

Lexical Semantic Recognition

Nelson Liu, Daniel Hershcovich, Michael Kranzlein, Nathan Schneider

MWE 2021 — August 6th, 2021

Joint Lexical Semantic Segmentation and Disambiguation

- In lexical semantics, full-sentence segmentation and segment labeling are often treated separately.
 - e.g., multiword expression identification (segmentation) and classification / supersense tagging (segment labeling)
- **This work:** joint segmentation and disambiguation, unifying previously disparate layers of annotation.
 - By analogy to named entity recognition (NER), these are lexical semantic recognition (LSR) tasks.

This Talk

- An overview of the STREUSLE LSR task
- Instantiation as sequence tagging
- Model description: BERT-based neural CRF tagger
- Evaluating the model on STREUSLE LSR task
- Do STREUSLE-trained taggers generalize to PARSEME and DiMSUM?

The STREUSLE LSR Task

- We focus on the LSR task defined by the STREUSLE corpus
- STREUSLE contains comprehensive annotations of:
 - MWEs (+ syntactic status)
 - Noun, Verb, and Preposition / Possessive Supersense.
- Annotated on Reviews section of English Web Treebank (55K tokens)

STREUSLE Example

STREUSLE	SS												
	LexCat	PRON	v.Motion	p.Possessor	n.Artifact	p.Purpose	n.Act	p.Theme				n.Artifact	
	MWE	We	V.VPC.full	PRON.POSS	N	P	DET	N	P	DET	N	N	
		We	took	our	vehicle	in	for	a	repair	to	the	air	conditioning

STREUSLE Example

STREUSLE	SS												
	LexCat	PRON	v.Motion V.VPC.full	p.Possessor PRON.POSS	n.Artifact N	p.Purpose P	n.Act DET	n.Act N	p.Theme P	DET		n.Artifact N	
	MWE	We	took	our	vehicle	in	for	a	repair	to	the	air	conditioning

STREUSLE Example

STREUSLE	SS		v.Motion	p.Possessor	n.Artifact		p.Purpose	n.Act	p.Theme		n.Artifact
	LexCat	PRON	V.VPC.full	PRON.POSS	N		P DET	N	P DET		N
	MWE	We	took	our	vehicle	in	for a	repair	to the	air	conditioning

STREUSLE Example

STREUSLE	SS		v.Motion	p.Possessor	n.Artifact		p.Purpose	n.Act	p.Theme		n.Artifact		
	LexCat	PRON	V.VPC.full	PRON.POSS	N		P	DET	N		P	DET	N
	MWE	We	took	our	vehicle	in	for	a	repair	to	the	air	conditioning

Related Frameworks

- STREUSLE annotations generalize those of two previous shared tasks:
 - English PARSEME 1.1 VMWEs (Ramisch et al., 2018): STREUSLE VMWE lexcats are identical to PARSEME.
 - DiMSUM (Schneider et al., 2016): MWEs, noun and verb supersenses.

Modeling Approach: Token-Level Sequence Tagging

- Serialize MWE, lexcat, and supersense information into token-level tags.
 - Format: {MWE BIO tag}-{Lexcat}-{Supersense Label}
- Joint sequence segmentation and segment-labeling is thus just token-level tagging.
- 572 unique tags in training, 12 tags in dev. that aren't seen during training

Tag Serialization Example

STREUSLE	SS											
	LexCat	PRON	v.Motion	p.Possessor	n.Artifact	p.Purpose	n.Act	p.Theme		n.Artifact		
	MWE		V.VPC.full	PRON.POSS	N	P	DET	N	P	DET	N	
		We	took	our	vehicle	in	for	a	repair	to	the	air conditioning

We/**O** - PRON took/**B** - V.VPC.full - v.Motion
 our/**O** - PRON.POSS vehicle/**O** - N - n.ARTIFACT in/**I**₋
 for/**O** - P - p.Purpose a/**O** - DET repair/**O** - N - n.ACT
 to/**O** - P - p.Theme the/**O** - DET air/**B** - N - n.ARTIFACT
 conditioning/**I**₋

Sequence Tagger

- BERT-based neural CRF tagger.
 - Embed tokens with BERT / other model
 - Feed contextualized token embeddings through BiLSTM
 - Predicting output: affine transformation + linear-chain CRF

Sequence Tagger Output Constraints

- Constraints on CRF transitions used to enforce valid MWE BIO sequences.
- Lexcats are also constrained by a UPOS and lemma.
 - Tokens with AUX UPOS must take AUX lexcat
- Use off-the-shelf UPOS tagger and lemmatizer to predict UPOS and lemmas for use in constraints

STREUSLE Performance

- Experiment: train on STREUSLE, test on STREUSLE

STREUSLE Full-Tag Accuracy

BERT (Gold UPOS)	82.5
BERT (Predicted UPOS)	81.0
GloVe (Gold UPOS)	79.3
GloVe (Predicted UPOS)	77.5

STREUSLE Performance

(only preposition / possessive supersenses)

	Preposition / Possessive Supersense F1
• Experiment: train on STREUSLE, test on STREUSLE	BERT (Gold UPOS) 71.4
• Schneider et al. (2018): previous best STREUSLE preposition / possessive tagger	BERT (Predicted UPOS) 71.6
	Schneider et al., (2018) 55.7

Can the model generalize to other corpora?

- Experiment: train on STREUSLE, evaluate on PARSEME

	MWE-based F1	Token-based F1
BERT (Gold UPOS)	40.3	45.4
BERT (Predicted UPOS)	39.2	43.4
Nerima et al. (2017) *Rule-based shared-task winner	33.3	34.4
Taslimipoor et al. (2019) *Uses ELMo / dependency parses	36.0	40.2
Rohanian et al. (2019) *Uses ELMo / dependency parses	41.9	-

Can the model generalize to other corpora?

- Experiment: train on STREUSLE, evaluate on DiMSUM

	MWEs F1	Supersenses F1	Combined F1
BERT (Gold UPOS)	50.0	54.2	53.4
BERT (Predicted UPOS)	49.7	51.4	51.1
Kirilin et al., (2016) DiMSUM shared task winner, uses Twitter + STREUSLE training data	58.4	58.0	58.1

Summary

- LSR on STREUSLE: joint MWE segmentation and lexical category / supersense disambiguation
- This task generalizes previous annotations of previous shared tasks (PARSEME VMWEs and DiMSUM)
- BERT-based tagger trained on STREUSLE is competitive with prior work on PARSEME and DiMSUM, despite only training on STREUSLE.
- Thanks! Questions?