Hypernetwork Knowledge Graph Embeddings

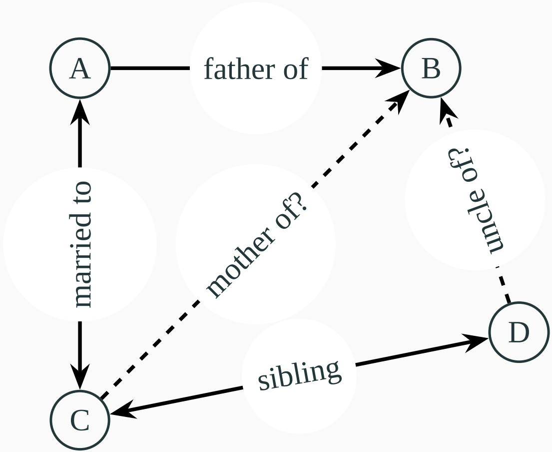
Ivana Balažević, Carl Allen and Timothy Hospedales

September 19, 2019

CDT in Data Science, School of Informatics, University of Edinburgh

28th International Conference on Artificial Neural Networks (ICANN), 2019

Task: Link Prediction on Knowledge Graphs



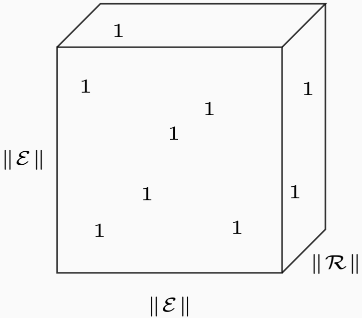
Entities $\mathcal{E} = \{A, B, C, D\}$

Relations $\mathcal{R} = \{$*married to, father of, uncle of, ...*$\}$

Knowledge Graph $\mathcal{G} = \{(A, \textit{father of}, B), (A, \textit{married to}, C), ...\}$

Background: Binary Tensor Representation

An alternative view of a knowledge graph: implicit sparse **third-order binary tensor representation** of known facts.



Background: Score Function

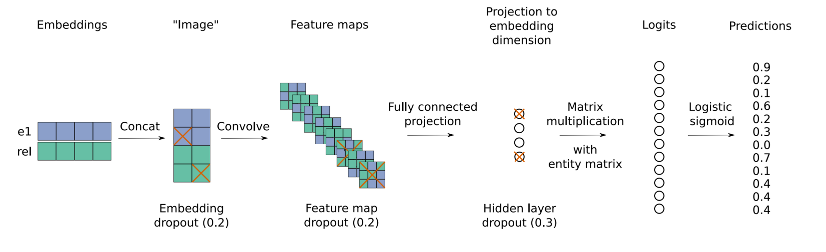
Typically in link prediction, a **score function** $\phi : \mathcal{E} \times \mathcal{R} \times \mathcal{E} \to \mathbb{R}$ is learned, that assigns a score $s = \phi(e_1, r, e_2)$ to each triple (e_1, r, e_2).

Type	$\phi(e_1, r, e_2)$	Models
Bilinear	$\mathbf{e}_1^\top \mathbf{M}_r \mathbf{e}_2 = \langle \mathbf{e}_1^{(r)}, \mathbf{e}_2 \rangle$	RESCAL (Nickel et al., 2011) DistMult (Yang et al., 2015) ComplEx (Trouillon et al., 2016)
Translational	$\lVert \mathbf{e}_1 \mathbf{M}_{r1} + \mathbf{r} - \mathbf{e}_2 \mathbf{M}_{r2} \rVert = \lVert \mathbf{e}_1^{(r)} + \mathbf{r} - \mathbf{e}_2^{(r)} \rVert$	TransE (Bordes et al., 2013) STransE (Nguyen et al., 2016))
Nonlinear	$f_r(\mathbf{e}_1, \mathbf{e}_2)$; f_r is nonlinear	ConvE (Dettmers et al., 2018)

Table 1: Types of score functions.

Background: ConvE - Nonlinear Model for Link Prediction

ConvE (Dettmers et al., 2018) - 2D convolution on the reshaped and concatenated subject entity and relation embeddings.



Background: ConvE - Nonlinear Model for Link Prediction

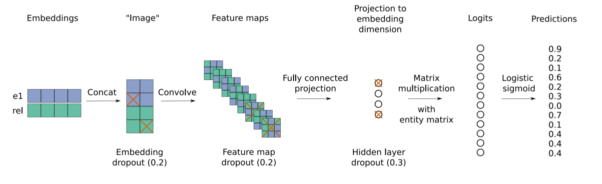
Oddities of ConvE:

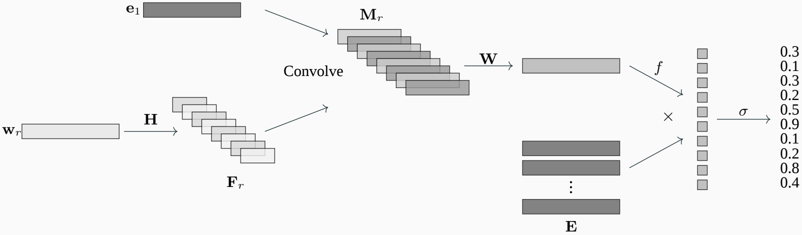
- reshaping and concatenation of embeddings; and
- using 2D convolution within word embeddings.

However, it achieves **state-of-the-art** results across standard link prediction datasets. How?

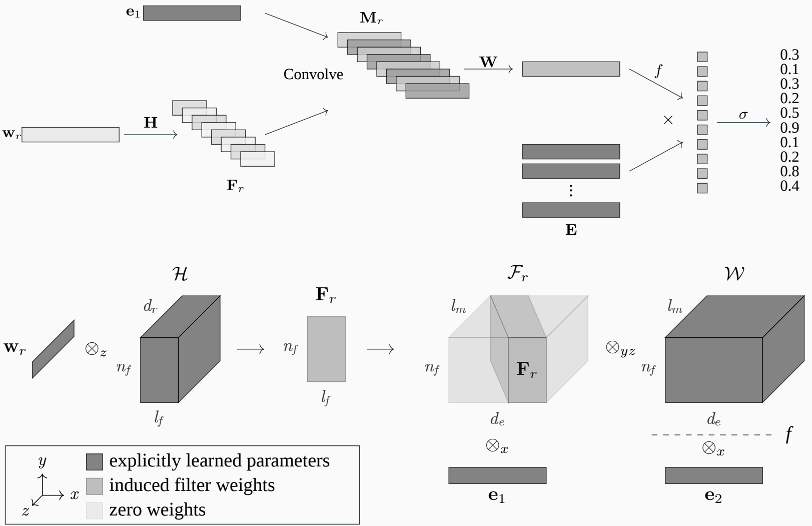
Maybe we can try to simplify it...

Hypernetwork Knowledge Graph Embeddings (HypER)





Hypernetwork Knowledge Graph Embeddings (HypER)



Hypernetwork Knowledge Graph Embeddings (HypER)

Main contributions of the paper:

- removes the concatenation and reshaping of embeddings;
- replaces 2D with 1D convolution;
- explains HypER in terms of **tensor factorization**, thus placing it within a well established family of bilinear models; and
- justifies using convolution as a convenient computational means of introducing **sparsity** and **parameter tying** (explicit regularization).

Results

	WN18RR					FB15k-237				
	MR	MRR	H@10	H@3	H@1	MR	MRR	H@10	H@3	H@1
DistMult (Yang et al., 2015)	5110	.430	.490	.440	.390	254	.241	.419	.263	.155
ComplEx (Trouillon et al., 2016)	5261	.440	.510	.460	.410	339	.247	.428	.275	.158
Neural LP (Yang et al., 2017)	–	–	–	–	–	–	.250	.408	–	–
R-GCN (Schlichtkrull et al., 2018)	–	–	–	–	–	–	.248	.417	.264	.151
MINERVA (Das et al., 2018)	–	–	–	–	–	–	–	.456	–	–
ConvE (Dettmers et al., 2018)	**4187**	.430	.520	.440	.400	**244**	.325	.501	.356	.237
HypER (ours)	5798	**.465**	**.522**	**.477**	**.436**	250	**.341**	**.520**	**.376**	**.252**

Table 2: Link prediction results on WN18RR and FB15k-237.

	WN18					FB15k				
	MR	MRR	H@10	H@3	H@1	MR	MRR	H@10	H@3	H@1
TransE (Bordes et al., 2013)	**251**	–	.892	–	–	125	–	.471	–	–
DistMult (Yang et al., 2015)	902	.822	.936	.914	.728	97	.654	.824	.733	.546
ComplEx (Trouillon et al., 2016)	–	.941	.947	.936	.936	–	.692	.840	.759	.599
ANALOGY (Liu et al., 2017)	–	.942	.947	.944	.939	–	.725	.854	.785	.646
Neural LP (Yang et al., 2017)	–	.940	.945	–	–	–	.760	.837	–	–
R-GCN (Schlichtkrull et al., 2018)	–	.819	**.964**	.929	.697	–	.696	.842	.760	.601
TorusE (Ebisu and Ichise, 2018)	–	.947	.954	.950	.943	–	.733	.832	.771	.674
ConvE (Dettmers et al., 2018)	374	.943	.956	.946	.935	51	.657	.831	.723	.558
HypER (ours)	431	**.951**	.958	**.955**	**.947**	**44**	**.790**	**.885**	**.829**	**.734**

Table 3: Link prediction results on WN18 and FB15k.

Conclusion

- HypER is **fast**, **robust to overfitting** and it consistently **outperforms all other models** across all datasets.
- **No benefit is gained from 2D convolutional filters** over 1D, dispelling the suggestion that 2D structure exists in entity embeddings, as implied by ConvE.
- Our results suggest that **convolution provides a good trade-off between expressiveness and number of parameters** compared to a dense network.

Code: `https://github.com/ibalazevic/HypER`

Thanks!

If you are interested:

- **TuckER: Tensor Factorization for Knowledge Graph Completion**
  Ivana Balažević, Carl Allen, Timothy Hospedales
  Empirical Methods in Natural Language Processing, 2019.
- **Multi-relational Poincaré Graph Embeddings**
  Ivana Balažević, Carl Allen, Timothy Hospedales
  Advances in Neural Information Processing Systems, 2019.
- **What the Vec? Towards Probabilistically Grounded Embeddings**
  Carl Allen, Ivana Balažević, Timothy Hospedales
  Advances in Neural Information Processing Systems, 2019.

Any questions?

References

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating Embeddings for Modeling Multi-relational Data. In *Advances in Neural Information Processing Systems*, 2013.

Rajarshi Das, Shehzaad Dhuliawala, Manzil Zaheer, Luke Vilnis, Ishan Durugkar, Akshay Krishnamurthy, Alex Smola, and Andrew McCallum. Go for a Walk and Arrive at the Answer: Reasoning over Paths in Knowledge Bases Using Reinforcement Learning. In *International Conference on Learning Representations*, 2018.

References ii

Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2D Knowledge Graph Embeddings. In *Association for the Advancement of Artificial Intelligence*, 2018.

Takuma Ebisu and Ryutaro Ichise. TorusE: Knowledge Graph Embedding on a Lie Group. In *Association for the Advancement of Artificial Intelligence*, 2018.

Hanxiao Liu, Yuexin Wu, and Yiming Yang. Analogical Inference for Multi-relational Embeddings. In *International Conference on Machine Learning*, 2017.

Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu, and Mark Johnson. STransE: a Novel Embedding Model of Entities and Relationships in Knowledge Bases. In *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016.

Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. A Three-Way Model for Collective Learning on Multi-Relational Data. In *International Conference on Machine Learning*, 2011.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. Modeling Relational Data with Graph Convolutional Networks. In *European Semantic Web Conference*, 2018.

Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. Complex Embeddings for Simple Link Prediction. In *International Conference on Machine Learning*, 2016.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In *International Conference on Learning Representations*, 2015.

Fan Yang, Zhilin Yang, and William W Cohen. Differentiable Learning of Logical Rules for Knowledge Base Reasoning. In *Advances in Neural Information Processing Systems*, 2017.