Sister Help: Data Augmentation for Frame-Semantic Role Labeling

Ayush Pancholy, Miriam R. L. Petruck, Swabha Swayamdipta

Background

FrameNet: database of semantic frames defined in terms of frame elements (semantic roles) that documents semantic-syntactic combinatorial possibilities of each Lexical Unit (LU) - pairing of a lemma and a frame

Problem: Expert annotation is resource intensive, resulting in relative paucity of data for NLP tasks (SRL, paraphrase, etc.)

Solution: Use frame-specific augmentation methods that leverage frame-specific parallel LU annotation

Commerce_buy

Method

Sister LU: in the same frame as an Empty LU, same POS, and non-empty annotation set

Copy annotation of Sister LU into that of Empty LU; replace each instance of Sister LU with correct word form of Empty LU

Use existing test sets to create Empty LUs artificially and repopulate their annotation sets

Results

Dev		Test			
ArgID	FrameID	ArgID			FrameID
F1	F1	P	R	F1	F1
59.3	84.7	61.8	56.9	59.2	81.1
61.9	**85.3**	**63.6**	**60.3**	**61.9**	**82.7**

Experimental outperforms baseline in every metric

Error analysis showed three categories of errors:

(a) *The moon was now *occlude* by clouds.

(b) *And he *complained* the endless squeeze on cash.

(c) ?The faint *flash* from a street light showed him the outline of a hedge.

Conclusions

Frame-specific augmentation *improves SRL performance.*
Future work: test on other SRL systems; modify augmentation approach; and pipeline back into FrameNet with human-in-the-loop.