

# Zero-shot Word Sense Disambiguation using Sense Definition Embeddings

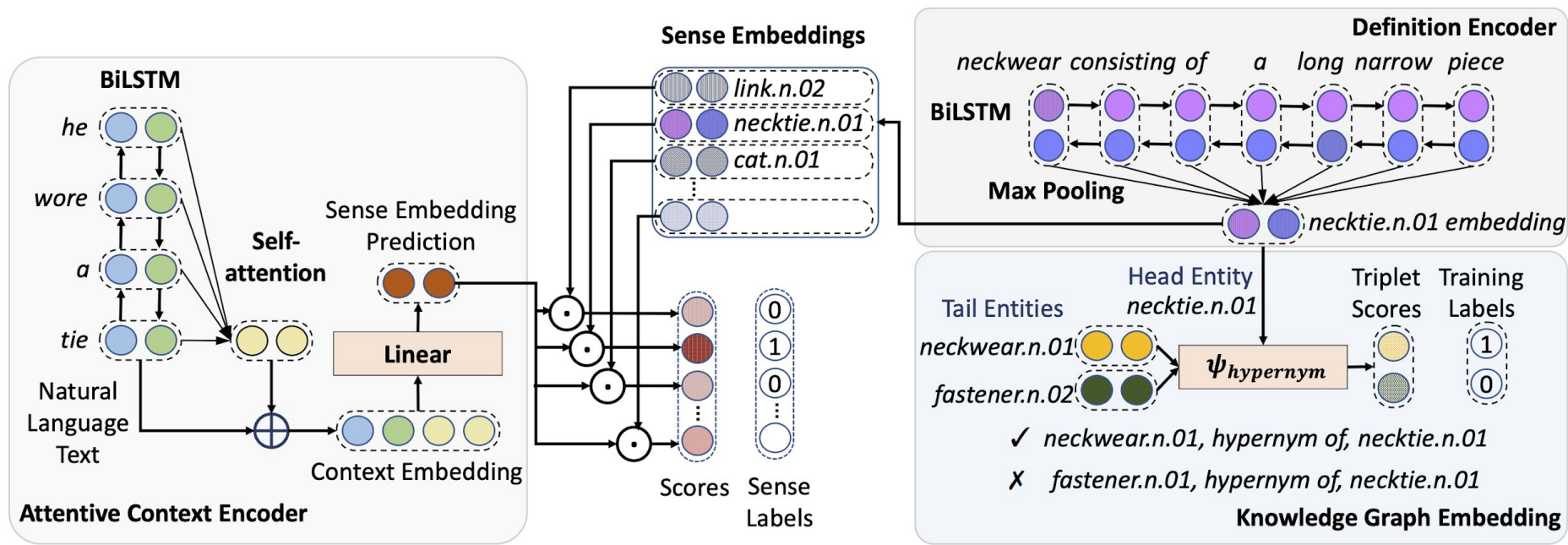
Sawan Kumar, Sharmistha Jat, Karan Saxena, Partha Talukdar

ACL 2019

# Motivation

- Supervised Learning treat sense as **discrete labels**.
- Supervised training predicts **most frequent sense** for unseen senses.
  
- Proposed to treat sense as continuous embedding
- Learn embedding from Knowledge Graph (e.g. WordNet)

# Framework



# Learning Sense Embedding

A Knowledge graph  $K=\{(h,l,t)\}$

TransE

$$d_{h,l,t} = -\text{cosine}(q(h) + e_l, q(t)) \quad (12) \quad L_T = \sum_{(h,l,t) \in K} \sum_{(h',l,t') \in K'} [\gamma + d_{h,l,t} - d_{h',l,t'}]_+, \quad (2)$$

ConvE

$$\psi_l(e_h, e_t) = f(\text{vec}(f([\overline{q(h)}; \overline{e_l}] * w))W) e_t \quad (13)$$

$$p = \sigma(\psi_l(e_h, e_t)). \quad (4)$$

$$L_C = -\frac{1}{N} \sum_i (t_i \cdot \log(p_i) + (1 - t_i) \cdot \log(1 - p_i)), \quad (5)$$

# Training

$$\hat{p}_j^i = \text{softmax}(\text{dot}(v^i, \rho_j) + \text{dot}(b, \rho_j)); \quad (9)$$
$$\rho_j \in S$$

# Results

	Dev	Test Datasets					Concatenation of All Test Datasets				
	SE7	SE2	SE3	SE13	SE15	Nouns	Verbs	Adj.	Adv.	ALL	
WordNet S1	55.2	66.8	66.2	63.0	67.8	67.6	50.3	74.3	80.9	65.2	
Non-neural baselines											
MFS (Using training data)	54.5	65.6	66.0	63.8	67.1	67.7	49.8	73.1	80.5	65.5	
IMS+emb (2016)^	62.6	72.2	70.4	<u>65.9</u>	71.5	71.9	<u>56.6</u>	75.9	84.7	<u>70.1</u>	
Lesk <sub>ext</sub> +emb (2014)*	<u>56.7</u>	63.0	63.7	66.2	64.6	70.0	51.1	51.7	80.6	64.2	
UKB <sub>gloss</sub> +w2w (2014)*	42.9	63.5	55.4	62.9	63.3	64.9	41.4	69.5	<u>69.7</u>	61.1	
Babelfy (2014)	51.6	67.0	63.5	66.4	<u>70.3</u>	68.9	50.7	<u>73.2</u>	79.8	66.4	
Context2Vec (2016) ^	61.3	71.8	69.1	65.6	71.9	71.2	57.4	75.2	82.7	69.6	
WSD-TM (2018)	55.6	<u>69.0</u>	<u>66.9</u>	65.3	69.6	69.7	51.2	76.0	80.9	66.9	
Neural baselines											
BiLSTM+att+LEX (2017b)	63.7	72.0	69.4	66.4	70.8	71.6	57.1	75.6	83.2	69.7	
BiLSTM+att+LEX+POS (2017b)	64.8	72.0	69.1	66.9	71.5	71.5	57.5	75.0	83.8	69.9	
GAS <sub>ext</sub> (Linear) (2018b)*	–	72.4	70.1	67.1	72.1	<u>71.9</u>	58.1	76.4	84.7	70.4	
GAS <sub>ext</sub> (Concatenation) (2018b)*	–	72.2	70.5	67.2	72.6	<u>72.2</u>	57.7	76.6	<b>85.0</b>	70.6	
CAN <sub>s</sub> (2018a)*	–	72.2	70.2	69.1	72.2	73.5	56.5	76.6	83.3	70.9	
HCAN (2018a)*	–	72.8	70.3	68.5	72.8	72.7	58.2	77.4	84.1	71.1	
<b>EWISÉ (ConvE)*</b>	<b>67.3</b>	<b>73.8</b>	<b>71.1</b>	<b>69.4</b>	<b>74.5</b>	<b>74.0</b>	<b>60.2</b>	<b>78.0</b>	82.1	<b>71.8</b>	

## Result

	MFS	LFS
WordNet S1	100.0	0.0
Lesk(ext)+emb	92.7	9.4
Babelfy	93.9	12.2
BiLSTM-A	93.4	22.9
EWISE	93.5	<b>31.2</b>