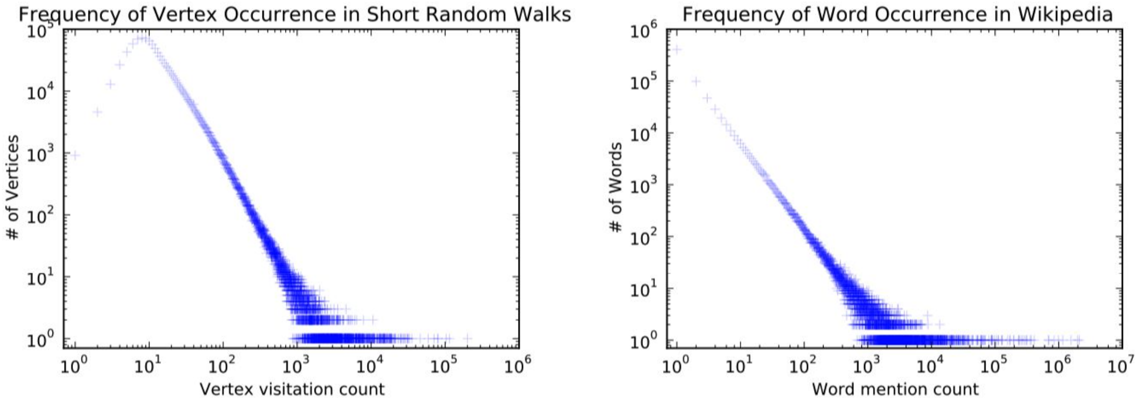
Dolores: Deep Contextualized Knowledge Graph Embeddings

Haoyu Wang, Vivek Kulkarni, William Yang Wang

Motivation: Language Modeling for Entity-Relation Chains

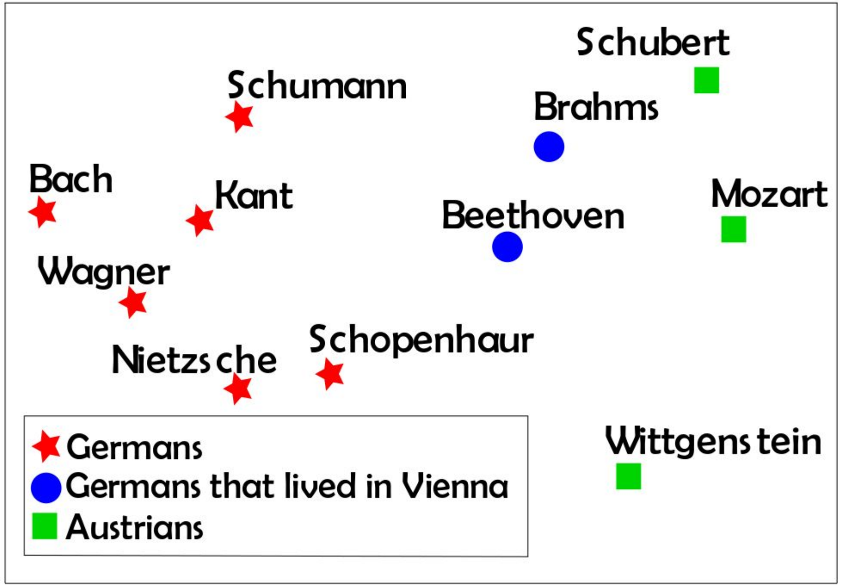
- Similar Power-Law distribution between vertices in graphs and words in natural language [Perozzi et al., 2014]



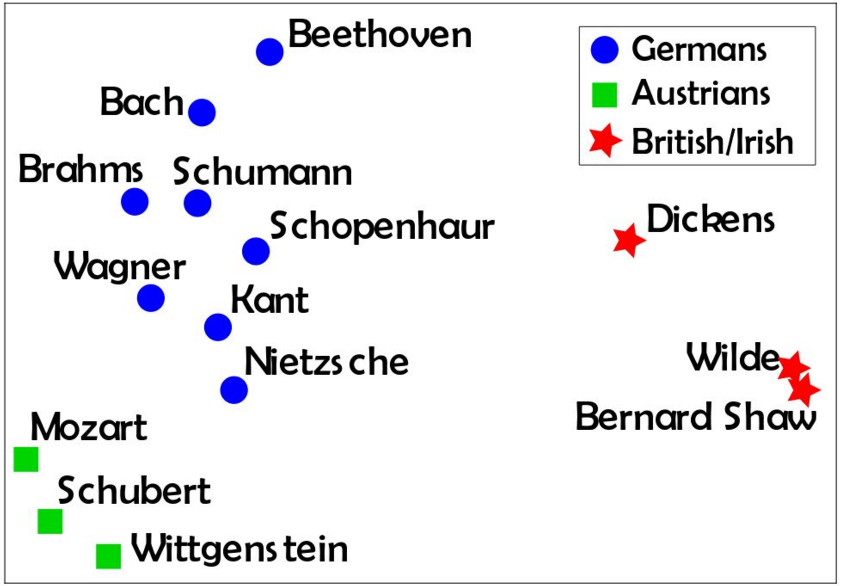
- Deep Contextualized Word Embeddings from language models [Peters et al., 2018]

$$\Pr([e_1, r_1], [e_2, r_2], \cdots, [e_N, r_N]) = \prod_{t=1}^{N} \Pr([e_t, r_t] \mid [e_1, r_1], [e_2, r_2], \cdots, [e_{t-1}, r_{t-1}]).$$

Visualization: Contextualized Knowledge Graph Embeddings



Relation: ‘people/place_lived’



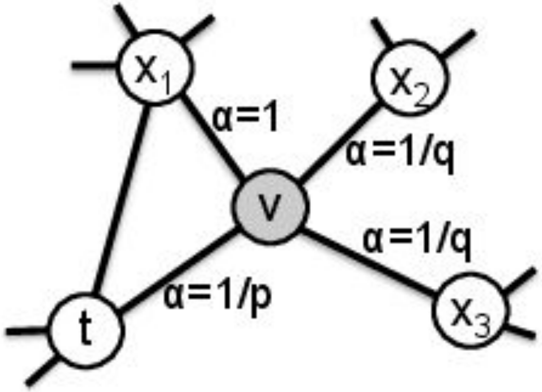
Relation: ‘people/nationality’

Dolores Component 1: Path Generator

- Random Walk on the Graph
  - p: the likelihood of immediately revisiting a node
  - q: the likelihood that a walk is biased towards nodes close to starting node
  - 20 chains for each node

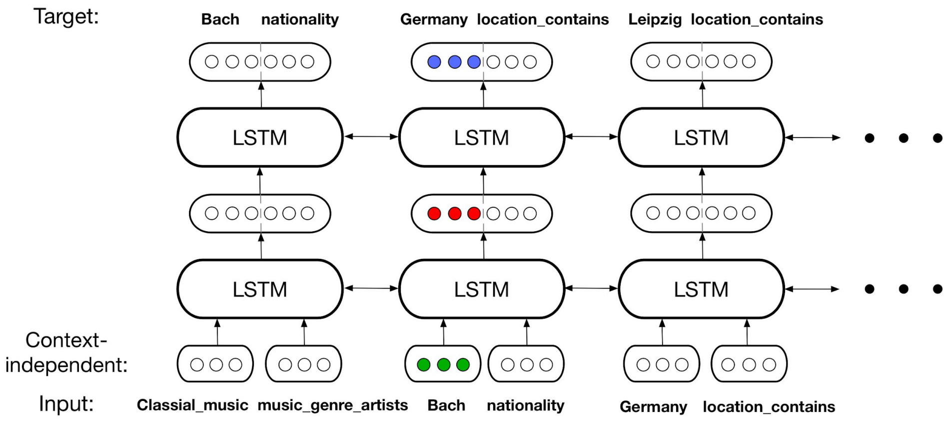
    $$(e_1, r_1, e_2, r_2, \cdots, e_k)$$

  - length of each chain: 21 (10 entities and 11 relations alternatively, k = 10)



Dolores Component 2: Embedding Learner

- Network Architecture



$$\textsc{Dolores}_t = [x_t, \sum_{i=1}^{L} \lambda_i \cdot h_{t,i}]$$

$h_{t,i} = [\ \overrightarrow{h_{t,i}},\ \overleftarrow{h_{t,i}}\ ]$ corresponds to the context-dependent embeddings from layer i

λ_i's denote task-specific learnable weights of the linear combination

Extracting Dolores Embeddings & Evaluation Results

- Training the Dolores Learner using chains from Path Generator
- Accepting task corpus as the input to Dolores Learner, and generate contextualized embedding for each entity/relation in the task corpus
- Utilize the embedding as the input of embedding layers of task-specific models
- Results on three KBC tasks:
  - Link Prediction: FB15K237
  - Triple Classification: WN11, FB13
  - Multi-hop KB Completion: dataset released by Neelakantan et al. [2015]

TASK	BASELINE		DOLORES+ BASELINE	INCREASE (ABSOLUTE/ RELATIVE)
Link Prediction (head)	[Nguyen et al., 2018b]	15.7	18.7	3.0 / 3.56%
Link Prediction (tail)	[Nguyen et al., 2018b]	32.8	37.2	4.4 / 6.55%
Triple Classification	[Nguyen et al., 2018b]	87.00	87.55	0.55 / 4.23%
Multi-hop KB Completion	[Yin et al., 2018]	76.16	78.28	2.12 / 8.9%

Conclusion

- We present a new method of learning deep contextualized knowledge graph embeddings using a deep neural sequential model.
- These embeddings are functions of hidden states of the deep neural model and can capture both context-independent and context-dependent cues.
- We show empirically that Dolores can easily be incorporated into existing predictive models on knowledge graphs to advance performances on several tasks like link prediction, triple classification, and multi-hop knowledge base completion.