Knowledge Graph Embedding with Diversity of Structures

Wen Zhang

Zhejiang University, China

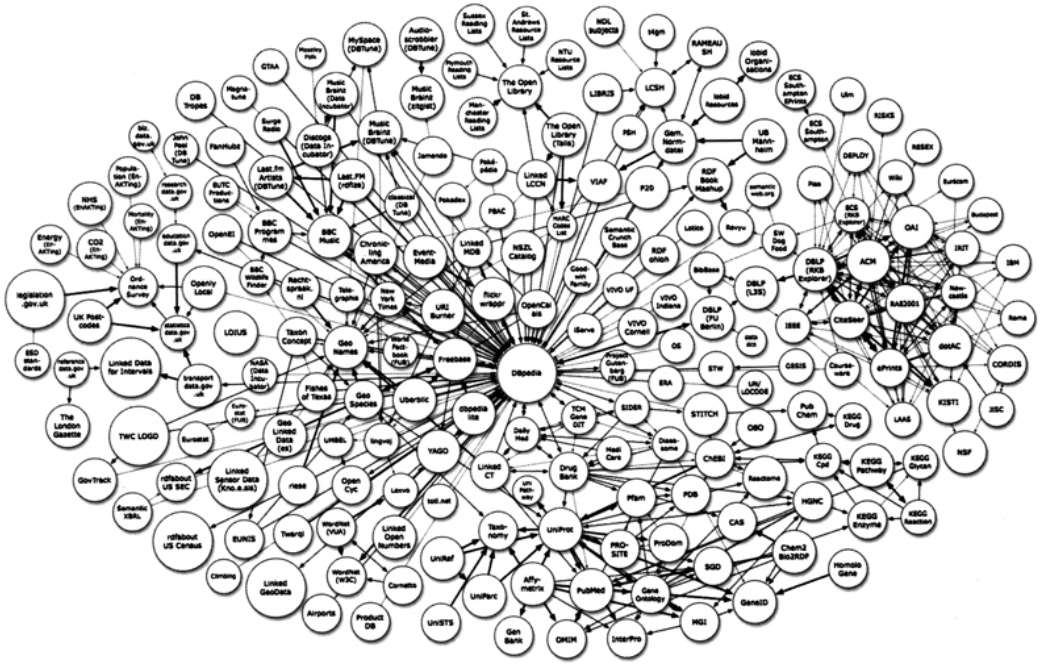
April 5, 2017

What is knowledge graph?

What is Knowledge graph embedding?

What is diversity of structures?

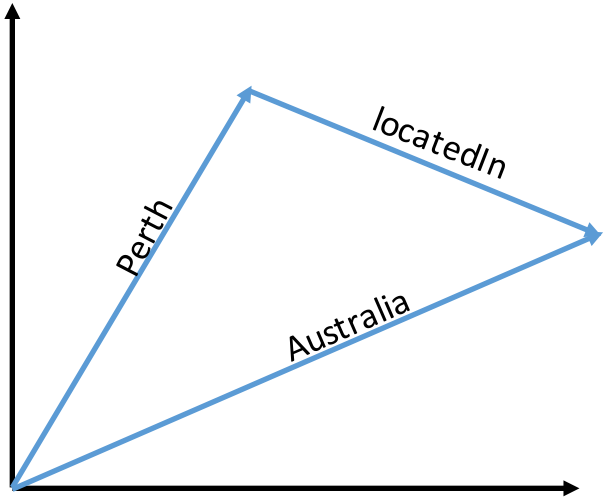
What is knowledge graph?



Triple example:
<Perth, locatedIn, Australia>

Famous large KGs:
WordNet,
Freebase,
YAGO,
Nell,
…

What is knowledge graph embedding?

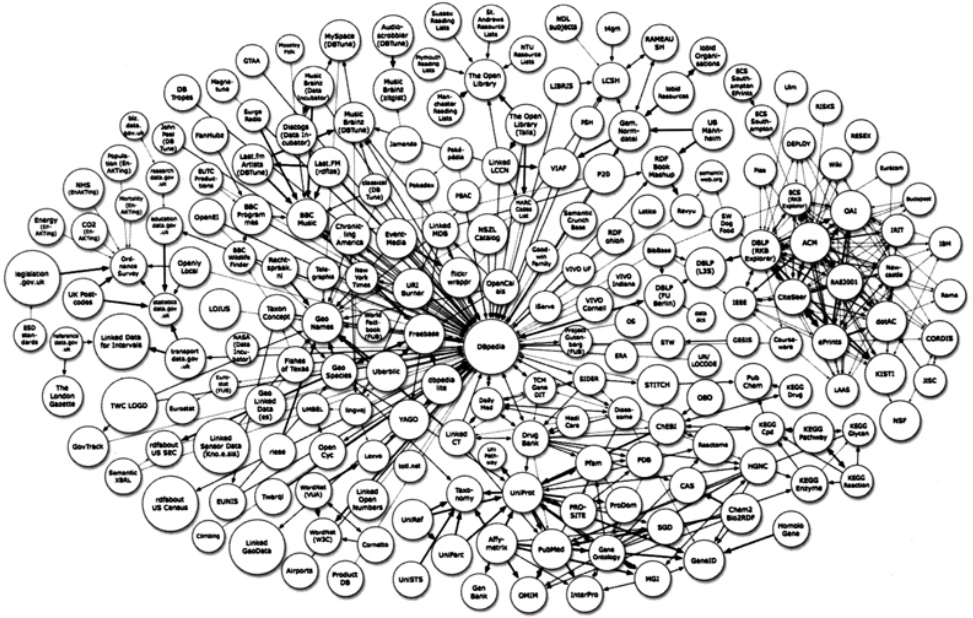


Translate-based methods

for true triple <e_1 , r , e_2>: $\boldsymbol{e}_1 + \boldsymbol{r} = \boldsymbol{e}_2$

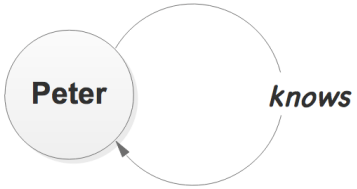
TransE, TransH, TransR, TransD

What is diversity of structures?



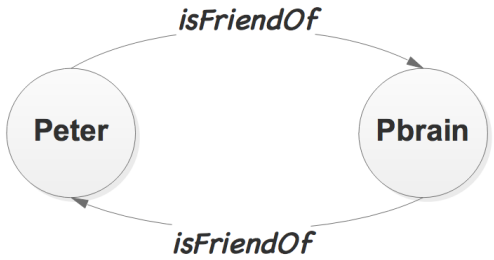
“One-relation-circle”

What is diversity of structures ? ---- ORC structures



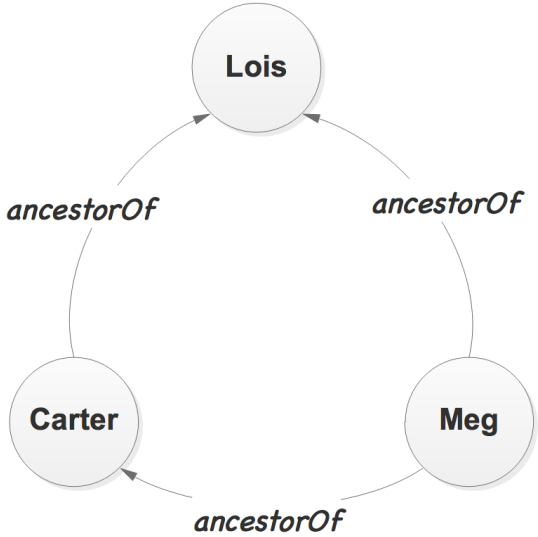
(a) example for C1

reflexive object properties



(b) example for C2

symmetric object properties



(c) example for C3

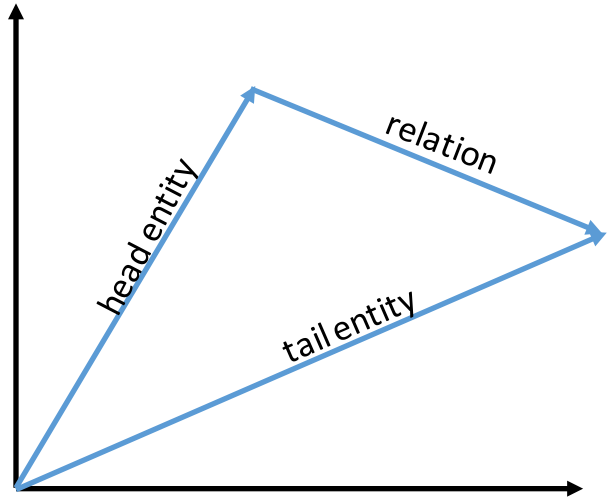
transitive object properties

Diversity of structures ---- Is ORC structures common in KG ?

	C1	C2	C3	ORC
WN18	9(0.00%)	30048(19.84%)	2237(1.48%)	32294(21.68%)
FB15k	2250(0.38%)	52376(8.84%)	59620(10.07)	114241(19.29%)

Table: the number and percent of entities which construct at least one of C1, C2 or C3

Why previous translate-based methods can't encode ORC structures?



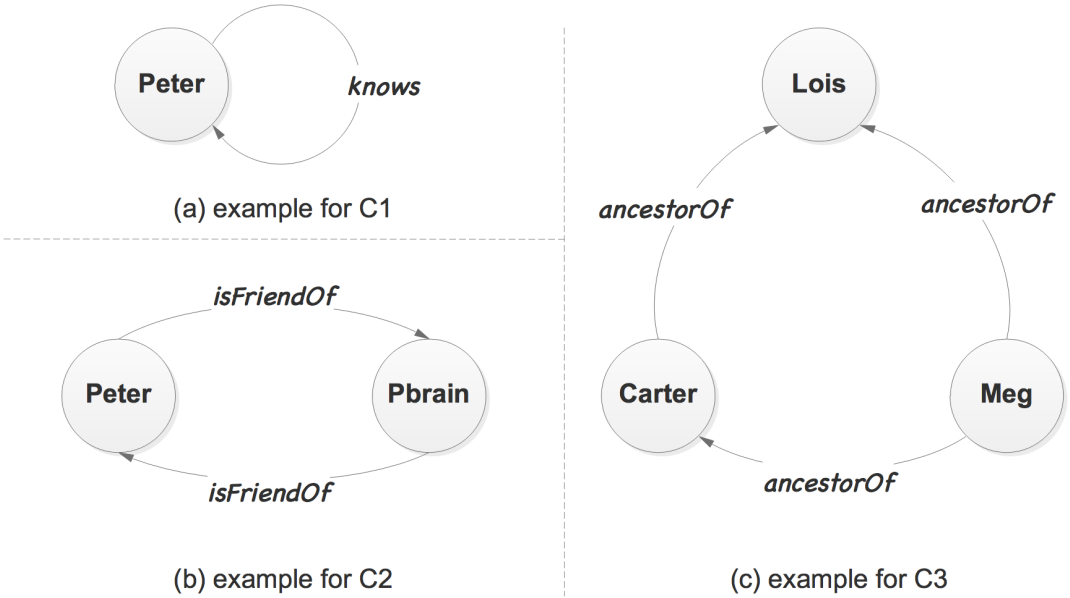
for true triple <e_1 , r , e_2>: $\boldsymbol{e_1 + r = e_2}$

C1: <e_1, r, e_1>
$\boldsymbol{e_1 + r = e_1} \Rightarrow \boldsymbol{r = 0}$

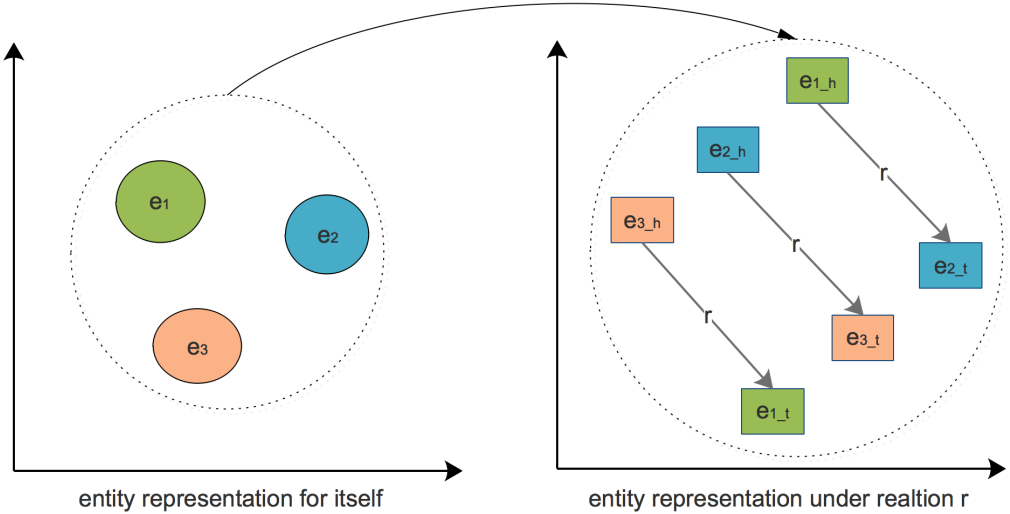
C2: <e_1, r, e_2>, <e_2, r, e_1>
$\boldsymbol{e_1 + r = e_2,\ e_2 + r = e_1} \Rightarrow \boldsymbol{r = 0}$

C3: <e_1, r, e_2>, <e_2, r, e_3>, <e_1, r, e_3>
$\boldsymbol{e_1 + r = e_1,\ e_2 + r = e_3,\ e_1 + r = e_3} \Rightarrow \boldsymbol{r = 0}$

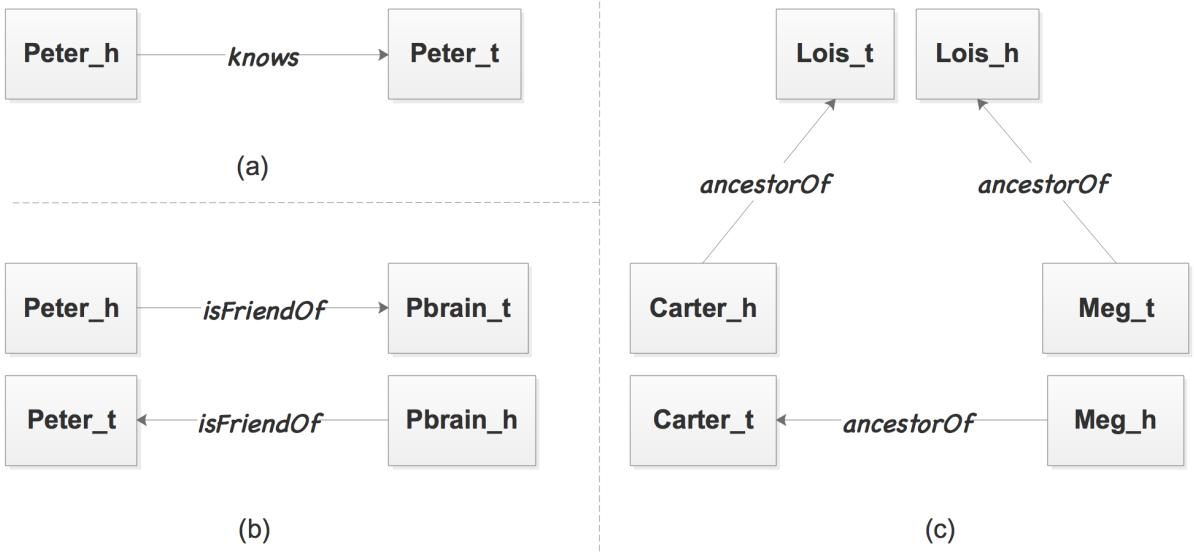
Why previous translate-based methods can't encode ORC structures?



Key idea to decompose ORC structures



Key idea to decompose ORC structures



Our method

Score function for triple <e_1, r, e_2>:

$$f_r(e_1, e_2) = \| e_{1h} + r - e_{2t} \|_{L1/L2}$$

Loss function:

$$L = \sum_{(e_1, r, e_2) \in \Delta} \sum_{(e_1', r', e_2') \in \Delta'} [f_r(e_1, e_2) + \gamma - f_r(e_1', e_2')]_+$$

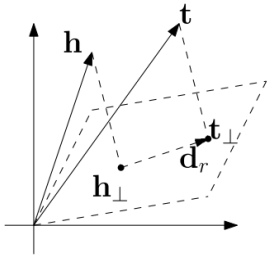
where:

$$[x]_+ = \max(0, x)$$

Apply our method to TransH, TransR and TransD

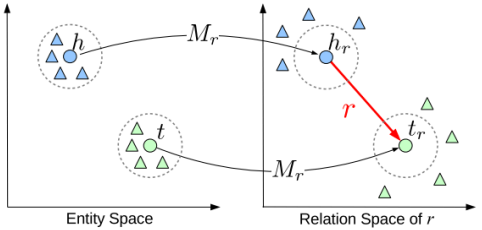
TransH(dORC)

$$e_{1h} = e_1 - w_{rh}^{\top} e_1 \cdot w_{rh}$$
$$e_{2t} = e_2 - w_{rt}^{\top} e_2 \cdot w_{rt}$$



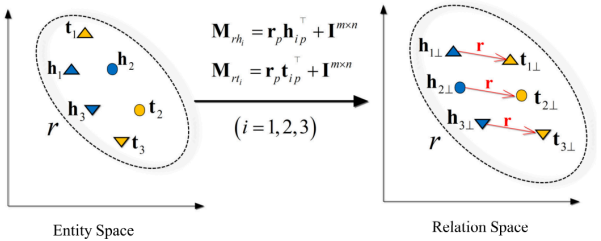
TransR(dORC)

$$e_{1h} = M_{rh} \cdot e_1$$
$$e_{2t} = M_{rt} \cdot e_2$$



TransD(dORC)

$$e_{1h} = M_{rh} \cdot e_1, \quad e_{2t} = M_{rt} \cdot e_2$$
$$M_{rh} = r_h e_{1p}^{\top} + I^{d \times d}, \quad M_{rt} = r_t e_{2p}^{\top} + I^{d \times d}$$



Experiment results

Link prediction:
<e1, r, ?> or <?, r , e2>

mean rank & hit@10

Table 3: Results on WN18 for Link Prediction

Method	Mean Rank		Hit@10(%)	
	Raw	Filter	Raw	Filter
Unstructured [1]	315	304	35.3	38.2
RESCAL [13]	1180	1163	37.2	52.8
SE [2]	1011	985	68.5	80.5
SME(linear) [1]	545	533	65.1	74.1
SME(Bilinear) [1]	526	509	54.7	61.3
LFM [17]	469	456	71.4	81.6
TransE [6]	263	251	75.4	89.2
TransH(unif) [29]	318	303	75.4	86.7
TransH(bern) [29]	401	388	73.0	82.3
TransR(bern) [26]	238	225	79.8	92.0
TransR(unif) [26]	232	219	78.3	91.7
CTransR(unif) [26]	243	230	78.9	92.3
CTransR(bern) [26]	231	218	79.4	92.3
TransD(unif) [9]	242	229	79.2	92.5
TransD(bern) [9]	224	212	79.6	92.2
TransH(dORC)(unif)	298	286	79.4	**93.3**
TransH(dORC)(bern)	278	271	80.2	93.0
TransR(dORC)(unif)	224	212	79.3	92.1
TransR(dORC)(bern)	231	219	**80.9**	92.5
TransD(dORC)(bern)	**205**	**192**	79.7	92.4

Thanks