**pos tag NNSSätze/Phänomene**

* Synonme
* Gleiche Semantik, keine Synonyme
* Unterschiedliche Anzahl Wörter
* Negation
* Idiome

**Scores**

S2match besser, aber oft auch nicht sher gut. Gründe: schlechte AMRs, MWEs. Gibt A young cat walks fast vs a kitten runs höheren Score als a young cat sprints vs a kitten runs

**AMRs**

((schwierig manuell zu beurteilen))

* MWEs: key\JJ appeals\NNS court\NN 🡪 egal, weil wird 1 ARG

Oder frernch\JJ (country) fries\NNS

* Aber andere falsche Tags resultieren in schlechten AMRs:
* Sprints\NNS 🡪 (cruise-01)/
* Schlagzeilen: Groß/Kleinschreibung ist relevant 🡪 House\(government-organization) vs house\(location)
* ((it got out of hand (hinder-01)) vs it got out of control (get, ARG1 control, ARG2 out)
* - Idiome werden nicht als solche erkannt, sondern in der wörtlichen Bedeutung getaggt. 🡪 hang\(hang-01 = suspend) in there.
* Negationen werden gut durch :polarity erfasst. The competition was not fair == the competition was unfair

Aber: do not need to worry: need (need-01) vs dont have to worry: have(obligate-01) ((unschönheit in propbank, die von s2match erfasst wird’s)

* Weitere Phänomene: less time 🡪 (quant-91) vs. Quicker (have-degree-91)

They decided/made the decision to protect and polish what remained: polish(polish vs meddle-01)

**Smatch-Alignments**

**S2match-Alignments**

Besser als Smatch-Alignments. Zb application- apply statt way-apply , more-realistic statt more-null, break up – split statt break up - null

Probleme: MWEs: french fries🡪 mod-konkatenation, Idiome

Walö-run, cat – kitten, young – null, fast - null

Kick-die, bucket- null, he-he

Findet manchmal dadurch schlechte alignments: The methodology.. methodology

**BERT Matrix**

Be located – have (sim = 0.0) zwar aligned, doch die token-embeddings zu berücksichtigen würde sich lohnen (sim = 0.71)

nicht-lemmatisieren lohnt sich zb chips-fries.

**Lösungsvorschläge für unterschiedliche Phänomeme**

**(Mail Juri, Update Denis)**