HyConvE: A Novel Embedding Model for Knowledge Hypergraph Link Prediction with Convolutional Neural Networks

Chenxu Wang
College of Intelligence and Computing
Tianjin University
Tianjin, China
Tianjin Key Laboratory of Cognitive Computing and Application
Tianjin, China
cxwang1998@tju.edu.cn

Xin Wang*
College of Intelligence and Computing
Tianjin University
Tianjin, China
Tianjin Key Laboratory of Cognitive Computing and Application
Tianjin, China
wangx@tju.edu.cn

Zhao Li
College of Intelligence and Computing
Tianjin University
Tianjin, China
Tianjin Key Laboratory of Cognitive Computing and Application
Tianjin, China
lizh@tju.edu.cn

Zirui Chen
College of Intelligence and Computing
Tianjin University
Tianjin, China
Tianjin Key Laboratory of Cognitive Computing and Application
Tianjin, China
zrchen@tju.edu.cn

Jianxin Li
School of Information Technology
Deakin University
Geelong, Victoria, Australia
jianxin.li@deakin.edu.cn

https://github.com/CarlllWang/HyConvE/tree/master

—— WWW 2023

1.Introduction

2.Method

3.Experiments

Introduction

The knowledge hypergraph link prediction task aims at **predicting missing component** in n-ary facts, where the missing component can be either an entity in the tuple $r(e_1, e_2, ..., ?, ..., e_k)$ or a n-ary relation $?(e_1, e_2, ..., e_k)$.

Tuple	`KidsOf`	(Jeffrey Jordan, Marcus Jordan, Jeffrey Jordan)
	`BestHelper`	(Scottie Pippen, Michael Jordan, 1998)
	`PlayRoleIn`	(Michael Jordan, Scoring Guard, Chicago Bulls)

Introduction

Due to the exponential growth of multi source information, it becomes challenging, even impossible, for the large-scale n-ary knowledge base to be updated in an appropriate way, resulting in incomplete and outdated knowledge hypergraphs.

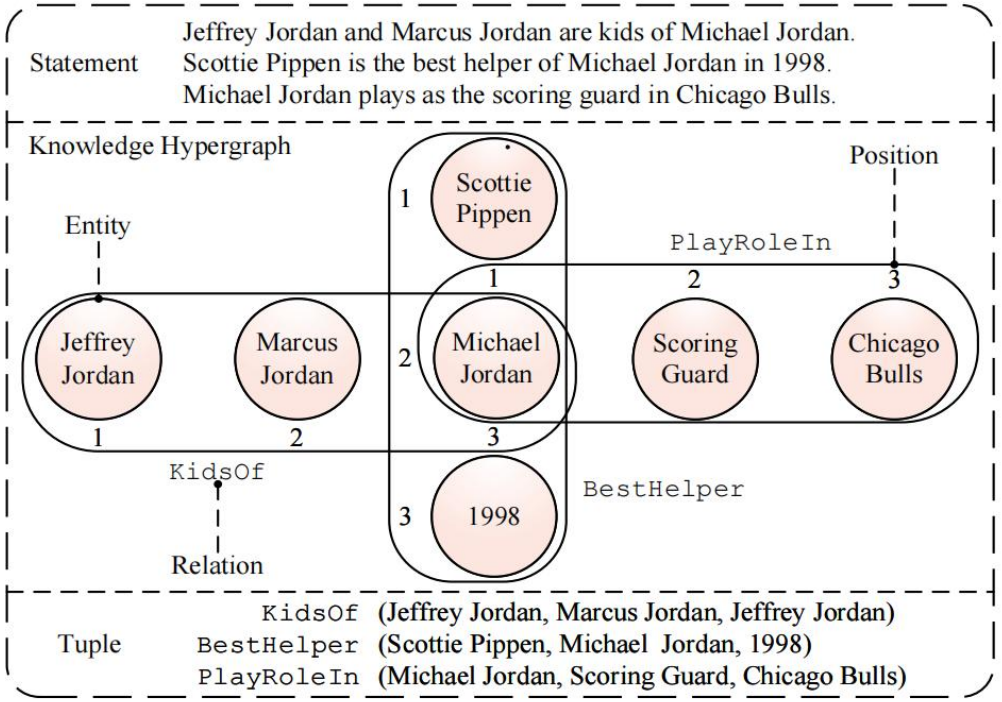


Figure 1: A real-world example of knowledge hypergraph about a set of facts related to *Michael Jordan*, where each tuple is accompanied by different positional information.

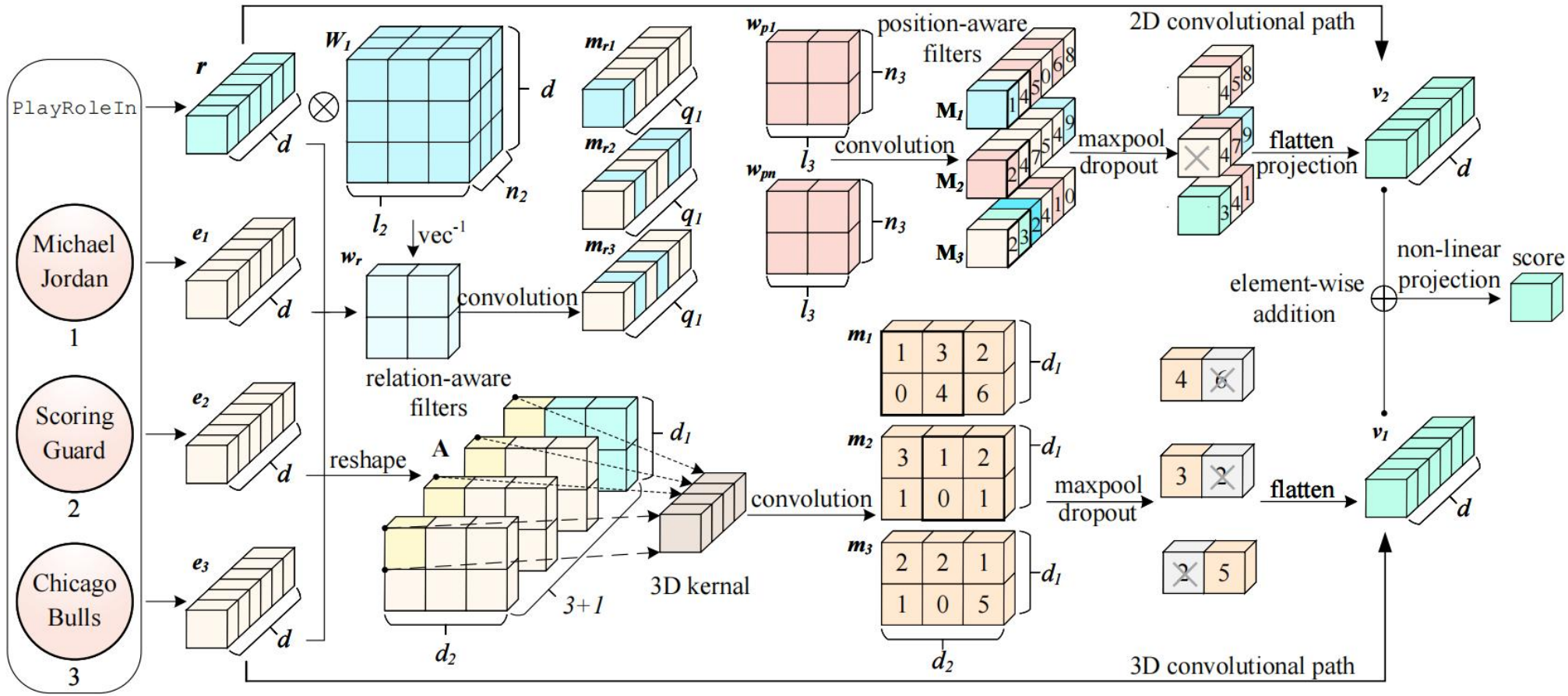
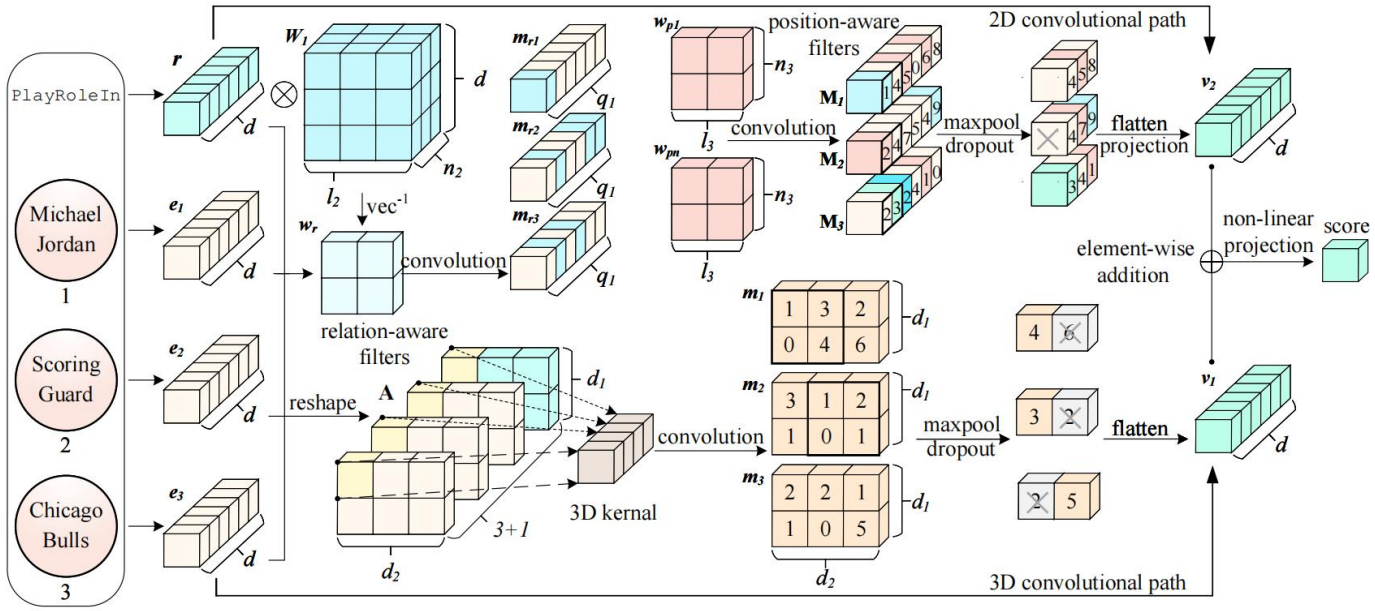


Figure 2: The framework of the HyConvE model.

Method



$$\mathcal{H} = (\mathcal{E}, \mathcal{R}, \mathcal{T}_O) \qquad t = r(e_1, e_2, ..., e_k)$$

$$m_i = \boldsymbol{w_i} * \mathbf{A} + b = \boldsymbol{w} * [\overline{\boldsymbol{r}} \,||\, \overline{\boldsymbol{e_1}}, ..., ||\, \overline{\boldsymbol{e_k}}] + b \qquad (1) \qquad \overline{\boldsymbol{r}} \in \mathbb{R}^{d_1 \times d_2} \qquad \overline{\boldsymbol{e_i}} \in \mathbb{R}^{d_1 \times d_2} \qquad A = [\overline{\boldsymbol{r}} \,||\, \overline{\boldsymbol{e_1}}, ..., ||\, \overline{\boldsymbol{e_k}}] \in \mathbb{R}^{(k+1) \times d_1 \times d_2}$$

$$\boldsymbol{v_1} = \mathbf{maxpool}(\boldsymbol{w} * [\overline{\boldsymbol{r}} \,||\, \overline{\boldsymbol{e_1}}, ..., ||\, \overline{\boldsymbol{e_k}}]) \qquad (2)$$

$$\boldsymbol{w_r} = \mathbf{vec}^{-1}(\boldsymbol{r} \cdot \mathbf{W_1}) \qquad (3) \qquad \mathbf{W_1} \in \mathbb{R}^{d \times l_2 n_2}$$

$$m_{ri} = \boldsymbol{e_i} * \boldsymbol{w_r} = \boldsymbol{e_i} * \mathbf{vec}^{-1}(\boldsymbol{r} \cdot \mathbf{W_1}) \qquad (4)$$

$$m_{ri} \in \mathbb{R}^{q_2 \times n_2} \qquad q_2 = (d - l_2)/s + 1$$

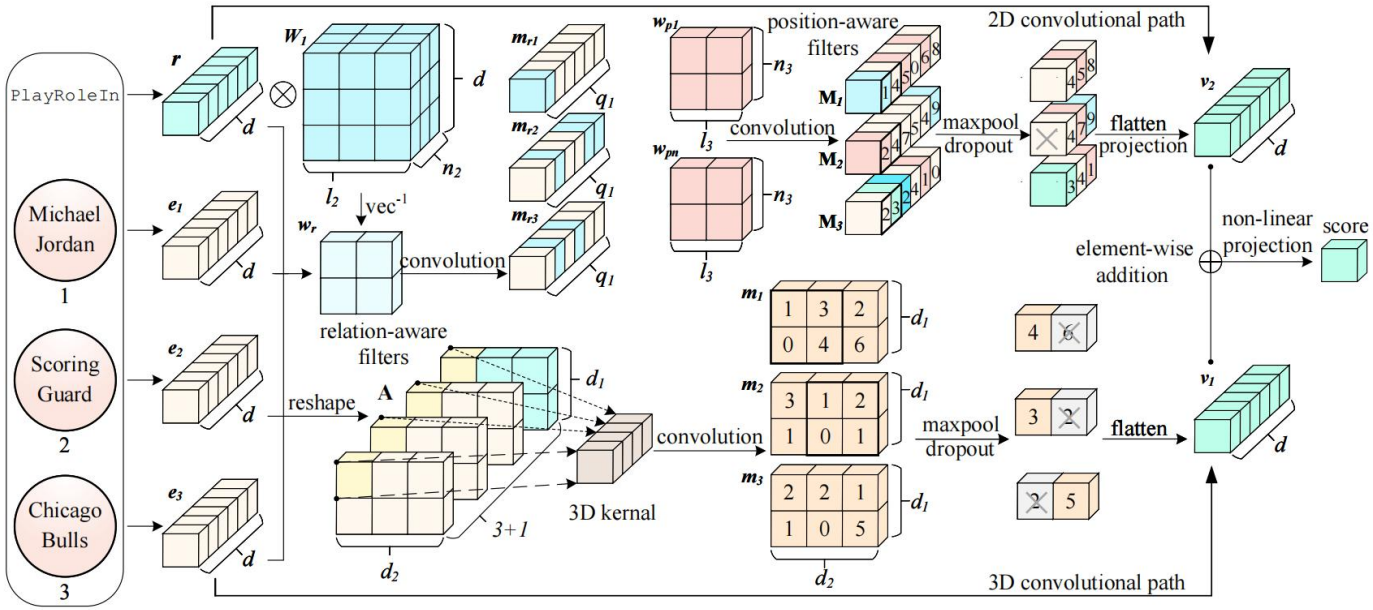
$$\mathbf{M_i} = \left[m_{ri} * \boldsymbol{w_{pi1}} \,||\, m_{ri} * \boldsymbol{w_{pi2}} \,||\, ... \,||\, m_{ri} * \boldsymbol{w_{pin_3}}\right] \qquad (5)$$

$$w_{pi} \in \mathbb{R}^{n_3 \times l_3}$$

$$\boldsymbol{v_2} = \mathbf{maxpool}(\mathbf{vec}([\mathbf{M_i} \,||\, \mathbf{M_2} \,||\, ... \,||\, \mathbf{M_k}]))\mathbf{W_2} \qquad (6)$$

$$\boldsymbol{score} = \mathrm{g}(\boldsymbol{v_1} + \boldsymbol{v_2})\mathbf{W_3} \qquad (7)$$

Method



$$\bigcup_{i=1}^{k} \mathcal{N}_x^{(i)} \equiv \bigcup_{i=1}^{k} \{e_1, \cdots, \bar{e}_i, \cdots, e_k \notin \mathcal{F} | \bar{e}_i \in \mathcal{E}, \bar{e}_i \neq e_i\} \tag{8}$$

$$\mathcal{L} = \sum_{r(e_1,\ldots,e_k) \in \{\mathcal{H} \cup \mathcal{H}'\}} \log(1+\exp(l_{r(e_1,\ldots,e_k)} \cdot f(r(e_1,\ldots,e_k)))) + \phi \tag{9}$$

$$l_{(r,e_1,\ldots,e_k)} = \begin{cases} 1 & \text{for } r(e_1,\ldots,e_k) \in \mathcal{H} \\ -1 & \text{for } r(e_1,\ldots,e_k) \in \mathcal{H}' \end{cases} \tag{10}$$

$$\phi = \lambda(\|\boldsymbol{c}\|_2^2 + \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{p}\|_2^2 + \sum_{i=1}^{k} \|\boldsymbol{e_i}\|_2^2 + \|\boldsymbol{r}\|_2^2 \tag{11}$$

Experiments

Table 1: Dataset Statistics. The size of the train, valid, and test columns represent the number of triples or tuples, respectively. "Arity" denotes the involved arities of relations.

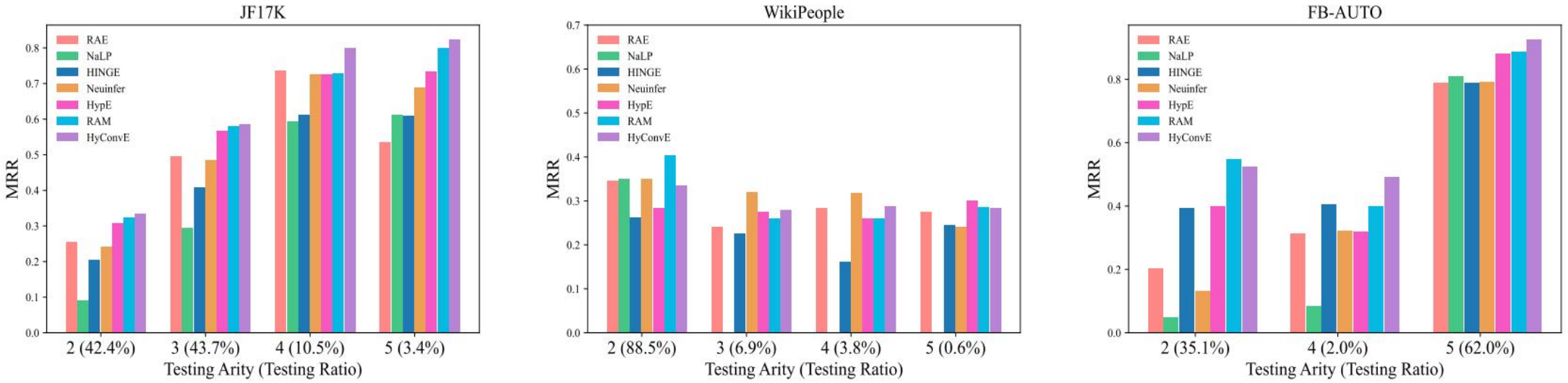
Dataset	$\|\mathcal{E}\|$	$\|\mathcal{R}\|$	Arity	# train	# valid	# test	# arity=2	# arity=3	# arity=4	# arity $\geq$ 5
FB15k-237	14, 541	237	2	272, 115	17, 535	20, 466	310, 116	–	–	–
WN18RR	40, 943	11	2	86, 835	3, 034	3, 134	93, 003	–	–	–
JF17K	29, 177	327	2-6	61, 104	15, 275	24, 568	56, 332	34, 550	9, 509	2, 267
WikiPeople	47, 765	707	2-9	305, 725	38, 223	38, 281	337, 914	25, 820	15, 188	3, 307
FB-AUTO	3, 388	8	2, 4, 5	6, 778	2, 255	2, 180	3, 786	–	125	7, 212
JF17K-3	11, 541	104	3	27, 645	3, 454	3, 455	–	34, 544	–	–
JF17K-4	6, 536	23	4	7, 607	951	951	–	–	9509	–
WikiPeople-3	12, 270	66	3	20, 656	2, 582	2, 582	–	25, 820	–	–
WikiPeople-4	9, 528	50	4	12, 150	1, 519	1, 519	–	–	15188	–

Experiments

Table 2: Results of Link Prediction on Knowledge Hypergraph Datasets. The best results are in boldface and the second best are underlined. Experimental results with "-" are those results that were not presented in the original paper. All experimental results are obtained locally.

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
RAE [41]	0.392	0.312	0.433	0.561	0.253	0.118	0.343	0.463	0.703	0.614	0.764	0.854
NaLP [15]	0.310	0.239	0.334	0.450	0.338	0.272	0.362	0.466	0.672	0.611	0.712	0.774
HINGE[28]	0.473	0.397	0.490	0.618	0.333	0.259	0.361	0.477	0.678	0.630	0.706	0.765
NeuInfer [14]	0.451	0.373	0.484	0.604	0.351	0.274	0.381	0.467	0.737	0.700	0.755	0.805
HypE [10]	0.494	0.399	0.532	0.650	0.263	0.127	0.355	0.486	0.804	0.774	0.824	0.856
tNaLP+ [13]	0.449	0.370	0.484	0.598	0.339	0.269	0.369	0.473	0.729	0.645	0.748	0.826
S2S [8]	0.528	0.457	0.570	0.690	**0.364**	0.273	0.402	**0.503**	-	-	-	-
RAM [22]	0.539	0.463	0.572	0.690	0.363	0.271	**0.405**	0.500	0.830	0.803	0.851	0.876
HyConvE (ours)	**0.590**	**0.478**	**0.610**	**0.729**	0.362	**0.275**	0.388	0.501	**0.847**	**0.820**	**0.872**	**0.901**

Experiments



(a) JF17K (b) WikiPeople (c) FB-AUTO

Figure 3: Breakdown performance across relations with different arities. x-axis identifies the relation arity and the ratio of testing samples. 6-ary relational facts and beyond are few and unreliable, thus not reported.

Experiments

Table 3: Results on fixed arity datasets. The best results are in boldface and the second best are underlined.

Model	JF17K-3			JF17K-4			WikiPeople-3			WikiPeople-4		
	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10
RAE [41]	0.505	0.430	0.644	0.707	0.636	0.835	0.239	0.168	0.379	0.150	0.080	0.273
NaLP [15]	0.515	0.431	0.679	0.719	0.673	0.805	0.301	0.226	0.445	0.342	0.237	0.540
n-CP [21]	0.669	0.613	0.801	0.754	0.701	0.855	0.313	0.237	0.476	0.253	0.163	0.432
n-tucker [21]	0.727	0.664	0.852	0.786	0.723	0.851	0.315	0.236	0.478	0.335	0.225	0.536
GETD [21]	0.725	0.660	0.858	0.822	0.761	0.924	**0.363**	**0.272**	**0.545**	0.346	0.229	0.542
RAM [22]	0.578	0.505	0.722	0.743	0.701	0.845	0.254	0.190	0.383	0.226	0.161	0.367
HyConvE (ours)	**0.729**	**0.670**	**0.861**	**0.827**	**0.770**	**0.931**	0.318	0.240	0.482	**0.386**	**0.271**	**0.607**

Experiments

Table 4: Results of Link Prediction on Knowledge Graph Datasets. The best results are in boldface and the second best are underlined. Experimental results with "-" are those results that were not presented in the original paper. All experimental results are obtained locally.

Model	FB15k-237			WN18RR			JF17K			WikiPeople			FB-AUTO		
	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10
TransE [5]	0.294	-	0.561	0.226	-	0.501	0.276	0.167	0.495	0.312	0.146	0.574	0.313	0.132	0.562
DistMult[28]	0.241	0.155	0.419	0.431	0.390	0.490	0.228	0.144	0.411	0.275	0.193	0.388	0.494	0.444	0.566
ComplEx [14]	0.253	0.158	0.428	0.440	0.411	0.512	0.308	0.219	0.498	0.326	0.232	0.461	0.487	0.442	0.568
HypE [10]	0.240	0.160	0.400	0.363	0.332	0.473	-	-	-	-	-	-	-	-	-
S2S [8]	**0.348**	**0.256**	**0.540**	**0.498**	**0.455**	**0.577**	-	-	-	-	-	-	-	-	-
RAM [22]	-	-	-	-	-	-	0.324	0.232	0.515	**0.408**	**0.313**	**0.568**	0.489	0.444	**0.576**
HyConvE (ours)	0.339	0.212	0.458	0.461	0.432	0.534	**0.338**	**0.246**	**0.525**	0.388	0.281	0.556	**0.493**	**0.445**	0.572

Experiments

Table 5: Results of ablation study. The best results are in boldface. HyConvE-path1-only means to use only the 3D path of HyConvE when conducting experiments and HyConvE-path2-only means the other.

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
HyConvE-path1-only	0.528	0.457	0.570	0.690	0.323	0.227	0.344	0.478	0.831	0.796	0.851	0.899
HyConvE-path2-only	0.102	0.054	0.094	0.168	0.072	0.048	0.094	0.172	0.145	0.082	0.164	0.212
HyConvE (ours)	**0.590**	**0.478**	**0.610**	**0.729**	**0.352**	**0.275**	**0.388**	**0.501**	**0.847**	**0.820**	**0.872**	**0.901**

Thank you!