

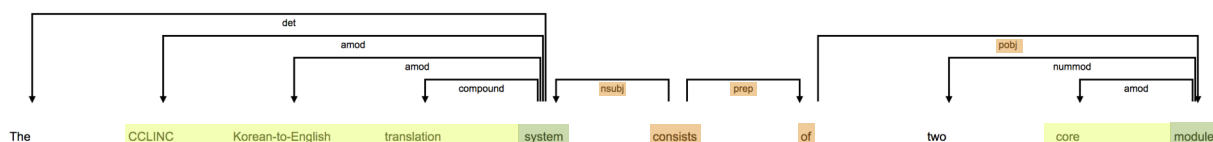
Problem Statement

- **Task:** Extract semantic relations from scientific literature
- **Relations:** Model-Feature, Part-Whole, Compare, Result, and Topic

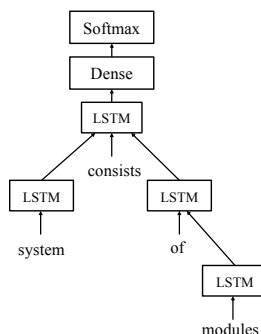
The CCLINC Korean-to-English translation system consists of two core modules. ← Part-Whole

- **Data:** Manually-annotated abstracts from the ACL Anthology Reference Corpus
- **Results:** Subtask 1.1: 9th (of 28), Subtask 1.2: 5th (of 20)
- Code available at: <https://github.com/Georgetown-IR-Lab/semEval2018-task7>

Method

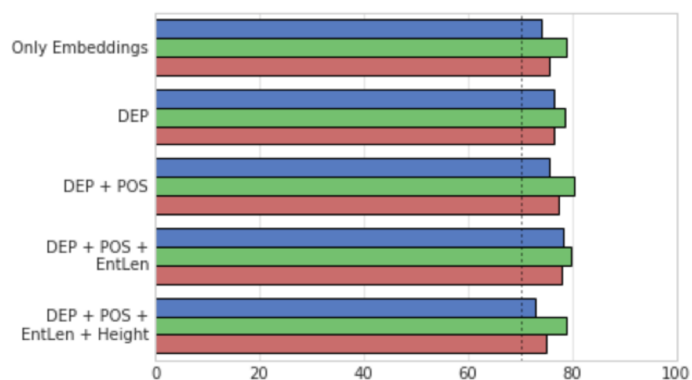


- **Approach:** Feed dependency paths between entities into Tree-LSTM



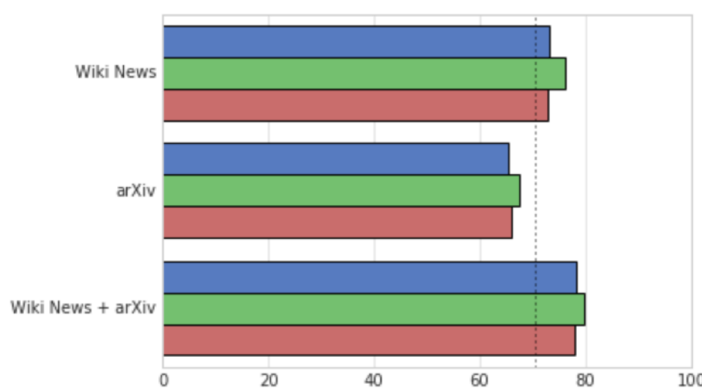
- Parsing: Stanford dependencies (spaCy)
- Child-sum variant tree LSTM
- Only syntactic root of entity
- Features:
 - Wiki + arXiv embeddings
 - Dependency labels
 - Part-of-speech tags
 - Entity length (where applicable)
 - Height in tree

Results & Conclusions



..... Median F1
█ P
█ R
█ F1

- Height feature causes model to overfit to training data



- Combination of general-language embeddings and domain-specific helps

True label \ Predicted label	U	M-F	P-W	C	R	T
U	143	15	9	6	1	1
M-F	13	38	10	0	3	1
P-W	15	8	44	3	0	0
C	1	1	2	14	3	0
R	4	2	0	1	13	0
T	2	0	0	0	0	1

- Often confuses Usage Model-Feature, and Part-Whole
- Difficulty identifying Topic