How to Turn your Knowledge Graph Embeddings into Generative Models

Paper

Code

Lorenzo Loconte
University of Edinburgh, UK

Nicola Di Mauro
University of Bari, Italy

Robert Peharz
TU Graz, Austria

Antonio Vergari
University of Edinburgh, UK

TL;DR *"We reinterpret KGE models into generative models of triples, making them to scale to large knowledge graphs, reliable with logical constraints and supporting sampling."*

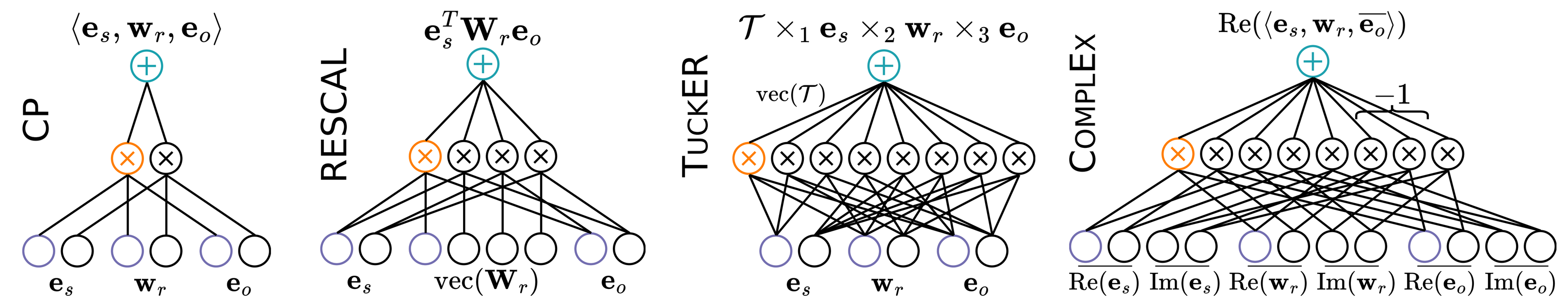
Knowledge graph embedding (KGE) models

Issues?

1 Scores are difficult to interpret, combine, compare [1].
How to measure the confidence of predictions?

2 Link predictions violate logical constraints.
How to guarantee the satisfaction of constraints?

3 KGE models are expensive to learn.
How to scale to KGs with millions of entities?

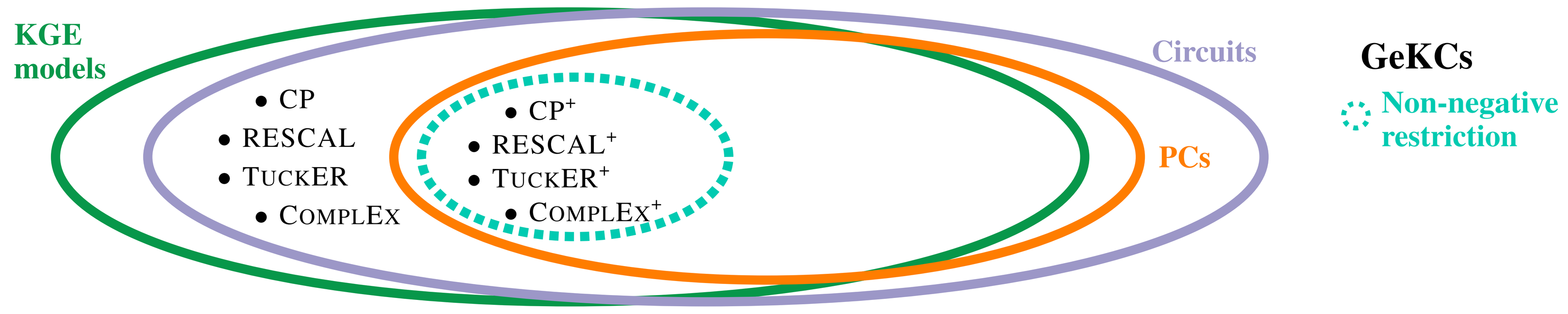


Interpreting the score functions of KGE models as constrained computational graphs: circuits [2]

1 From KGE models to probabilistic circuits (PCs)

We convert score functions (i.e., circuits) into ***probabilistic circuits*** (PCs) [2, 3] without additional memory requirements.

$$p(s,r,o) = \frac{1}{Z}\,\phi_{\mathsf{pc}}(s,r,o) \qquad \text{s.t.} \qquad \phi_{\mathsf{pc}}(s,r,o) \geq 0$$



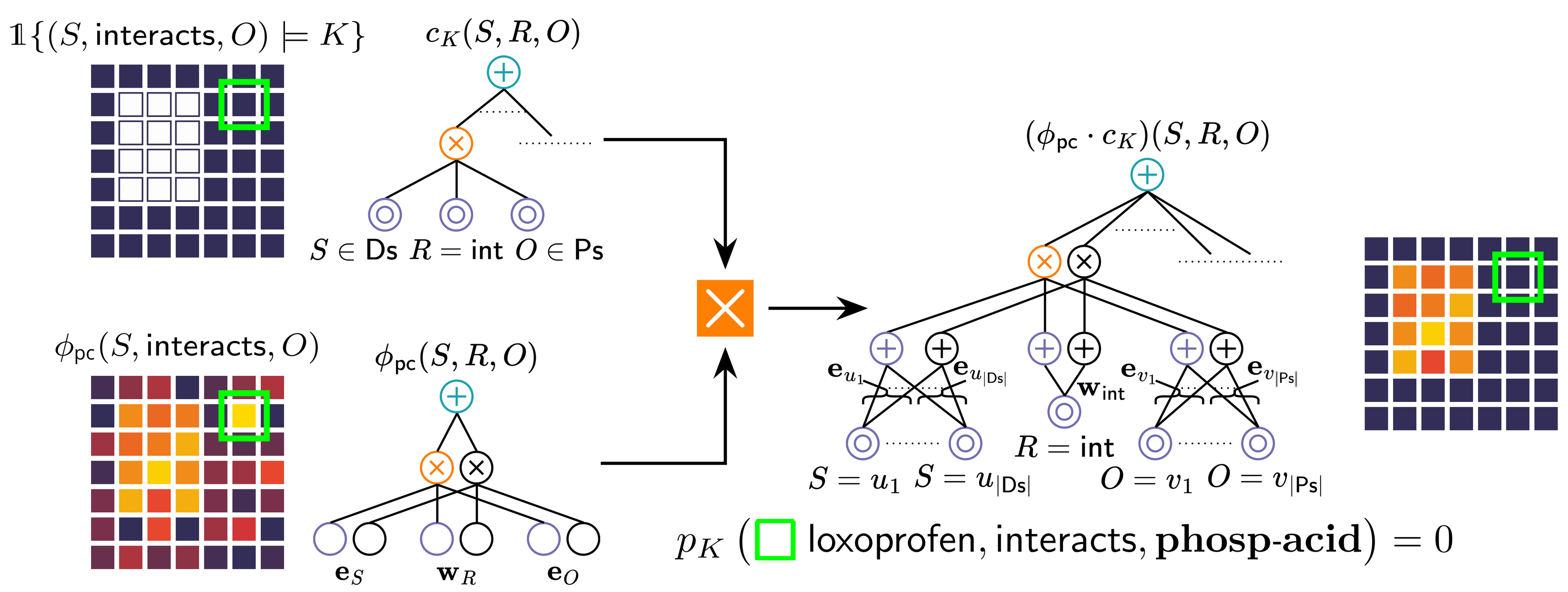
Generative KGE circuits (GeKCs) obtained via:

Non-negative restriction make the embeddings and computational unit activations non-negative;

or **Squaring the score function** square the circuit [4] without restricting the parameters domain, e.g.,

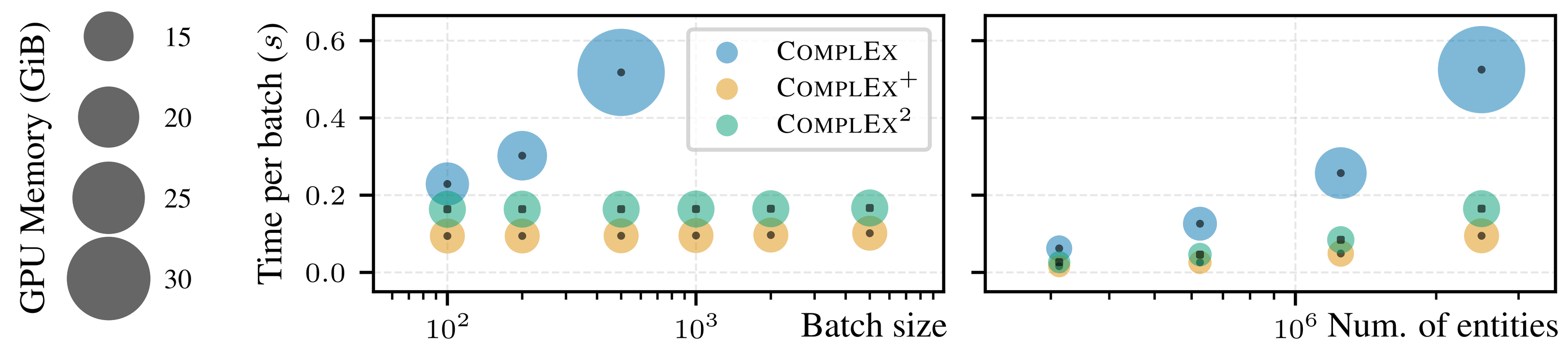
$$\phi_{\mathsf{CP}^2}(s,r,o) = \left(\sum_{i=1}^{d} \mathbf{e}_{si}\mathbf{w}_{ri}\mathbf{e}_{oi}\right)^2 = \sum_{i=1}^{d}\sum_{j=1}^{d} \mathbf{e}_{si}\mathbf{e}_{sj}\mathbf{w}_{ri}\mathbf{w}_{rj}\mathbf{e}_{oi}\mathbf{e}_{oj}$$

2 Integration of constraints with guarantees



[5]

3 Scaling to KGs with millions of entities



Link prediction benchmarks

Model	FB15k-237		WN18RR		ogbl-biokg	
	PLL	MLE	PLL	MLE	PLL	MLE
CP	0.310	—	**0.105**	—	0.831	—
CP+	0.237	0.230	0.027	0.026	0.496	0.501
CP²	**0.315**	0.282	**0.104**	0.091	**0.848**	0.829
ComplEx	**0.342**	—	**0.471**	—	0.829	—
ComplEx+	0.214	0.205	0.030	0.029	0.503	0.516
ComplEx²	0.334	0.300	0.420	0.391	**0.858**	0.840

Bonus *Sampling triples*

$$\mathrm{KTD}(\mathbb{P},\mathbb{Q}) = \|\, \mathbb{E}_{\mathbf{x}\sim\mathbb{P}}[\, \varphi(\psi(\mathbf{x}))\,] - \mathbb{E}_{\mathbf{z}\sim\mathbb{Q}}[\, \varphi(\psi(\mathbf{z}))\,]\,\|^2$$

References

[1] Erik Arakelyan, Pasquale Minervini, and Isabelle Augenstein. *Adapting Neural Link Predictors for Complex Query Answering*. 2023. arXiv: 2301.12313 [cs.LG].

[2] Antonio Vergari, Nicola Di Mauro, and Guy Van den Broeck. "Tractable probabilistic models: Representations, algorithms, learning, and applications". In: *Tutorial at the 35th Conference on Uncertainty in Artificial Intelligence (UAI)* (2019).

[3] YooJung Choi, Antonio Vergari, and Guy Van den Broeck. "Probabilistic Circuits: A Unifying Framework for Tractable Probabilistic Modeling". In: (2020).

[4] Antonio Vergari et al. "A Compositional Atlas of Tractable Circuit Operations for Probabilistic Inference". In: *Advances in Neural Information Processing Systems 34 (NeurIPS)*. Curran Associates, Inc., 2021, pp. 13189–13201.

[5] Kareem Ahmed et al. "Semantic probabilistic layers for neuro-symbolic learning". In: *Advances in Neural Information Processing Systems 35 (NeurIPS)*. Vol. 35. Curran Associates, Inc., 2022, pp. 29944–29959.