How to Turn your Knowledge Graph Embeddings

into Generative Models

Lorenzo Loconte University of Edinburgh, UK

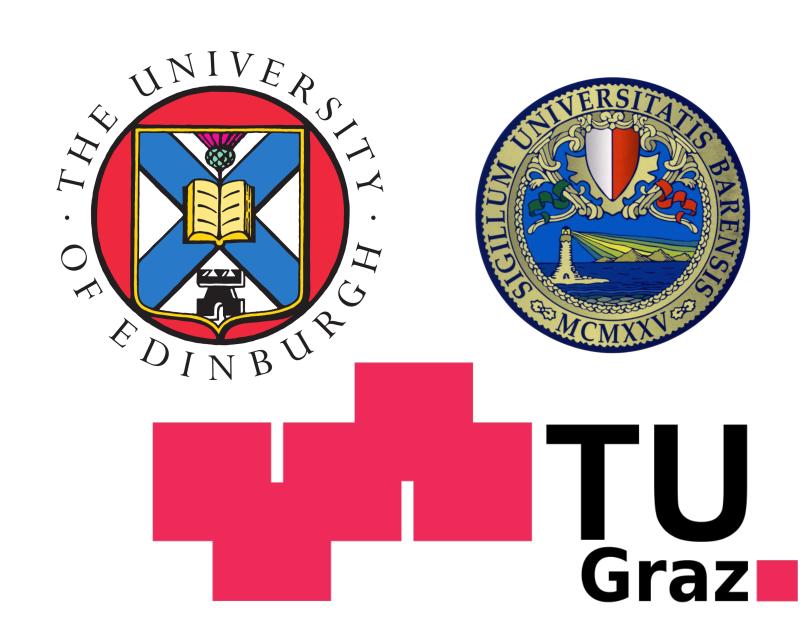
Nicola Di Mauro University of Bari, Italy **Robert Peharz** TU Graz, Austria

Antonio Vergari University of Edinburgh, UK

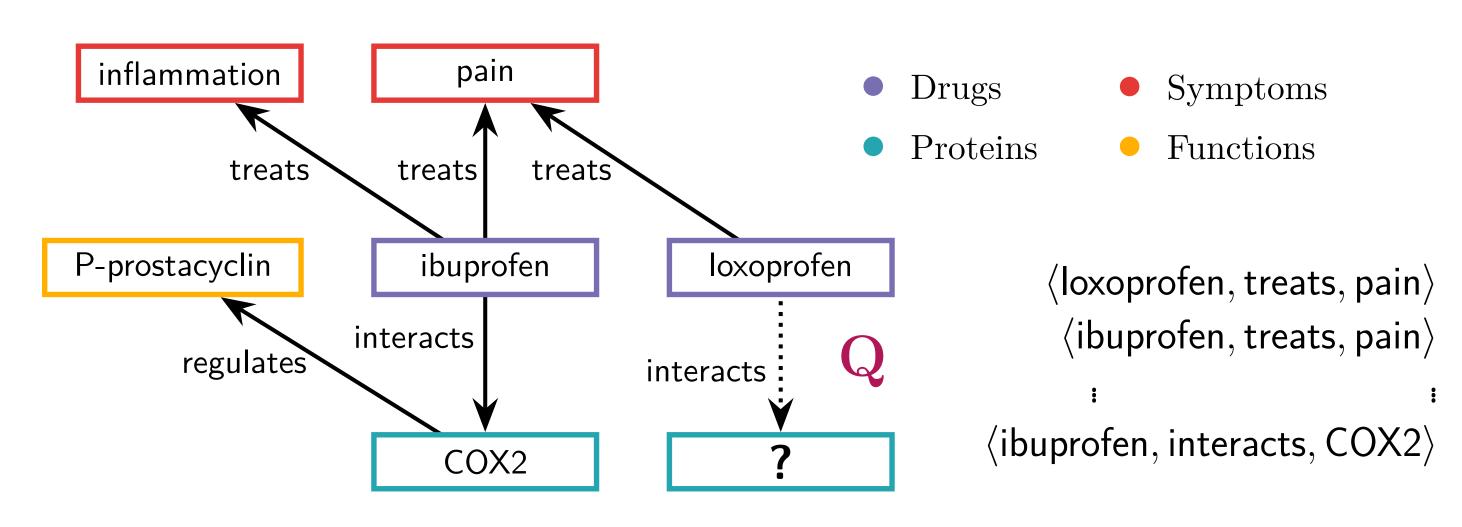


Paper





Knowledge graphs (KGs)



Link prediction

Q: (loxoprofen, interacts, ?)

KG embeddings (KGEs)

A KGE model maps entities ${\mathcal E}$ and predicates ${\mathcal R}$ to embeddings and defines a score function ϕ over triples (s, r, o).

E.g., ComplEx [1] defines

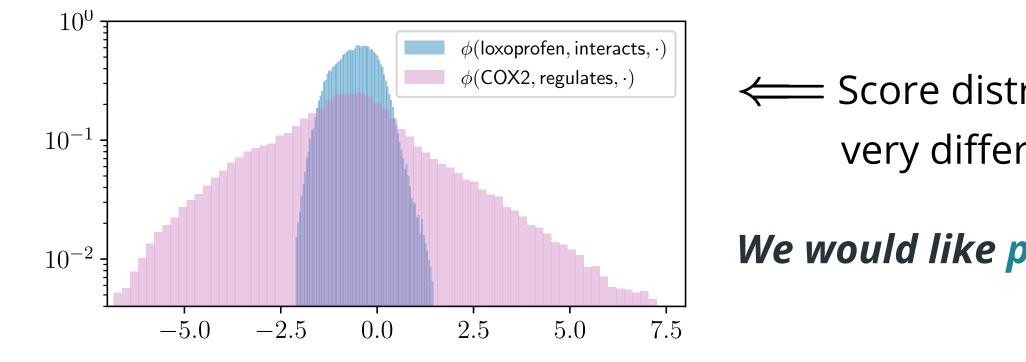
 $\phi_{\mathsf{Complex}}(s, r, o) = \operatorname{Re}\left(\langle \mathbf{e}_s, \mathbf{w}_r, \overline{\mathbf{e}_o} \rangle\right) \in \mathbb{R} \quad \mathbf{e}_s, \mathbf{w}_r, \mathbf{e}_o \in \mathbb{C}^d$

Issues?

...Opportunities! 1 2 3



Scores are difficult to interpret, combine, compare [2]



Score distributions have very different supports

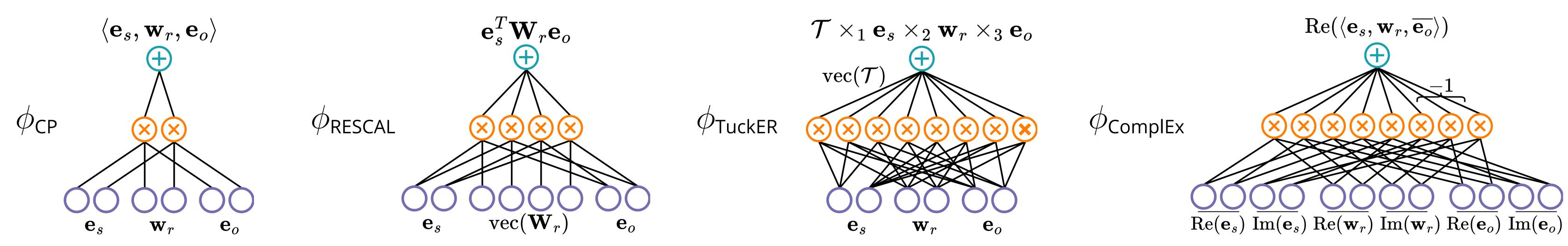
We would like probabilities instead

KGEs systematically violate logical constraints ComplEx predicts (loxoprofen, interacts, **phosp-acid**)

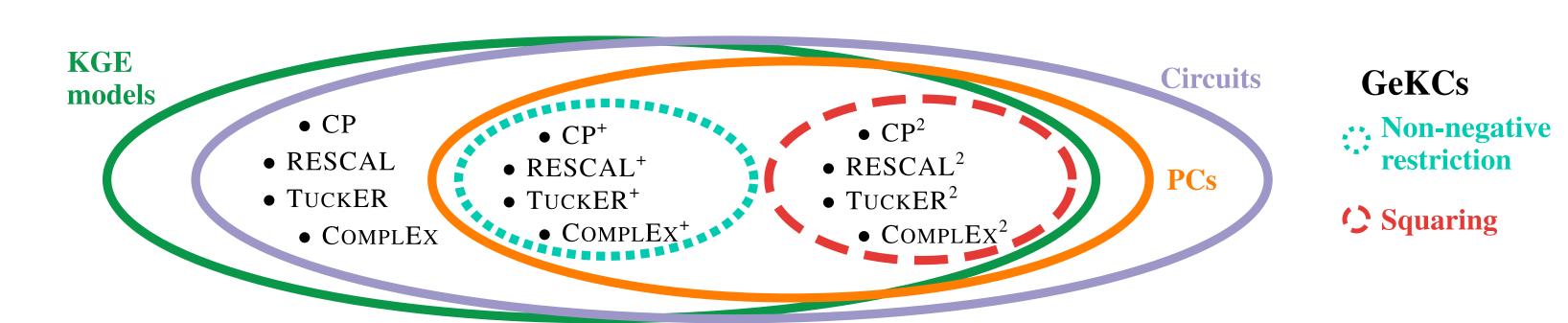
interacts can only hold between drugs and proteins

KGEs do not scale gracefully to KGs with millions of entities e.g., ComplEx takes 3 days to learn on WikiData (1 GPU)

"We convert KGE models into generative models of triples (GeKCs), making them scale to very large KGs and be reliable with constraints.



Interpreting KGE score functions as constrained computational graphs: circuits [3]



Converting KGE score functions into generative models of triples via probabilistic circuits [3]

Retrieving generative KGE circuits (GeKCs)

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

Enforce non-negative embeddings: CP⁺, ComplEx⁺, . . . OR Square the score function: CP², ComplEx², . . . ⇒ squaring circuits results in other circuits [4]

Linear-time renormalization

E.g., for ComplEx⁺

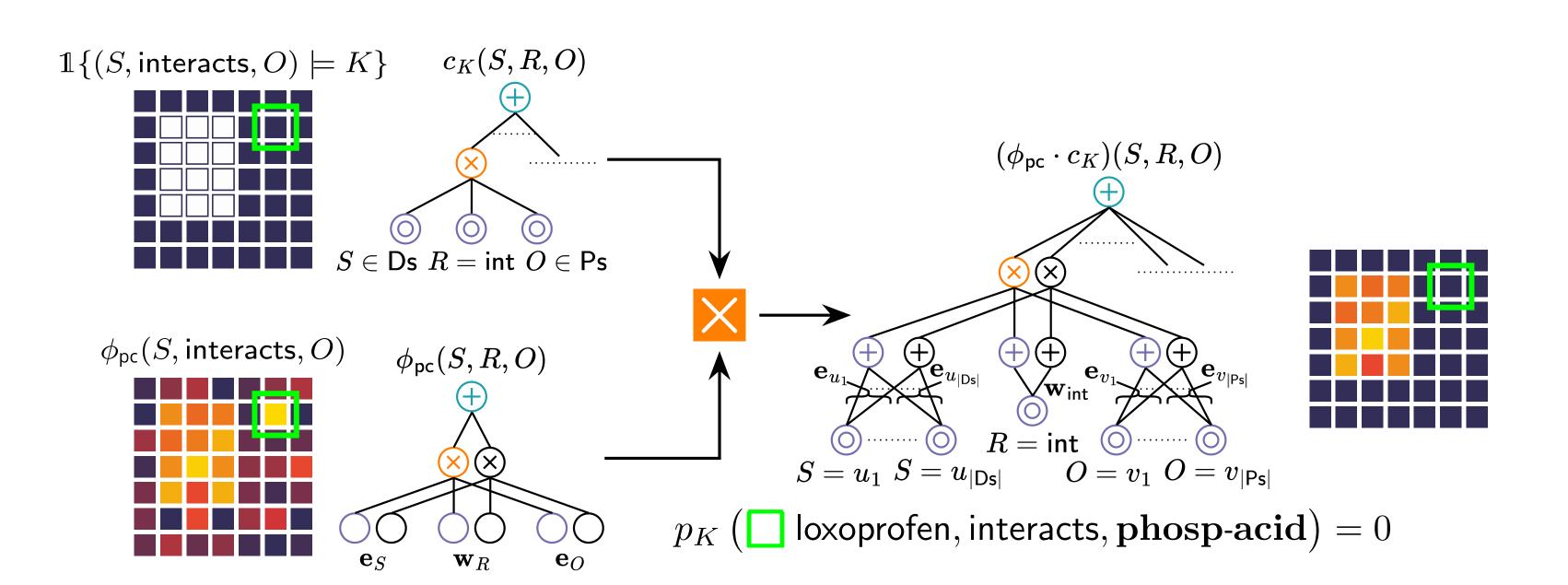
Learning GeKCs from data

Discriminative objectives [5, 6] \Longrightarrow weighted pseudo-log-likelihood

$$\mathcal{L}_{\mathsf{PLL}} := \sum\nolimits_{(s,r,o) \in \mathcal{D}} \omega_s \log p(s \mid r,o) + \omega_r \log p(r \mid s,o) + \omega_o \log p(o \mid s,r)$$

Generative objectives \Longrightarrow maximum-likelihood estimation

$$\mathcal{L}_{\mathsf{MLE}} := \sum_{(s,r,o) \in \mathcal{D}} \log p(s,r,o)$$
 New!!



2 Integration of logical constraints* with guarantees [7]

K: interacts can only hold between drugs and proteins

X (loxoprofen, interacts, phosp-acid)

✓ **(loxoprofen, interacts, COX2)**

*Any constraint over triples

Link prediction benchmarks

X	Model	FB15k-237		WN18RR		ogbl-biokg	
better) D D D D D D D D D D D D D	MOGEL	PLL	MLE	PLL	MLE	PLL	MLE
cal	СР	0.310		0.105		0.831	
pro is b	CP ⁺	0.237	0.230	0.027	0.026	0.496	0.501
recipr her is	CP^2	0.315	0.282	0.104	0.091	0.848	0.829
ean recij (higher	ComplEx	0.342		0.471		0.829	
Mean (hig	$ComplEx^{\scriptscriptstyle{\dagger}}$	0.214	0.205	0.030	0.029	0.503	0.516
Š	ComplEx ²	0.334	0.300	0.420	0.391	0.858	0.840

3 Scaling training to very large KGs

Sampling triples

Kernel triple distance (KTD) to evaluate their quality

discrepancy measure between distributions over triples [8]

Model	FB15k-237 0.589		WN18RR 0.766		ogbl-biok 1.822	
Uniform						
	PLL	MLE	PLL	MLE	PLL	M
ComplEx ⁺	0.336	0.323	0.456	0.478	0.175	0.0
ComplEx ²	0.326	0.102	0.338	0.278	0.104	0.0

References

- Théo Trouillon et al. "Complex Embeddings for Simple Link Prediction". In: ICML. Vol. 48. JMLR Workshop and Conference Proceedings. JMLR.org, 2016,
- Erik Arakelyan, Pasquale Minervini, and Isabelle Augenstein. Adapting Neural Link Predictors for Complex Query Answering. 2023. arXiv: 2301.12313
 - YooJung Choi, and Robert Peharz. "Probabilistic Circuits: Representations, Inference, Learning and Applications". In: Advances in Neural
- 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. Daniel Ruffinelli, Samuel Broscheit, and Rainer Gemulla. "You CAN Teach an Old Dog New Tricks! On Training Knowledge Graph Embeddings". In: ICLR.
- Yihong Chen et al. "Relation Prediction as an Auxiliary Training Objective for Improving Multi-Relational Graph Representations". In: CoRR abs/2110.02834

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. [8] Mikolaj Binkowski et al. "Demystifying MMD GANs". In: ICLR (Poster). OpenReview.net, 2018.