Representation Learning of Knowledge Graphs with Entity Descriptions

Ruobing Xie[1,2], Zhiyuan Liu[1,2,3], Jia Jia[1], Huanbo Luan[1,2], Maosong Sun[1,2,3]

[1] Department of Computer Science and Technology,
[2] State Key Lab on Intelligent Technology and Systems,
National Lab for Information Science and Technology, Tsinghua University, Beijing, China
[3] Jiangsu Collaborative Innovation Center for Language Ability,
Jiangsu Normal University, Xuzhou 221009 China

Representor Tang Ning

November 20, 2017

Overview

Motivation

- Most methods concentrate on learning representations with knowledge triples indicating relations between entities. In fact, in most knowledge graphs there are usually concise descriptions for entities
- In zero-shot setting, previous triple embedding cannot work

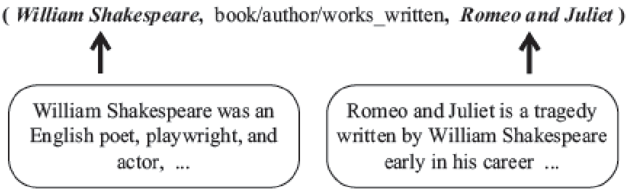


Figure 1: Example of entity descriptions in Freebase.

Problem Formulation

Definition 1. Structure-based Representations

h_s and t_s are the structure-based representations for head and tail which can directly represent entities.

Definition 2. Description-based Representations

h_d and t_d are the description-based representations for head and tail which are built from entity descriptions.

Continuous Bag-of-words Encoder

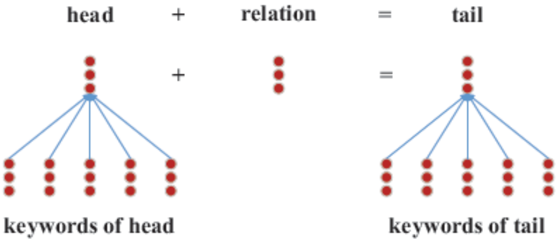


Figure 2: The CBOW Encoder

$$e_d = x_1 + x_2 + \cdots + x_k \tag{1}$$

Convolutional Neural Network Encoder

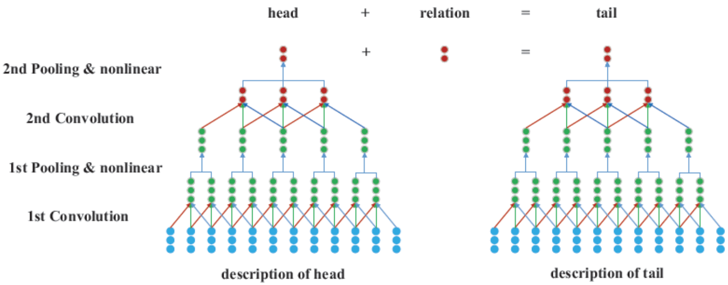


Figure 3: The Convolutional Neural Network Encoder

Convolutional Neural Network Encoder

Convolution

$$x_i'^{(1)} = x_{i:i+k-1} = [x_i^T, x_{i+1}^T, \ldots, x_{i+k-1}^T]^T \tag{2}$$

$$z_i^{(l)} = \sigma(\mathbf{W}^{(l)} x_i'^{(l)} + b_i^{(l)}) \tag{3}$$

$$\mathbf{W}^{(l)} \in \mathcal{R}^{n_2^{(l)} \times n_1^{(l)}} \tag{4}$$

$$n_1^{(l)} = k \times n_0^{(l)} \tag{5}$$

Pooling

$$x_i^{(2)} = max(z_{n\times i}^{(1)}, \ldots, z_{n\times(i+1)-1}^{(1)}) \tag{6}$$

$$x_i^{(3)} = \sum_{i=1,\ldots,m} \frac{z_i^{(2)}}{m} \tag{7}$$

Training

Energy Function

$$\mathbf{E} = \mathbf{E_S} + \mathbf{E_D} \tag{8}$$

$$\mathbf{E_D} = \mathbf{E_{DD}} + \mathbf{E_{DS}} + \mathbf{E_{SD}} \tag{9}$$

$$\mathbf{E}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = ||\mathbf{h} + \mathbf{r} - \mathbf{t}|| \tag{10}$$

objective function

$$L = \sum_{(h,r,t)} \sum_{(h',r',t') \in T'} max(\gamma + d(h + r, t) - d(h' + r', t'), 0) \tag{11}$$

DataSet

Table 1: Statistics of data sets

Dataset	#Rel	#Ent	#Train	#Valid	#Test
FB15K	1,341	14,904	472,860	48,991	57,803

Dataset	#Ent	$\#e-e$	$\#d-e$	$\#e-d$	$\#d-d$
FB20K	19,923	57,803	18,753	11,586	151

Knowledge Graph Completion

Table 2: Evaluation results on entity prediction

Metric	Mean Rank Raw	Mean Rank Filter	Hits@10(%) Raw	Hits@10(%) Filter
TransE	210	119	48.5	66.1
DKRL(CBOW)	236	151	38.3	51.8
DKRL(CNN)	200	113	44.3	57.6
DKRL(CNN)+TransE	**181**	**91**	**49.6**	**67.4**

Table 3: Evaluation results on relation prediction

Metric	Mean Rank Raw	Mean Rank Filter	Hits@1(%) Raw	Hits@1(%) Filter
TransE	2.91	2.53	69.5	90.2
DKRL(CBOW)	2.85	2.51	65.3	82.7
DKRL(CNN)	2.91	2.55	**69.8**	89.0
DKRL(CNN)+TransE	**2.41**	**2.03**	**69.8**	**90.8**

Entity Classification

Table 4: Evaluation results on entity classification

Metric	FB15K	FB20K
TransE	87.9	-
BOW	86.3	57.5
DKRL(CBOW)	89.3	52.0
DKRL(CNN)	**90.1**	**61.9**

Zero-shot Scenario

Table 5: Evaluation results on entity prediction in zero-shot scenario

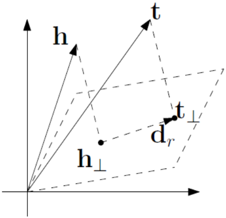
Metric	$d-e$	$e-d$	$d-d$	Total
Partial-CBOW	26.5	20.9	67.2	24.6
CBOW	27.1	21.7	66.6	25.3
Partial-CNN	26.8	20.8	69.5	24.8
CNN	**31.2**	**26.1**	**72.5**	**29.5**

Table 6: Evaluation results on relation prediction in zero-shot scenario

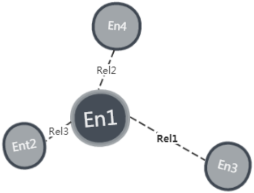
Metric	$d-e$	$e-d$	$d-d$	Total
Partial-CBOW	49.0	42.2	0.0	46.2
CBOW	52.2	47.9	0.0	50.3
Partial-CNN	56.6	52.4	4.0	54.8
CNN	**60.4**	**55.5**	**7.3**	**58.2**

Two basic assumptions

1. If a relation is important to an entity, most of its connections will be connected through that relation
2. If an entity is important to a relation, the relation can be well defined by the entity.

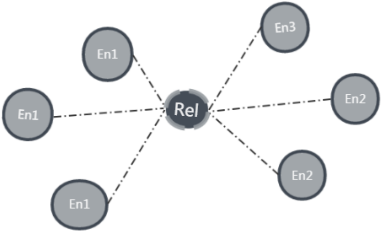


Relation Importance to Entity



$$\alpha_j = \frac{En_j \cdot Rel_i}{\sum_{i=1}^{k} En_j \cdot Rel_i} \tag{12}$$

Entity Importance to Relation



$$\alpha_j = \frac{En_i \cdot Rel_j}{\sum_{i=1}^{k} En_i \cdot Rel_j} \tag{13}$$

Objective function and gradient update

Objective function

$$f_r(h, r, t) = ||r^T hr + d_r - r^T tr|| \tag{14}$$

$$\mathcal{L} = \sum_{(h,r,t) \in T} \sum_{(h',r',t')} [f_r(h, t) + \gamma - f_r(h', t')]_+ \tag{15}$$

gradient update

$$h = h - \alpha_{(h,r)} \frac{\partial L}{\partial h} \tag{16}$$

$$r = r - \alpha_{(h,r)} \frac{\partial L}{\partial r} \tag{17}$$