Contextual Parameter Generation for Knowledge Graph Link Prediction

George Stoica*, Otilia Stretcu*, Emmanouil Antonios Platanios*, Tom M. Mitchell, Barnabás Póczos

{gis, ostretcu, e.a.platanios, tom.mitchell, bapoczos}@cs.cmu.edu

MACHINE LEARNING DEPARTMENT

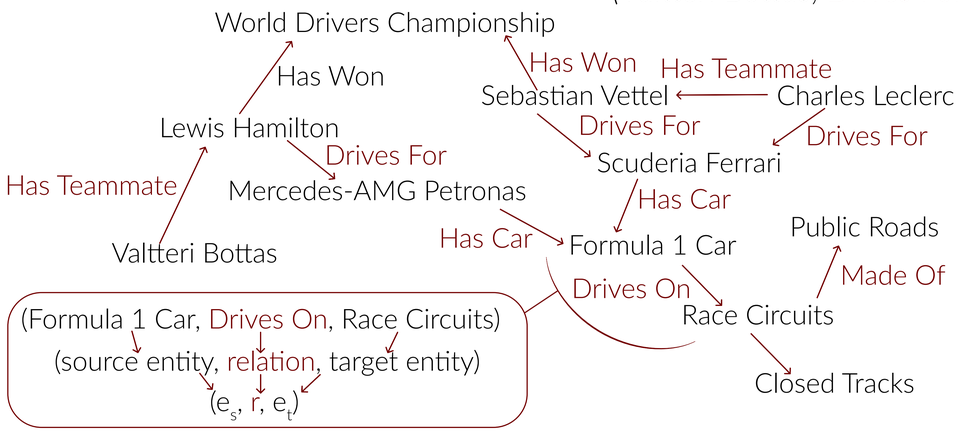
Knowledge Graphs

KGs offer a concise way in which to store factual information.

Link Prediction

Predict missing *links* to questions,

(Valtteri Bottas, Drives For, ?)



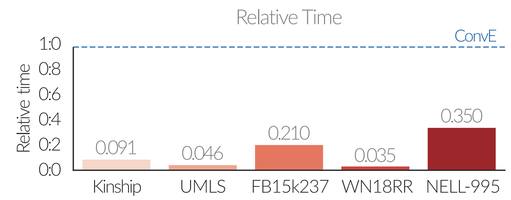
Knowledge Graph and Representation Example

Datasets

Dataset	# Train	N_e	N_r	$\bar{N}_a$	$\bar{d}$
Kinship	8,544	104	25	6.1	82.2
UMLS	5,216	135	46	7.8	26.6
FB15k237	272,115	14,541	237	3.0	17.9
WN18RR	86,835	40,945	11	1.4	2.2
NELL-995	154,213	75,492	200	3.6	4.1

Evaluation Datasets and Summary Statistics

Convergence Comparisons



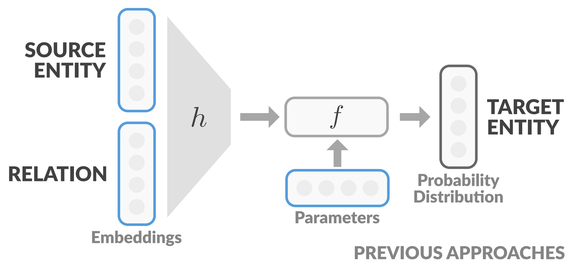
Convergence Ratio between CoPER-ConvE and ConvE

Results

Dataset	Metric	DistMult	ComplEx	NeuralLP	NTP-λ	MINERVA	MultiHop-KG	ConvE	CoPER-MINERVA	CoPER-ConvE
UMLS	Hits@1	82.1	82.3	64.3	84.3	75.3	90.2	92.9	77.8†	**95.5**‡
	Hits@10	96.7	99.5	96.2	**100.0**	96.7	99.2	99.7	97.4†	99.7‡
	MRR	86.8	89.4	77.8	91.2	84.1	94.0	95.4	85.4†	**97.1**‡
Kinship	Hits@1	48.7	75.4	47.5	75.9	60.5	78.9	74.2	66.2†	**83.6**†
	Hits@10	90.4	98.0	91.2	87.8	92.4	98.2	97.9	94.2†	**98.4**†
	MRR	61.4	83.8	61.9	79.3	72.0	86.5	83.0	76.0†	**89.5**†
WN18RR	Hits@1	43.1	41.0	37.6	–	41.3	41.8	41.9	42.7†	**44.1**†
	Hits@10	52.4	51.0	**65.7**	–	51.3	51.7	52.2	51.0†	56.1†
	MRR	46.2	44.0	46.3	–	44.8	45.0	45.2	46.5†	**48.3**†
FB15k237	Hits@1	32.4	15.8	16.6	–	22.3	**32.7**	30.3	29.5†	32.2†
	Hits@10	60.0	42.8	34.8	–	44.9	56.4	60.8	50.4†	**62.9**†
	MRR	41.7	24.7	22.7	–	29.2	40.7	40.5	36.5†	**42.6**†
NELL-995	Hits@1	55.2	64.3	–	–	64.0	65.6	67.0	65.5†	**72.2**†
	Hits@10	78.3	86.0	–	–	82.4	84.4	88.0	83.2†	**88.4**†
	MRR	64.1	72.6	–	–	71.0	72.7	75.4	72.5†	**78.7**†

CoPER Comparisons with Existing Work

Existing Solutions

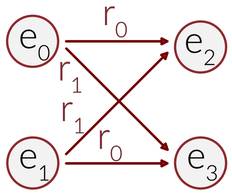


PREVIOUS APPROACHES

- *Single-Hop*: Infer answer directly from question
- *Multi-Hop*: Infer answer by traversing the KG

Additive Limitations

Existing approaches with additive interactions cannot directly represent this toy example,



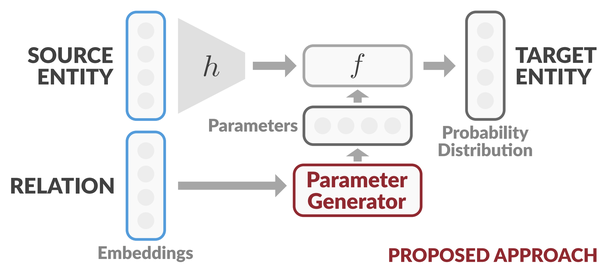
Model:

$h_\varphi(e_s, r) = \varphi[e_s; r] = \varphi_e e_s + \varphi_r r$

Solutions:

$\{e_3 = e_2 \wedge (\varphi_r = 0 \vee r_0 = r_1)\}$

CoPER



PROPOSED APPROACH

- **General**: Enhances expressive power of several models
- **Simple**: Implemented in just 10 LoCs
- **Scalable**: Improves convergence by up to 28x
- **Performance**: outperforms state-of-the-art

Parameter Generators

Three of many choices,

- **Parameter Lookup (PL)**: $\theta = W_r$
- **Linear Transformation**†: $\theta = W \cdot r$
- **Multilayer Perceptron**‡: $\theta = \text{MLP}(r)$

Design Comparisons

Dataset	Metric	ConvE	CoPER-PL-ConvE	CoPER-ConvE
UMLS	Hits@1	92.9	73.8	**95.5**‡
	Hits@10	**99.7**	99.1	**99.7**‡
	MRR	95.4	85.2	**97.1**‡
Kinship	Hits@1	74.2	74.9	**83.6**†
	Hits@10	97.9	96.6	**98.4**†
	MRR	83.0	83.2	**89.5**†
WN18RR	Hits@1	41.9	**44.1**	**44.1**†
	Hits@10	52.2	51.2	**56.1**†
	MRR	45.2	46.6	**48.3**†
FB15k237	Hits@1	30.3	30.7	**32.2**†
	Hits@10	60.8	60.0	**62.9**†
	MRR	40.5	40.5	**42.6**†
NELL-995	Hits@1	67.0	N/A	**72.2**†
	Hits@10	88.0	N/A	**88.4**†
	MRR	75.4	N/A	**78.7**†

CoPER Design Comparisons

Relation Similarities

NELL-995 Relation Pair	Cosine Distance
(topmemberoforganization, ceoof)	0.01
(citylocatedincountry, citycapitalofcountry)	0.01
(organizationheadquarteredincity, radiostationincity)	0.01
(airportincity, buildinglocatedincity)	0.02
(athleteplaysforteam, athleteplaysforteam)	0.03
(organizationheadquarteredincity, televisionstationincity)	0.05
⋮	⋮
(statelocatedincountry, teamalsoknownas)	1.47
(teamalsoknownas, athletewinsawardtrophytournament)	1.47
(countrylocatedingeopoliticallocation, agentinvolvedwithitem)	1.47
(synonymfor, teamplayssport)	1.48
(sportschoolincountry, countrycurrency)	1.49
(synonymfor, sportsgamesport)	1.52

Learned NELL-995 Relation Pairwise Cosine Distances

Repository

https://github.com/otiliastr/coper