



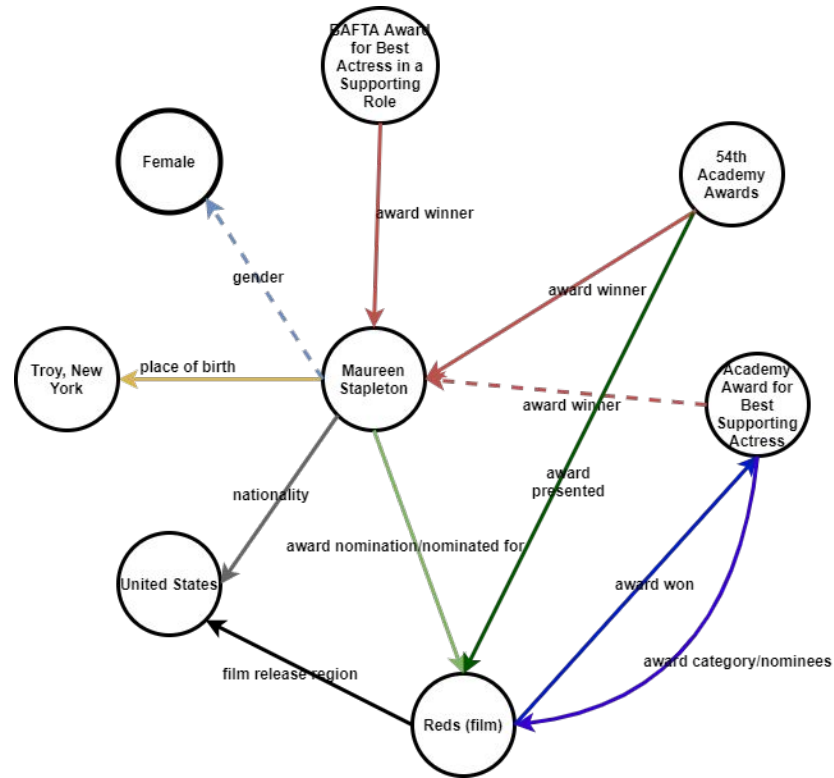
Relation Prediction as an Auxiliary Training Objective for Improving Multi-Relational Graph Representations

Yihong Chen, Pasquale Minervini, Sebastian Riedel, Pontus Stenetorp

University College London
Facebook AI Research London

Knowledge Base Completion (KBC)

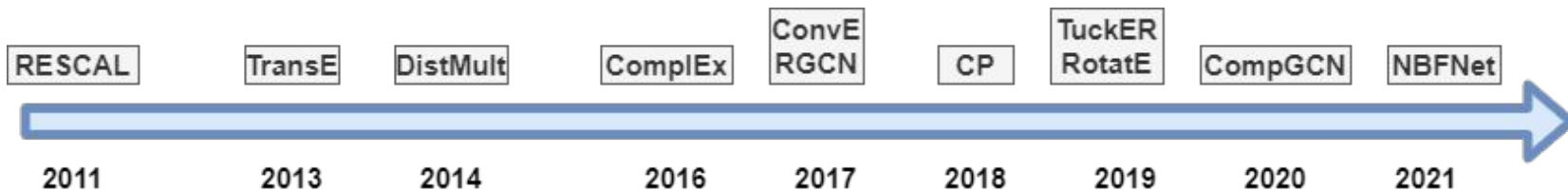
The task of KBC demands good representation learning on multi-relational graphs.





History of KBC Models

Models get more and more complex ...



Re-evaluation of KBC Models

A Re-evaluation of Knowledge Graph Completion Methods

Zhiqing Sun^{1*} Shikhar Vashishth^{1,2*} Soumya Sanyal^{2*}
Partha Talukdar² Yiming Yang¹

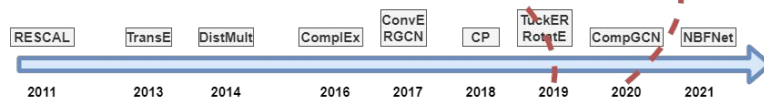
¹ Carnegie Mellon University, ² Indian Institute of Science

Knowledge Base Completion: Baseline Strikes Back (Again)

Prachi Jain¹, Sushant Rathi¹, Mausam¹ and Soumen Chakrabarti²

¹ Indian Institute of Technology Delhi

² Indian Institute of Technology Bombay



Knowledge Base Completion: Baselines Strike Back

Rudolf Kadlec and Ondrej Bajgar and Jan Kleindienst
IBM Watson
V Parku 4, 140 00 Prague, Czech Republic

YOU CAN TEACH AN OLD DOG NEW TRICKS! ON TRAINING KNOWLEDGE GRAPH EMBEDDINGS

Daniel Ruffinelli^{*}
Data and Web Science Group
University of Mannheim, Germany

Samuel Broschelt^{*}
Data and Web Science Group
University of Mannheim, Germany

Rainer Gemulla^{*}
Data and Web Science Group
University of Mannheim, Germany

Loss Function Reciprocal

1vsAll

Optimizer

Negative Sampling

Regularizer

Simple Models with Appropriate Training Strategies

≈

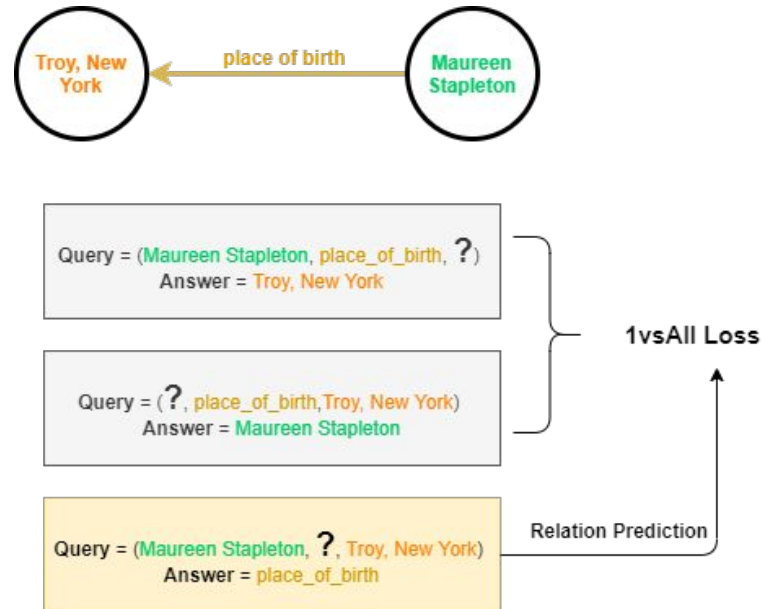
Complicated Models

?!

Relation Prediction as An Auxiliary Training Objective for KBC

A new self-supervised training objective

- not only predicting entities
- but also predicting relations





Including Relation Prediction into 1vsAll Objective

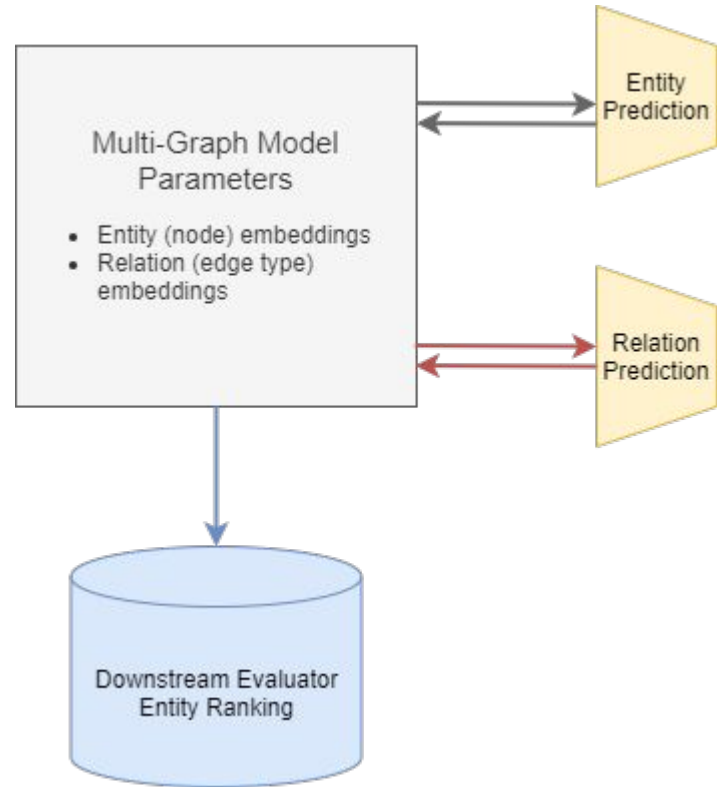
$$\arg \max_{\theta \in \Theta} \sum_{\langle s, p, o \rangle \in \mathcal{G}} [\log P_{\theta}(s \mid p, o) + \log P_{\theta}(o \mid s, p) + \lambda \log P_{\theta}(p \mid s, o)]$$

with $\log P_{\theta}(p \mid s, o) = \phi_{\theta}(s, p, o) - \log \sum_{p' \in \mathcal{R}} \exp [\phi_{\theta}(s, p', o)],$

λ : hyper-parameter balancing the entity prediction and relation prediction

Experiments

How does relation prediction impact various models on different datasets?





Experimental Setup

Models RESCAL, ComplEx, CP, Tucker

extensive hyper-parameters search
based on validation set
 $\approx 41,316$ runs

Dataset	$ \mathcal{E} $	$ \mathcal{R} $	#Train	#Validation	#Test
Nations	14	55	1,592	100	301
UMLS	135	46	5,216	652	661
Kinship	104	25	8,544	1,068	1,074
WN18RR	40,943	11	86,835	3,034	3,134
FB15k-237	27,395	237	272,115	17,535	20,466
Aristo-v4	44,950	1,605	242,594	20,000	20,000

Test Performance on All Datasets

Dataset	Entity Prediction	Relation Prediction	MRR	Hits@1	Hits@3	Hits@10
Kinship	✗	✓	0.920	0.867	0.970	0.990
	✓	✗	0.897	0.835	0.955	0.987
	✓	✓	0.916	0.866	0.964	0.988
Nations	✗	✓	0.686	0.493	0.871	0.998
	✓	✗	0.813	0.701	0.915	1.000
	✓	✓	0.827	0.726	0.915	0.998
UMLS	✗	✓	0.863	0.795	0.914	0.979
	✓	✗	0.960	0.930	0.991	0.998
	✓	✓	0.971	0.954	0.986	0.997

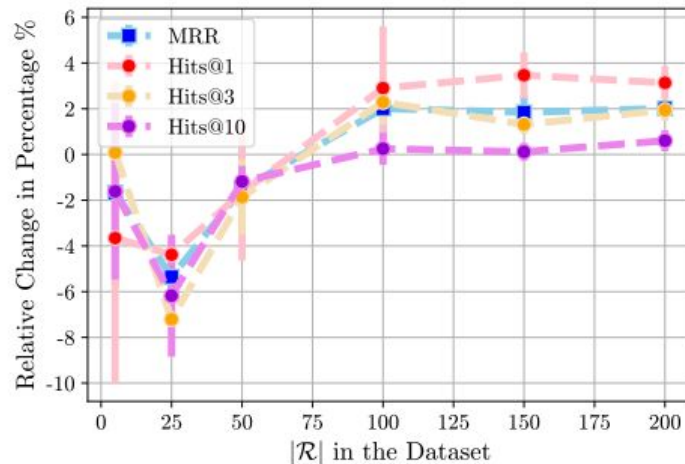
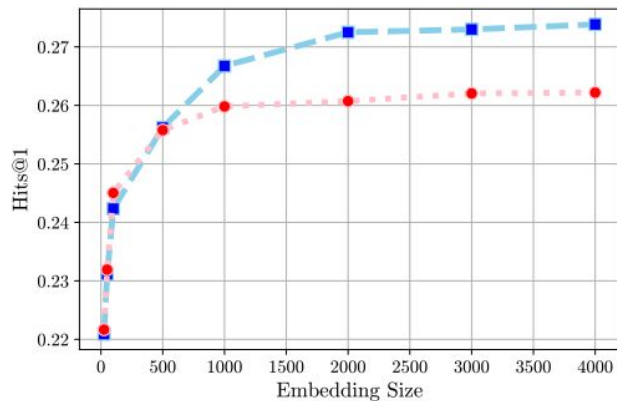
Dataset	Entity Prediction	Relation Prediction	MRR	Hits@1	Hits@3	Hits@10
WN18RR	✗	✓	0.258	0.212	0.290	0.339
	✓	✗	0.487	0.441	0.501	0.580
	✓	✓	0.488	0.443	0.505	0.578
FB15K-237	✗	✓	0.263	0.187	0.287	0.411
	✓	✗	0.366	0.271	0.401	0.557
	✓	✓	0.388	0.298	0.425	0.568
Aristo-v4	✗	✓	0.169	0.120	0.177	0.267
	✓	✗	0.301	0.232	0.324	0.438
	✓	✓	0.311	0.240	0.336	0.447



Test Performance Across Various Models, $\lambda=1$

Model	Relation Prediction	MRR	Hits@1	Hits@3	Hits@10
CP	✗	0.356	0.262	0.392	0.546
	✓	0.366	0.274	0.401	0.550
ComplEx	✗	0.366	0.271	0.401	0.557
	✓	0.382	0.289	0.419	0.568
RESCAL	✗	0.356	0.266	0.390	0.532
	✓	0.359	0.271	0.395	0.533
TuckER	✗	0.351	0.260	0.386	0.532
	✓	0.354	0.264	0.388	0.535

Ablation Study: Embedding Size & Number of Relation Types





Summary

Relation Prediction as an Auxiliary
Objective for Training KBC Models

Conclusion

- a new self-supervised objective for training KBC models
- up to 9.9% boost in Hits@1 on FB15k-237

Future Work

- extend to more complex models
- downstream applications besides link prediction
 - node classification



Thank you :)

Q&A