Workshop on Representation Learning for NLP
11 August, 2016, Berlin, Germany

TECHNISCHE UNIVERSITÄT DARMSTADT

Making Sense of Word Embeddings

Maria Pelevina[1], Nikolay Arefyev[2], Chris Biemann[1] and **Alexander Panchenko**[1]

[1]Technische Universität Darmstadt, LT Group, Computer Science Department, Germany
[2]Moscow State University, Faculty of Computational Mathematics and Cybernetics, Russia

Overview of the contribution

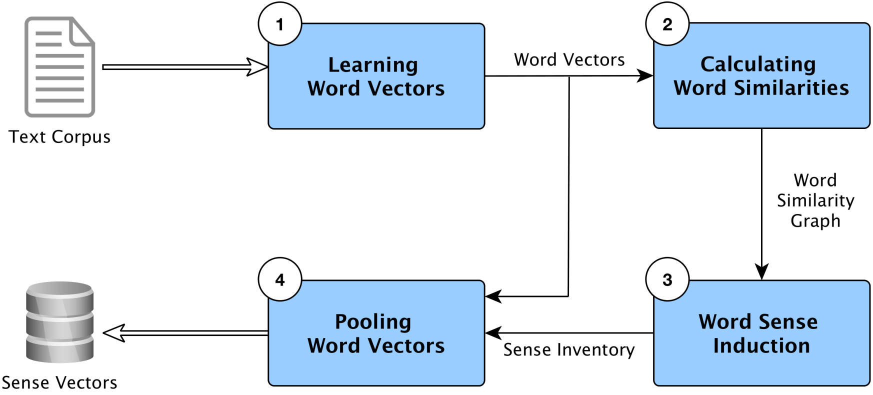
Prior methods:

- Induce inventory by clustering of word instances (Li and Jurafsky, 2015)
- Use existing inventories (Rothe and Schütze, 2015)

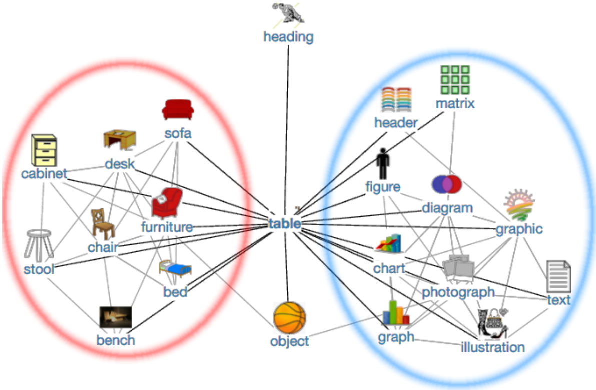
Our method:

- **Input:** word embeddings
- **Output:** word sense embeddings
- **Word sense induction** by clustering of word ego-networks
- **Word sense disambiguation** based on the induced sense representations

Learning Word Sense Embeddings



Word Sense Induction: Ego-Network Clustering



- The "furniture" and the "data" sense clusters of the word "table".
- Graph clustering using the **Chinese Whispers algorithm** (Biemann, 2006).

Neighbours of Word and Sense Vectors

TECHNISCHE UNIVERSITÄT DARMSTADT

Vector	Nearest Neighbours
table	tray, bottom, diagram, bucket, brackets, stack, basket, list, parenthesis, cup, trays, pile, playfield, bracket, pot, drop-down, cue, plate
table#0	leftmost#0, column#1, randomly#0, tableau#1, top-left0, indent#1, bracket#3, pointer#0, footer#1, cursor#1, diagram#0, grid#0
table#1	pile#1, stool#1, tray#0, basket#0, bowl#1, bucket#0, box#0, cage#0, saucer#3, mirror#1, birdcage#0, hole#0, pan#1, lid#0

- Neighbours of the word “table" and its senses produced by our method.
- The neighbours of the initial vector belong to **both senses**.
- The neighbours of the sense vectors are **sense-specific**.

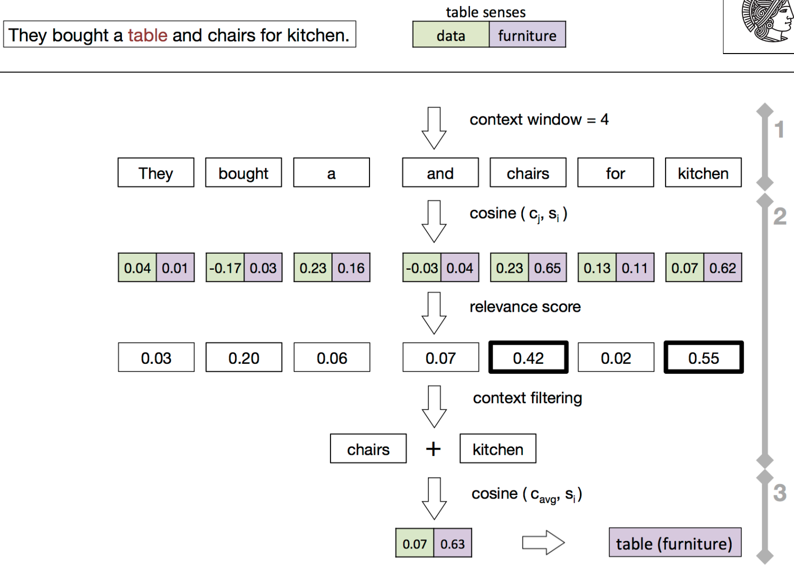
1. Context Extraction
   - use context words around the target word

2. Context Filtering
   - based on context word's relevance for disambiguation

3. Sense Choice
   - maximize similarity between context vector and sense vector

Word Sense Disambiguation: Example

TECHNISCHE UNIVERSITÄT DARMSTADT



Evaluation on SemEval 2013 Task 13 dataset: comparison to the state-of-the-art

Model	Jacc.	Tau	WNDCG	F.NMI	F.B-Cubed
AI-KU (add1000)	0.176	0.609	0.205	0.033	0.317
AI-KU	0.176	0.619	0.393	0.066	0.382
AI-KU (remove5-add1000)	0.228	0.654	0.330	0.040	0.463
Unimelb (5p)	0.198	0.623	0.374	0.056	0.475
Unimelb (50k)	0.198	0.633	0.384	0.060	0.494
UoS (#WN senses)	0.171	0.600	0.298	0.046	0.186
UoS (top-3)	0.220	0.637	0.370	0.044	0.451
La Sapienza (1)	0.131	0.544	0.332	–	–
La Sapienza (2)	0.131	0.535	0.394	–	–
AdaGram, α = 0.05, 100 dim	0.274	0.644	0.318	0.058	0.470
w2v	0.197	0.615	0.291	0.011	0.615
w2v (nouns)	0.179	0.626	0.304	0.011	0.623
JBT	0.205	0.624	0.291	0.017	0.598
JBT (nouns)	0.198	0.643	0.310	0.031	0.595
TWSI (nouns)	0.215	0.651	0.318	0.030	0.573

Conclusion

TECHNISCHE UNIVERSITÄT DARMSTADT

- Novel approach for learning **word sense embeddings**.
- Can use **existing word embeddings** as input.
- WSD performance **comparable to the state-of-the-art** systems.
- **Source code and pre-trained models:**
  `https://github.com/tudarmstadt-lt/SenseGram`

Thank you and welcome to our poster!

Evaluation based on the TWSI dataset: a large-scale dataset for development