Iterative Entity Alignment via Joint Knowledge Embeddings

Hao Zhu
Jointly work with Ruobing Xie, Zhiyuan Liu, Maosong Sun

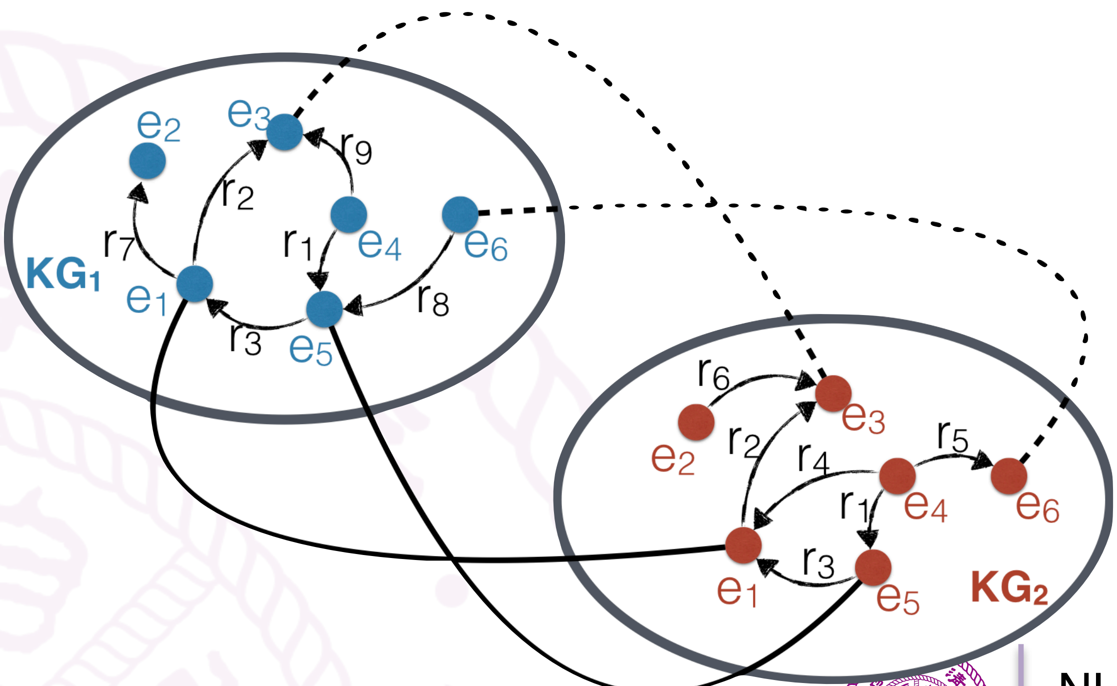
Natural Language Processing Lab,
Department of Computer Science and Technology,
Tsinghua University

ME

NLP Lab
Tsinghua University

Entity Alignment

- Goal
  – Align synonymous entity pairs from heterogeneous Knowledge Graphs



NLP Lab
Tsinghua University

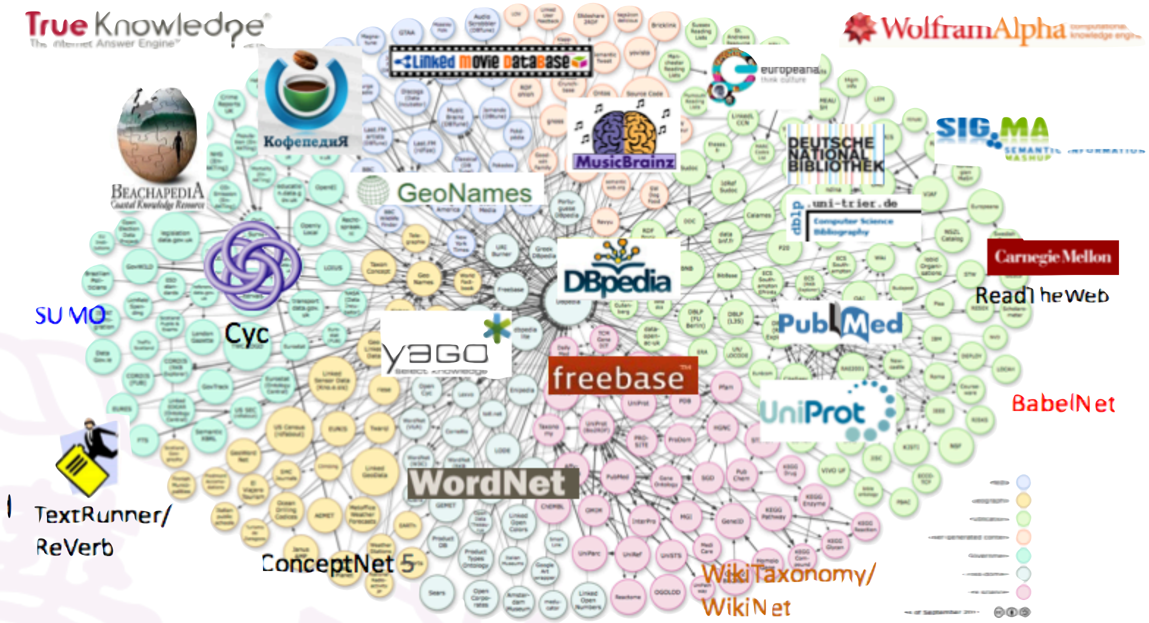
Knowledge Graphs

Knowledge Representation Learning

Knowledge Alignment

BACKGROUND

NLP Lab
Tsinghua University



Existing Knowledge Graphs

Knowledge Alignment

- Graph-based models
  - time-consuming on large-scale KGs
- Other conventional models
  - crowd-sourcing
  - well-designed hand-crafted features
- MTransE [Chen et al. 2016]
  - similar idea, but from different assumption
  - experimental result shows our method is better

NLP Lab
Tsinghua University

Knowledge Representation Learning

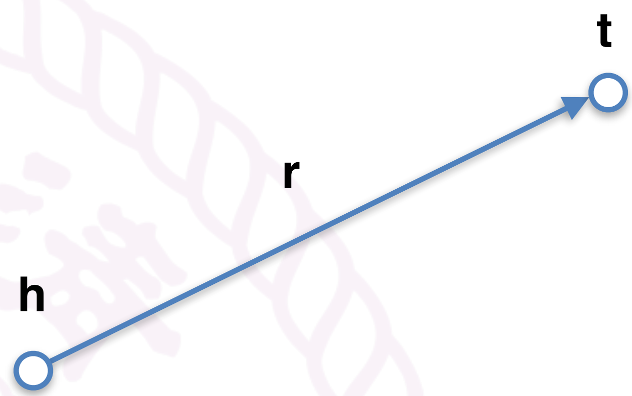
- Knowledge Representation Learning
  - TransE [Bordes *et al.*, 2013] and its extensions
  - RESCAL [Nickel *et al.*, 2011; 2012]
  - HOLE [Nickel *et al.*, 2016]
  - NTN [Socher *et al.*, 2013]

NLP Lab
Tsinghua University

TransE

IJCAI-17 MELBOURNE

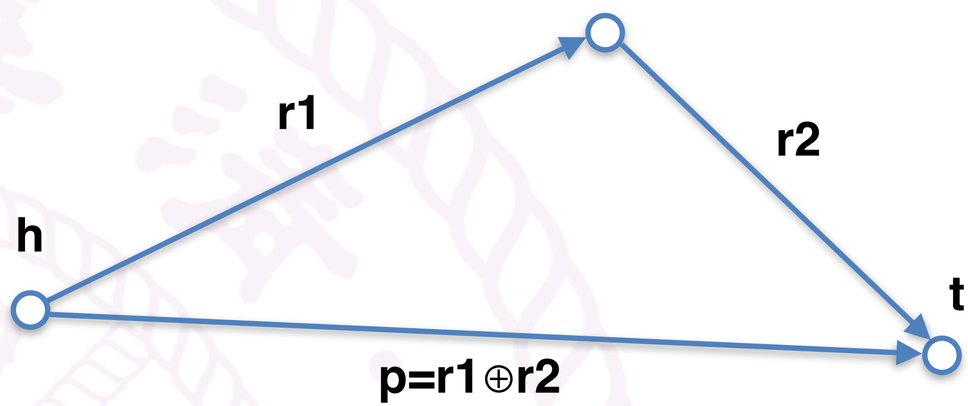
- Embedding:
  - Entity: vectors
  - Relation: translation vectors
- Goal: $\mathbf{h} + \mathbf{r} = \mathbf{t}$



NLP Lab
Tsinghua University

PTransE [Lin et. al 2015]

- Besides entities and relations, also embed relation path into the same space.



NLP Lab
Tsinghua University

What could KRL help us?

Representation ⟷ Intrinsic Meaning

Closer Representations ⟷ Higher Probability to be Synonymous

NLP Lab
Tsinghua University

Knowledge Representation Learning

Parameter Sharing Model

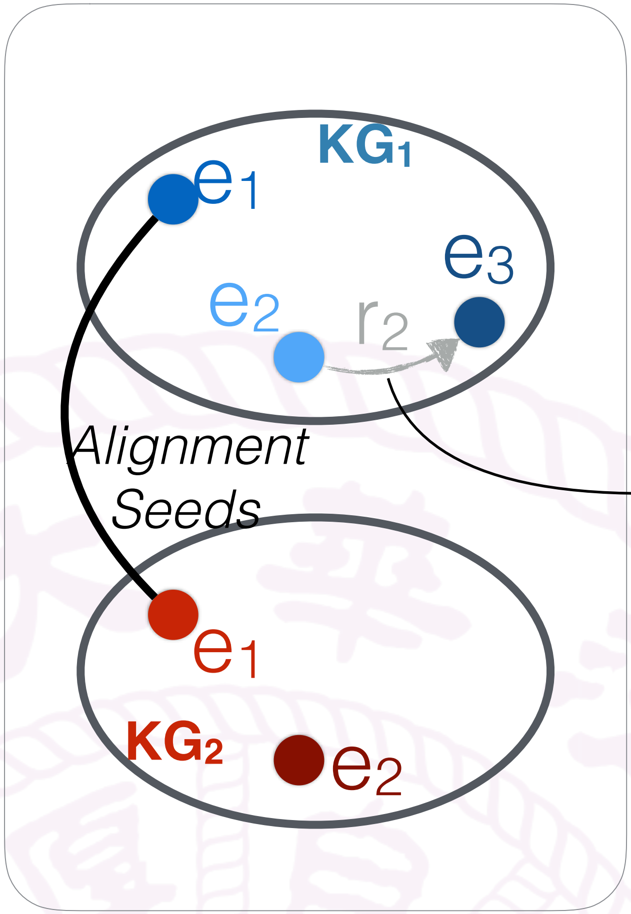
Iterative Alignment Model

Our Model(ITransE)

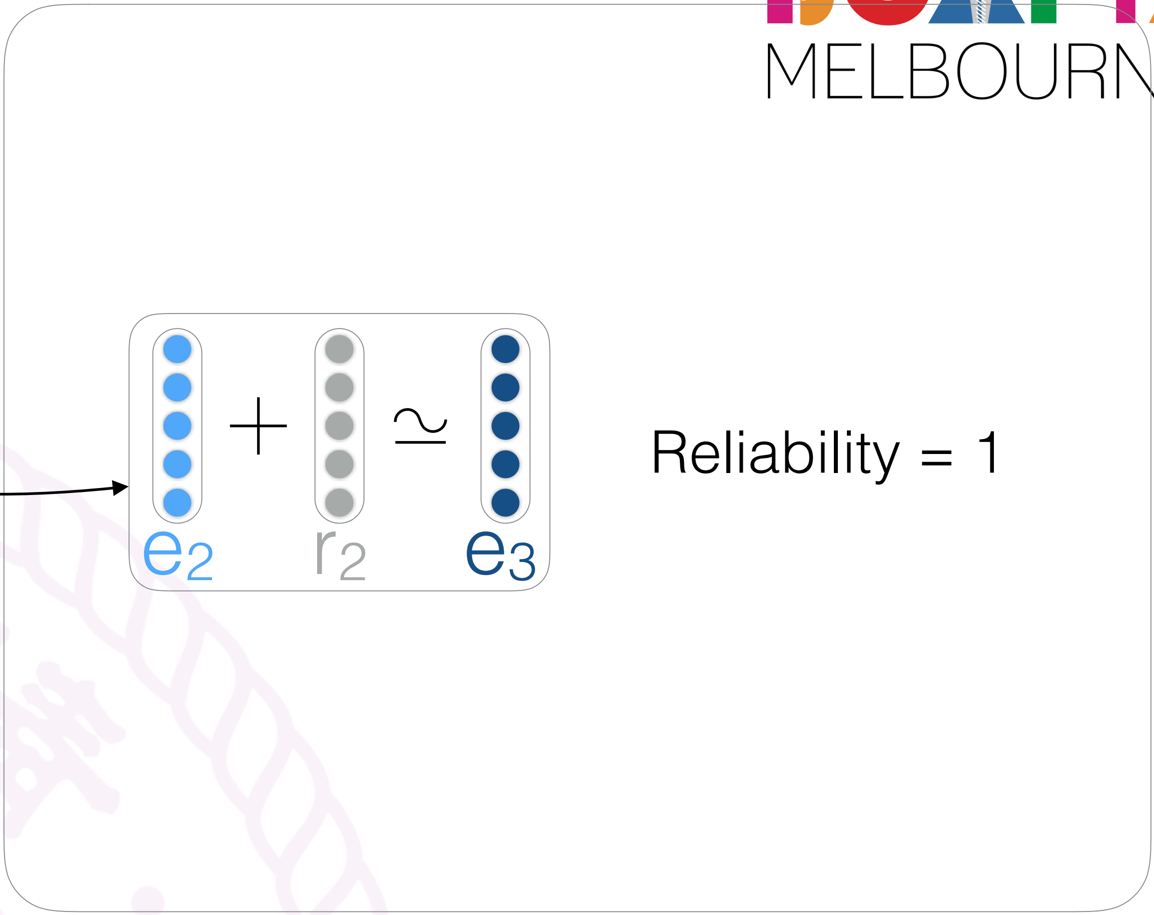
NLP Lab
Tsinghua University

Knowledge Representation Learning

IJCAI-17 MELBOURNE



Knowledge Graph

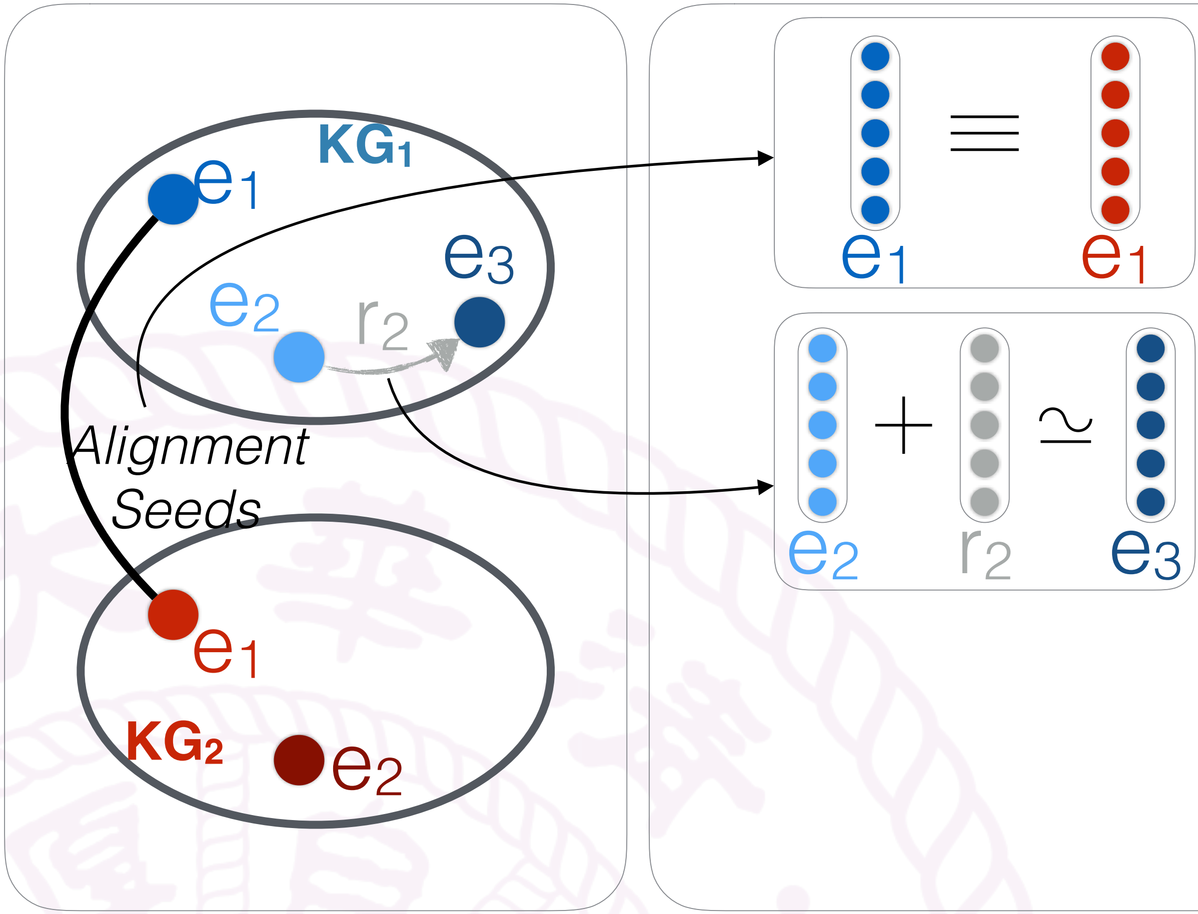


Relationship Among Embeddings

NLP Lab
Tsinghua University

Parameter Sharing Model

IJCAI-17 MELBOURNE



Reliability = 1

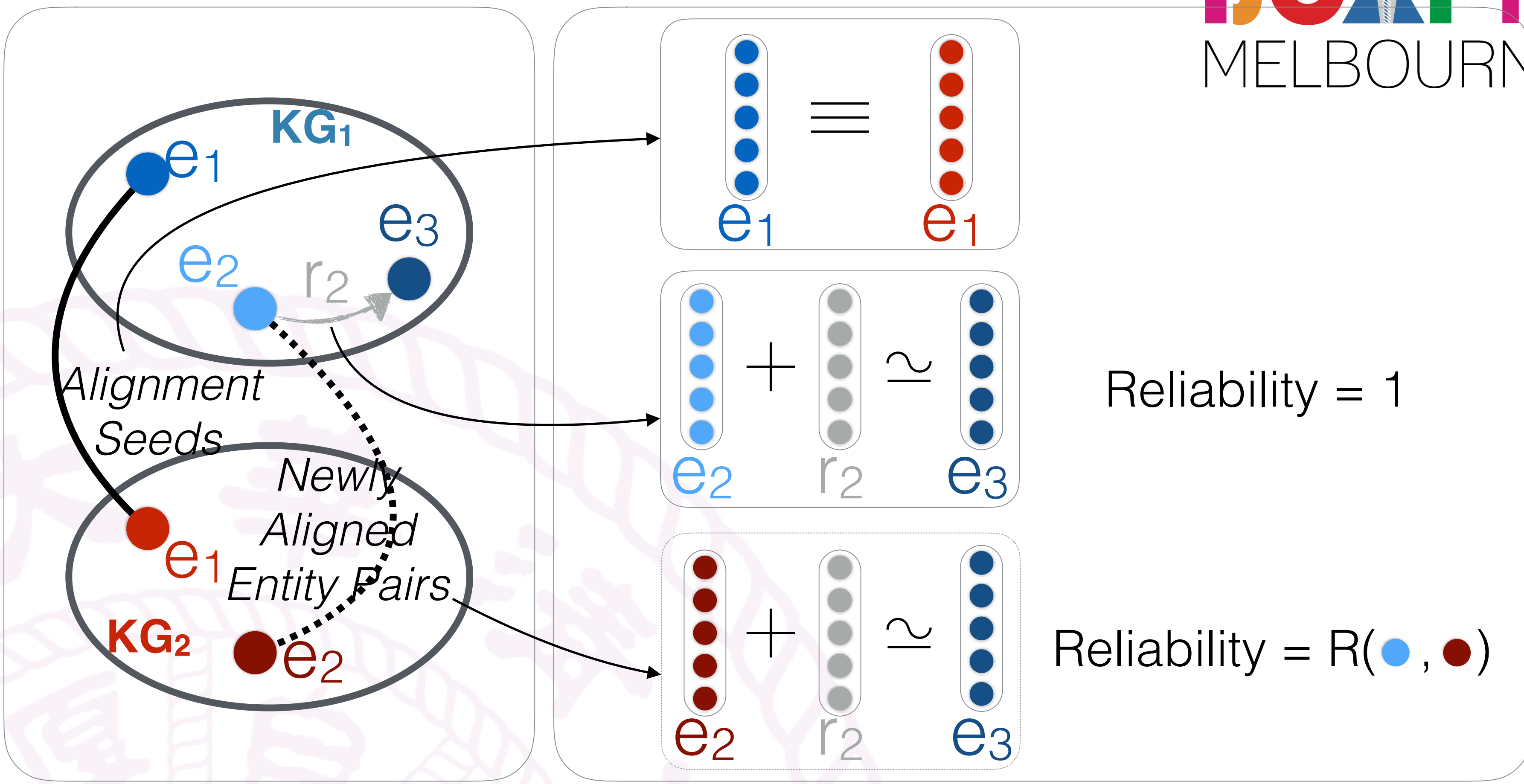
Knowledge Graph

Relationship Among Embeddings

NLP Lab
Tsinghua University

Iterative learning Model

IJCAI-17 MELBOURNE



Knowledge Graph

Relationship Among Embeddings

NLP Lab
Tsinghua University

Empirical Evaluation

- Data
  - DFB-1,2,3

Dataset	$\lvert R \rvert$	$\lvert E \rvert$	$\lvert T_1 \rvert$	$\lvert T_2 \rvert$	$\lvert \mathbb{L} \rvert$	#Valid	O
DFB-1	1,345	14,951	444,159	444,160	5,000	1,000	0.5
DFB-2	1,345	14,951	444,159	444,160	500	1,000	0.5
DFB-3	1,345	14,951	325,717	325,717	500	1,000	0.1

  - DFB-4: Training set, test set and auxiliary training set are 399, 856/59, 071/399, 857 respectively.
- Task
  - Entity Alignment
  - Knowledge Completion

NLP Lab
Tsinghua University

Entity Alignment

- Goal
  - infer the synonymous entity pairs

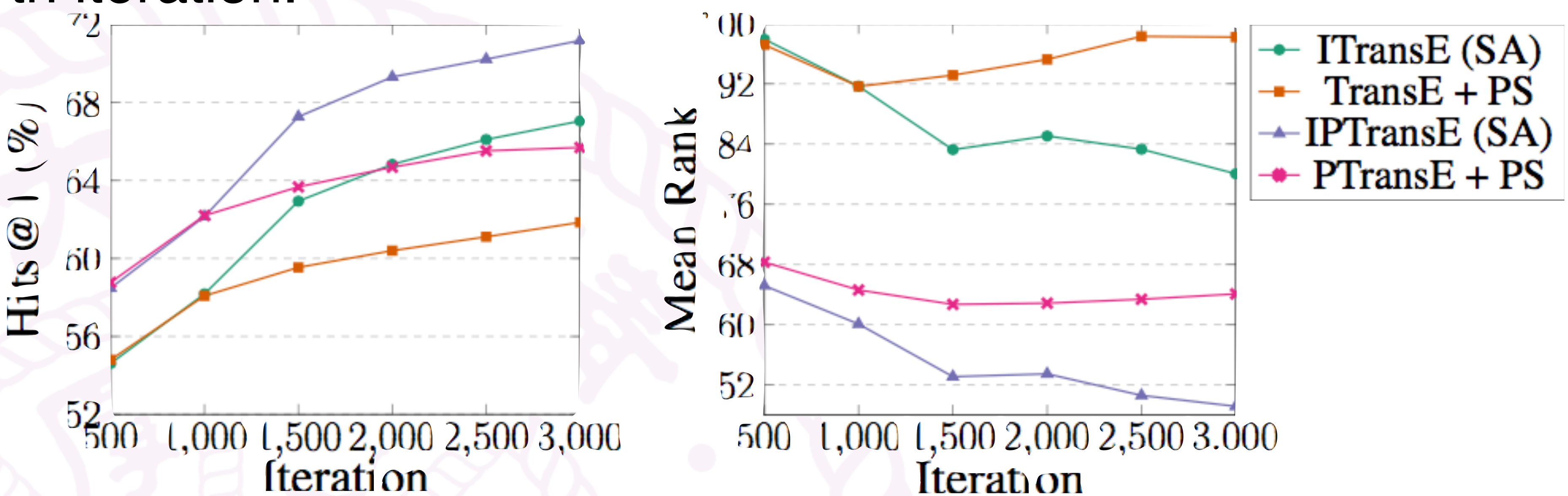
Metric	DFB-1			DFB-2			DFB-3		
	Hits@1	Hits@10	Mean Rank	Hits@1	Hits@10	Mean Rank	Hits@1	Hits@10	Mean Rank
MTransE (LT)	38.9	61.0	237.7	12.3	33.8	419.2	6.5	22.0	699.8
MTransE (TB)	13.6	35.1	547.7	13.9	35.4	675.7	4.5	16.1	1255.5
TransE + PS	61.9	79.2	105.2	41.1	67.0	154.9	12.2	34.6	431.9
ITransE (HA)	62.6	78.9	100.0	41.2	66.9	151.9	12.3	33.7	432.3
ITransE (SA)	**67.1**	**83.1**	**80.1**	**57.7**	**77.7**	**109.3**	**16.2**	**40.9**	**367.2**
PTransE + PS	65.8	83.4	62.9	46.3	72.1	96.8	15.8	40.2	346.9
IPTransE (HA)	66.1	83.3	59.1	46.2	72.6	94.2	15.1	39.7	337.6
IPTransE (SA)	**71.7**	**86.5**	**49.0**	**63.5**	**82.2**	**67.5**	**20.4**	**47.4**	**281.0**

NLP Lab
Tsinghua University

Entity Alignment

IJCAI-17 MELBOURNE

Hits@1 and Mean Rank of our methods through different iterations. (Hits@10 has similar trends to Hits@1.) We conduct soft alignment every 500 iterations from the 1000-th iteration.



NLP Lab
Tsinghua University

Knowledge Completion

- Goal
  - help learn better knowledge embeddings

Metric	Entity Prediction				Relation Prediction			
	Mean Rank		Hits@10		Mean Rank		Hits@1	
	Raw	Filter	Raw	Filter	Raw	Filter	Raw	Filter
MTransE (LT)	240.8	131.3	36.4	47.3	37.2	36.9	48.3	56.9
MTransE (TB)	851.3	759.7	9.4	10.8	293.7	293.4	27.4	27.7
TransE	246.1	131.6	42.5	54.3	55.9	55.6	44.2	50.7
TransE + Aux	232.8	121.5	43.3	54.9	50.1	49.8	44.4	50.9
ITransE (SA)	**209.2**	**101.0**	**44.2**	**55.1**	**19.8**	**19.6**	**54.2**	**60.7**
PTransE	213.0	97.2	50.9	72.1	2.33	1.96	67.4	86.9
PTransE + Aux	206.3	80.4	52.7	80.7	2.34	1.93	**68.8**	90.5
IPTransE (SA)	**197.5**	**70.6**	**53.0**	**80.8**	**2.03**	**1.62**	68.6	**90.8**

NLP Lab
Tsinghua University

Conclusion

- This paper presents iterative entity alignment via joint knowledge embeddings, by encoding both entities and relations of various KGs into a unified semantic space.
- A simple and effective Parameter Sharing Model
- An Iterative Alignment Model
- We evaluate on entity alignment and knowledge graph completion.
- Experiment results show the effectiveness of our methods as compared with other baselines.

NLP Lab
Tsinghua University

Future Work

- incorporate rich external information of KGs for entity alignment
- explore the effectiveness of other KRL models in our methods for entity alignment.

- Our **code** and **data** will be available at https://github.com/thunlp/IEAJKE

NLP Lab
Tsinghua University

Questions?

NLP Lab
Tsinghua University