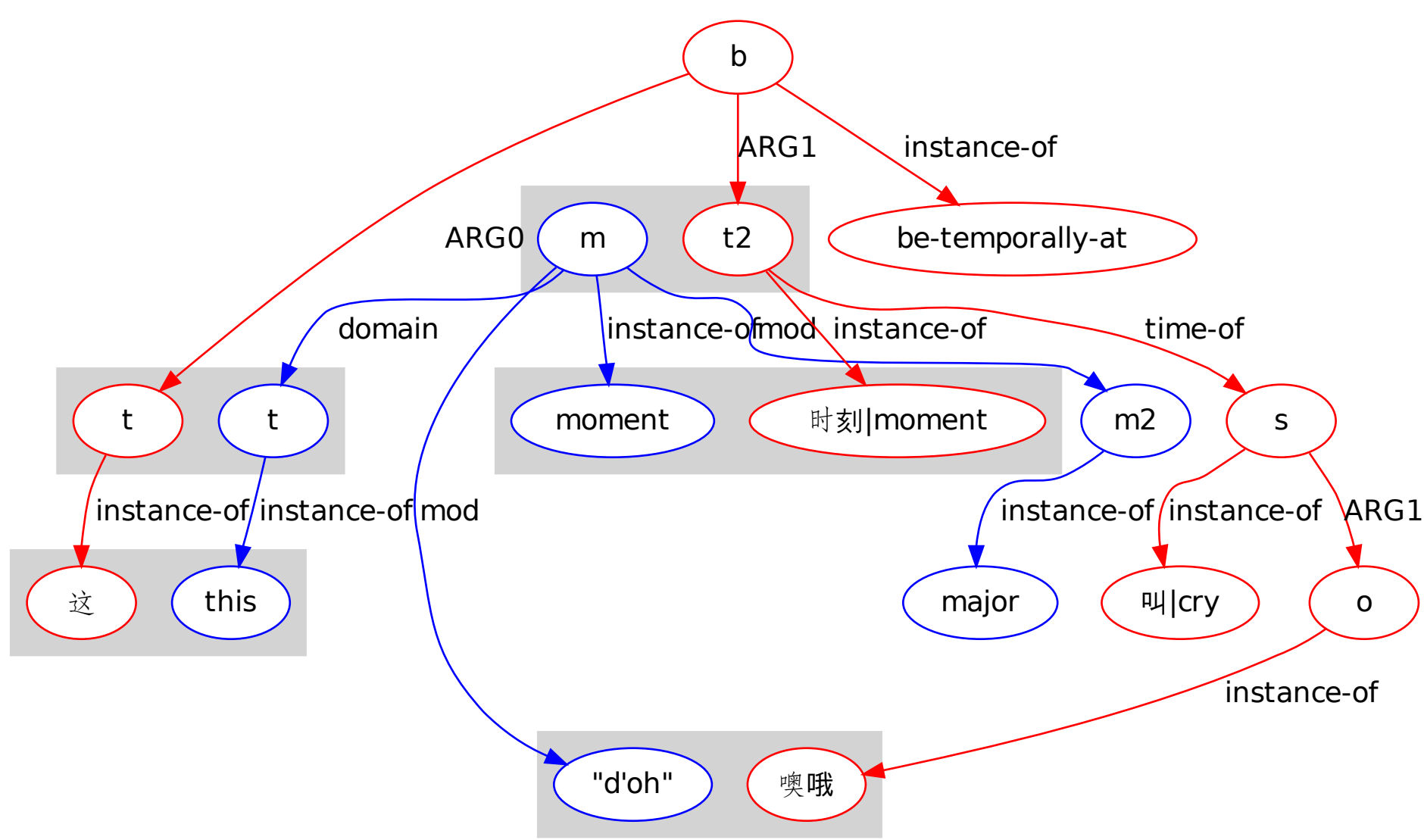
AMRs abstract away from word category, word order, morphological variation. Some function words are also not represented. Many sources of cross-lingual differences are removed, creating better alignment for bilingual sentence pairs.

当你需要他时，霍默辛普森在哪里？  
Where is Homer Simpson when you need him?

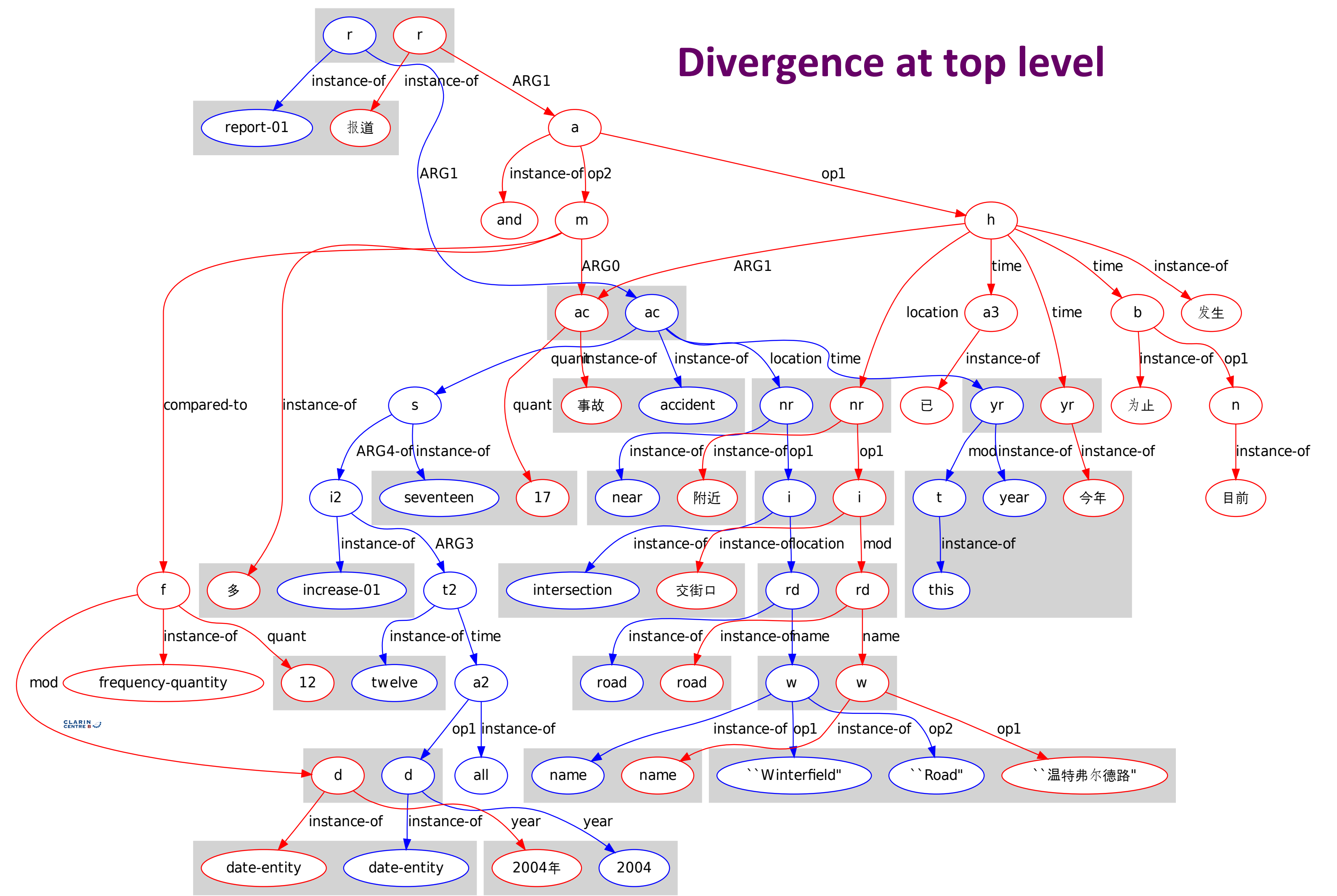


Perfect match

这是一个大叫“噢哦！”的时刻。  
This is a major "D'oh!" moment.



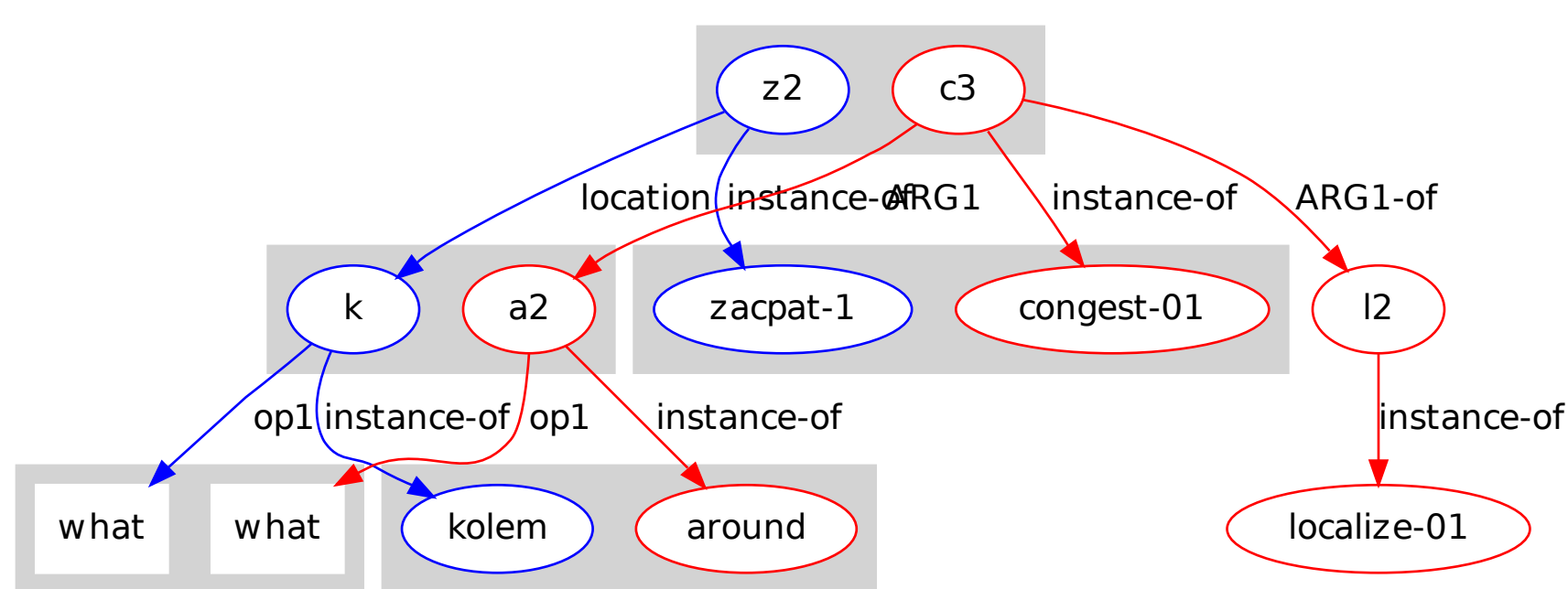
Difference in lexicalization and annotation choice



Divergence at top level

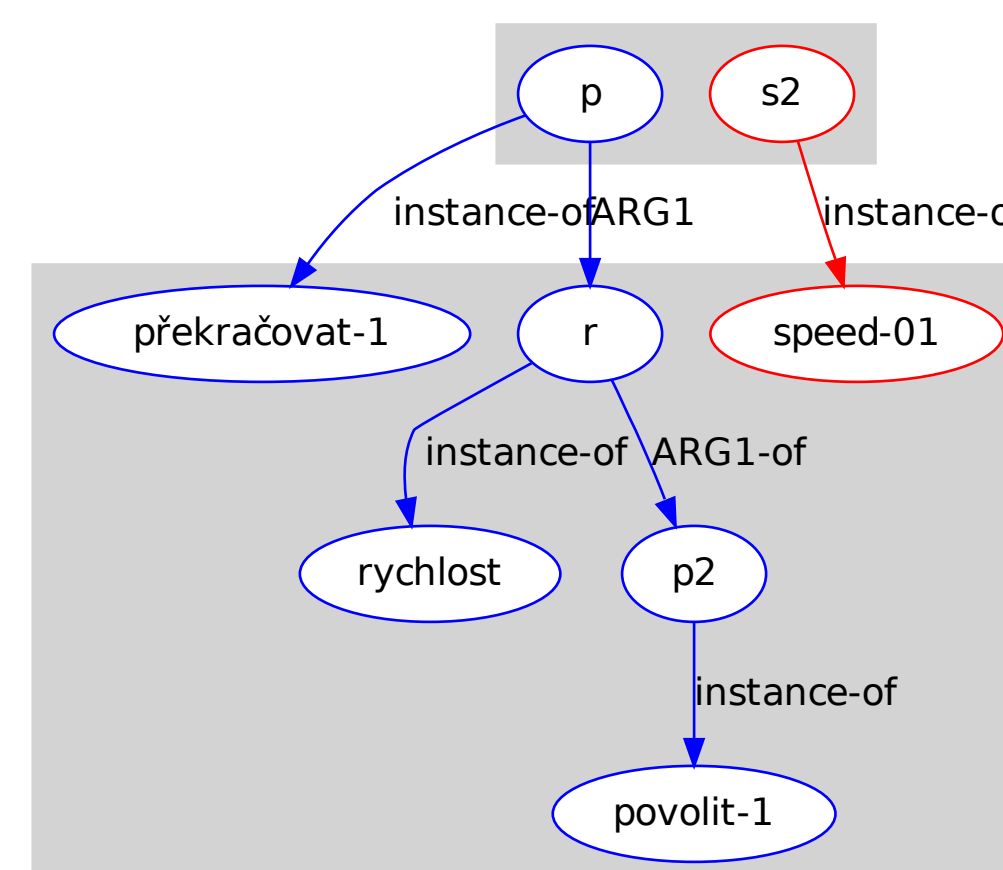
今年，到目前为止，温特弗尔德路的交叉口附近据报道已发生了17起事故，比2004年全年的12起还多。  
17 accidents have been reported near the intersection of Winterfield Road so far this year, up from 12 for all of 2004.

"Zácpa kolem čeho"  
"Localized congestion around what"

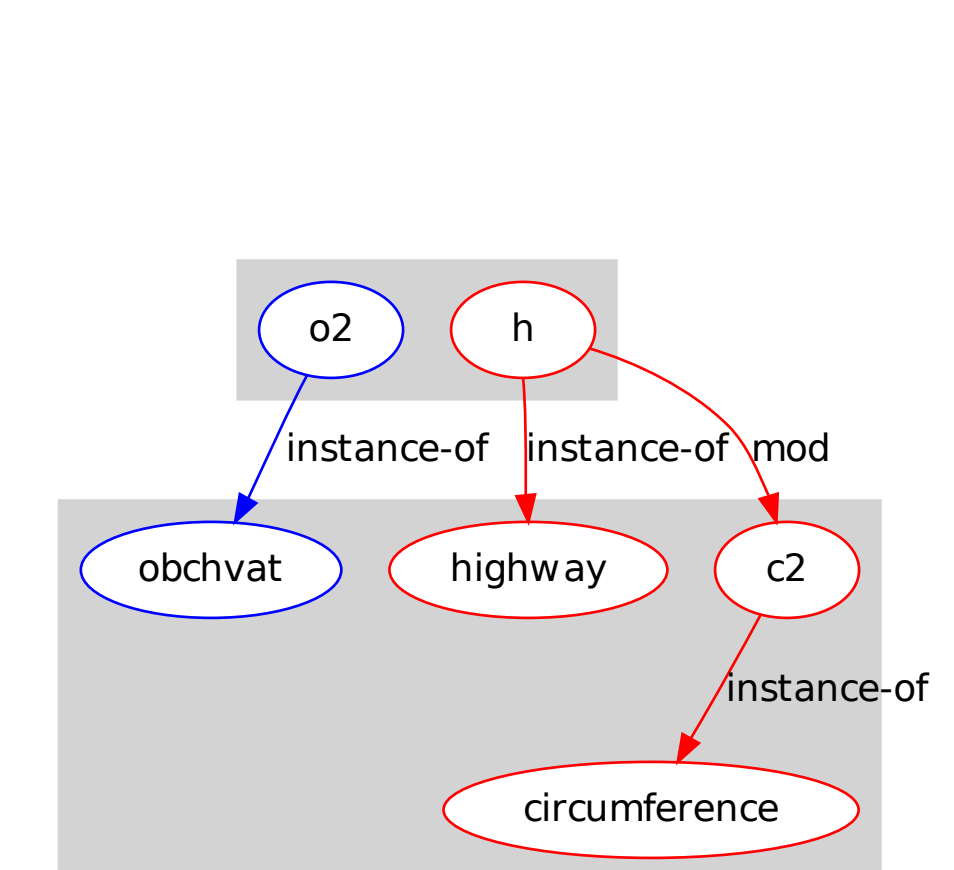


Difference in annotation choice

"překračovat povolenou rychlost"  
"speeding"

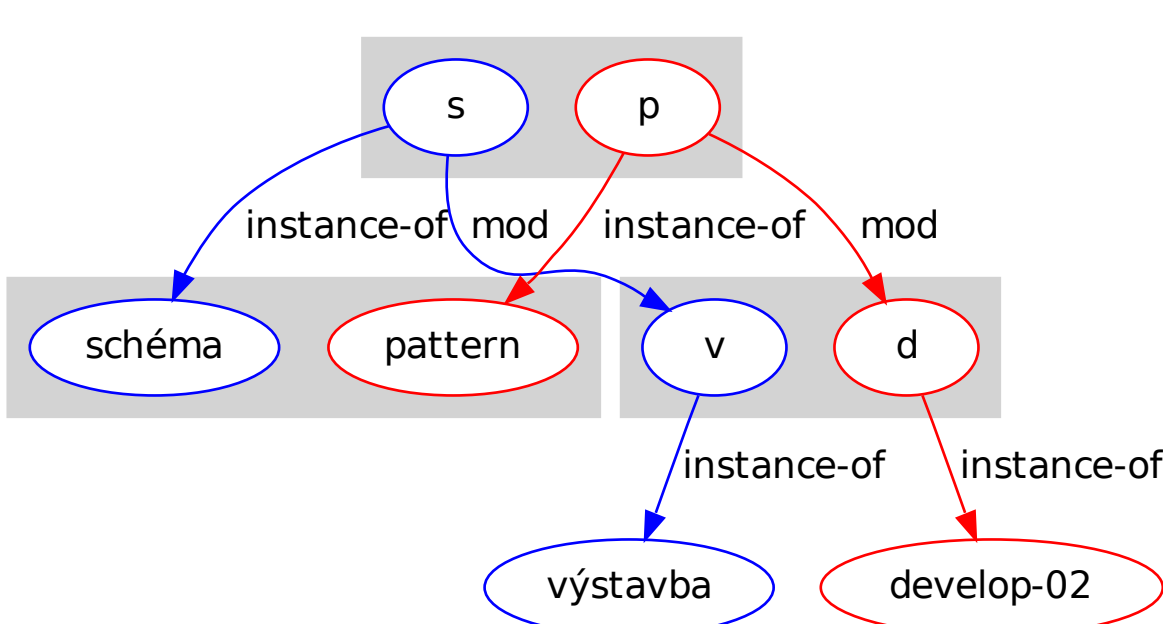


"obchvat"  
"Circumferential highway"



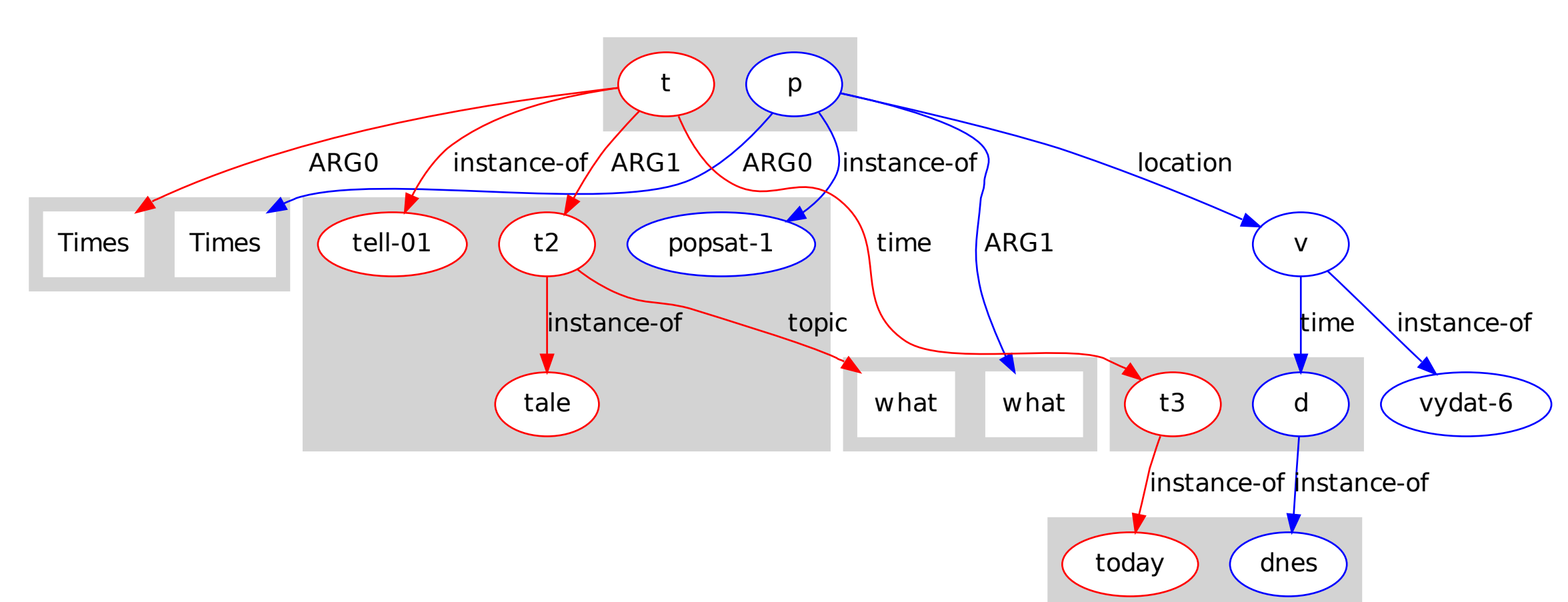
MWE vs. single word

"schémata výstavby",  
"Development patterns"



Event object choice and ontology divergence

„v dnešním vydání [Times] jsou popsány...“  
„[Times] tells the tale today...“



MWE and structural divergence

Preliminary comparison shows structural compatibility between Chinese and English AMRs, and between Czech and English AMRs. It also reveals some cross-lingual divergences in lexicalization of concepts, and in how sentences are structured.