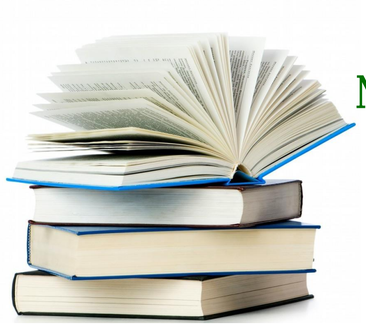
Survey on Information Extraction & Knowledge Base Reasoning

from Papers in Top Conference 2016~2017

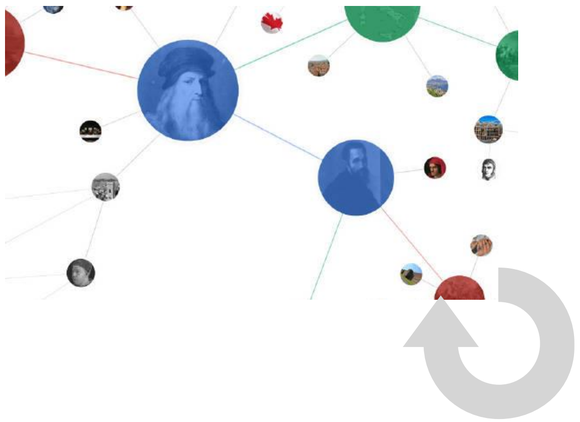
Tianwen Jiang

Outline



Named Entity Recognition
Entity Linking

Relation Classification
ERE Extraction
KB Population

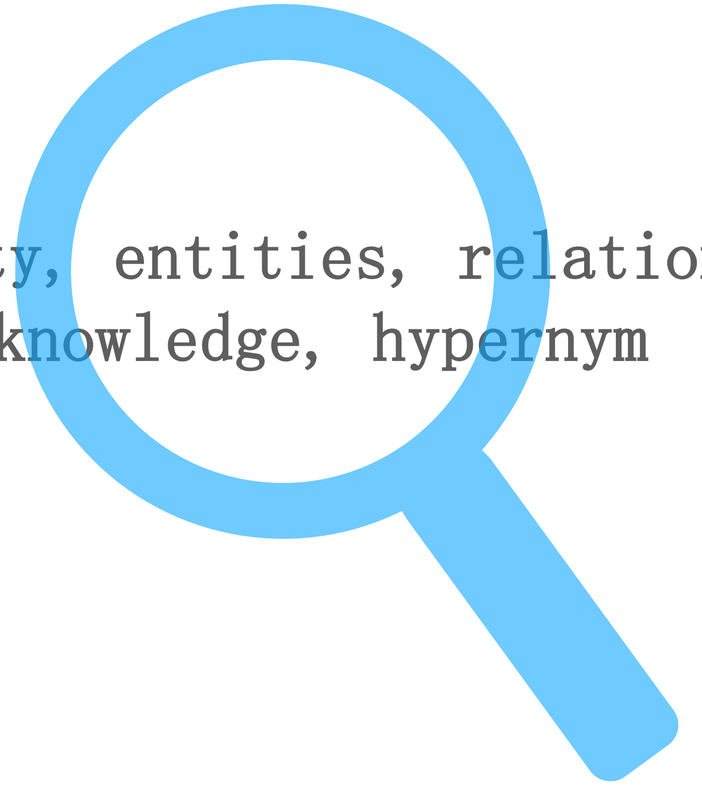


KB Completion

KB Embeddings

KB Reasoning

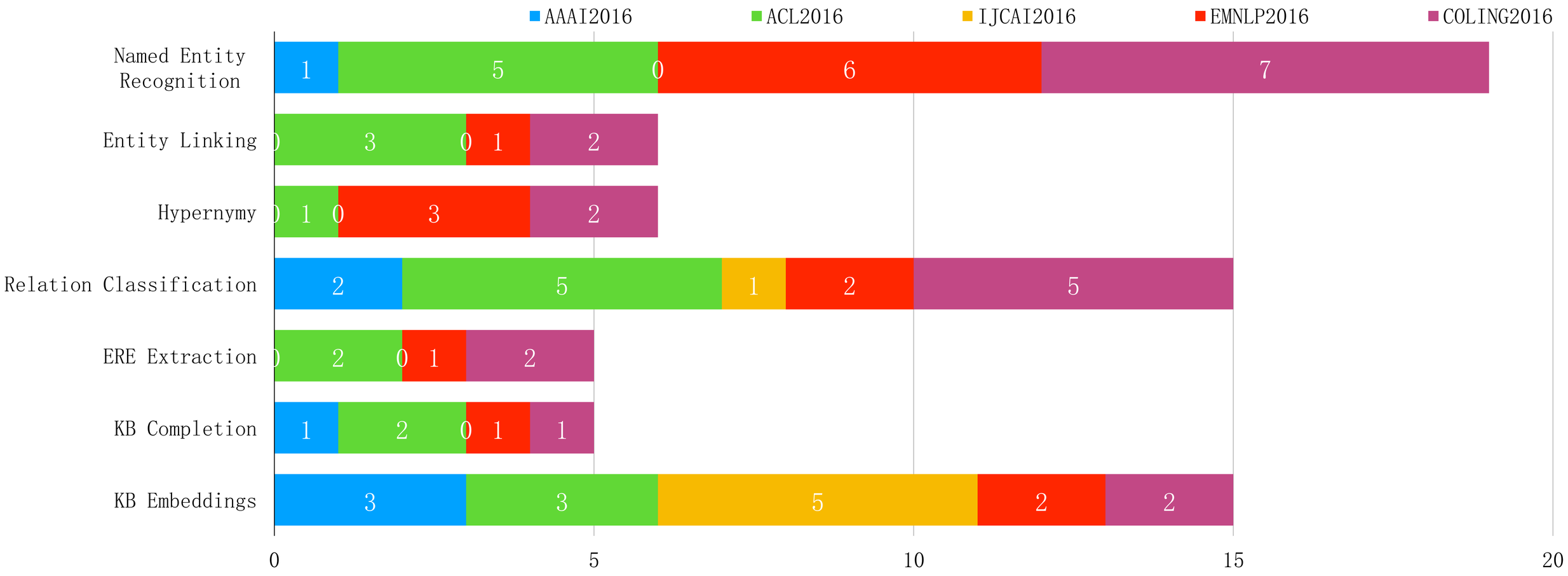
AAAI & ACL & IJCAI & EMNLP & COLING



entity, entities, relation,
knowledge, hypernym

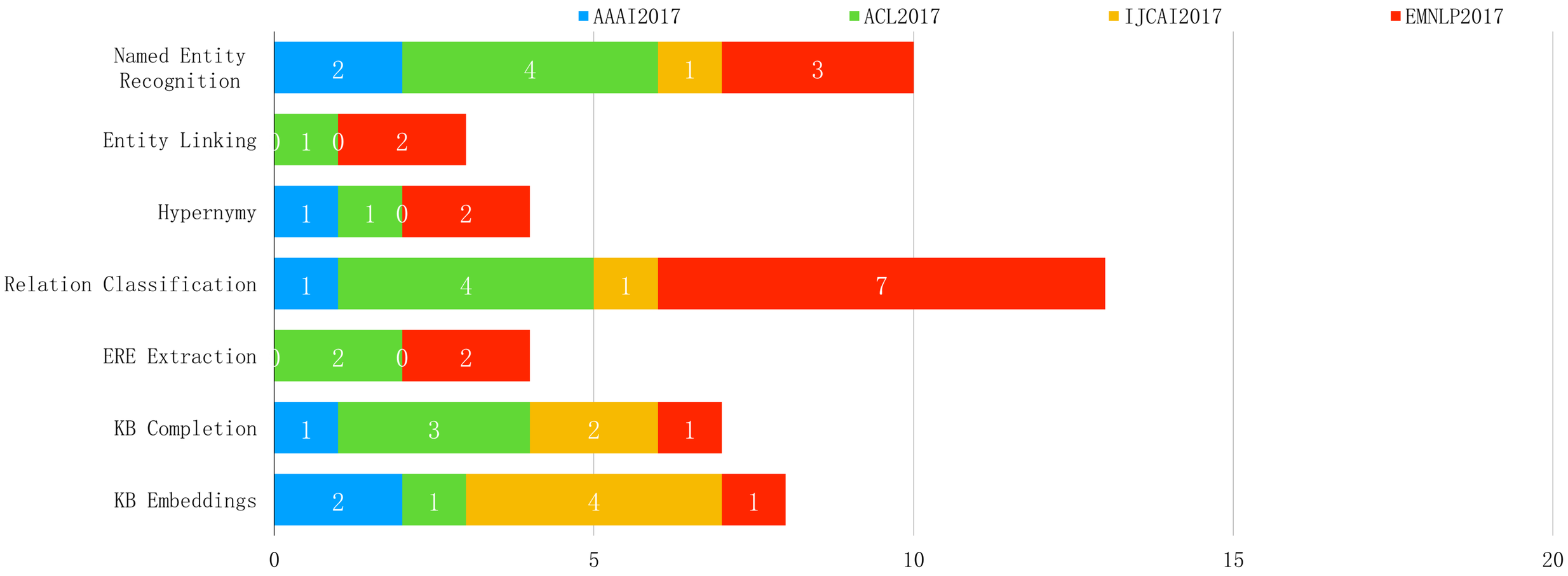
AAAI & ACL & IJCAI & EMNLP & COLING

2016



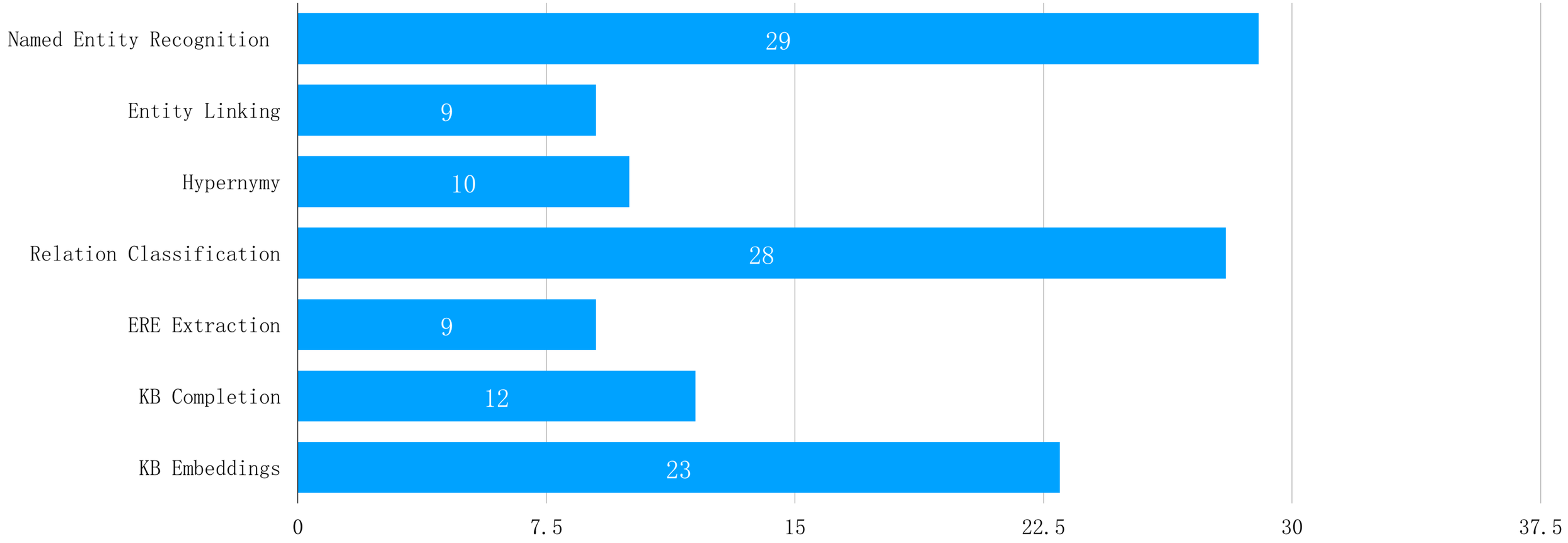
AAAI & ACL & IJCAI & EMNLP

2017



total

2016～2017



Named Entity Recognition

Research Focus

- domain-specific
- discontiguous entities
- fine-grained entity typing
- classification approach
- multilingual
- speed of recognition
- word segmentation
- new word discovery

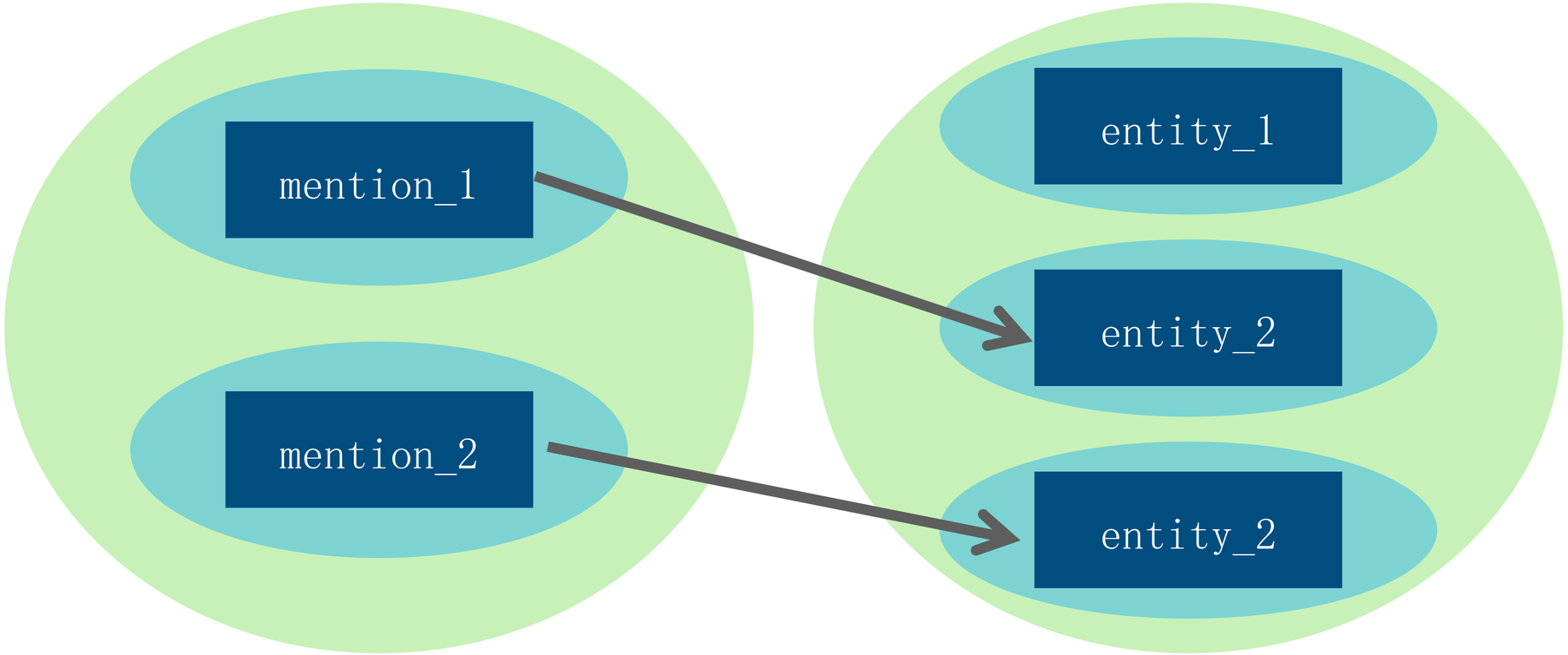
Entity Linking

Research Focus

- context information
- collective entity resolution
- using global information
- short text
- little known referents
- list only entity linking

Entity Linking

The goal of Named Entity Disambiguation (NED) is to link each mention of named entities in a document to a knowledge-base of instances.



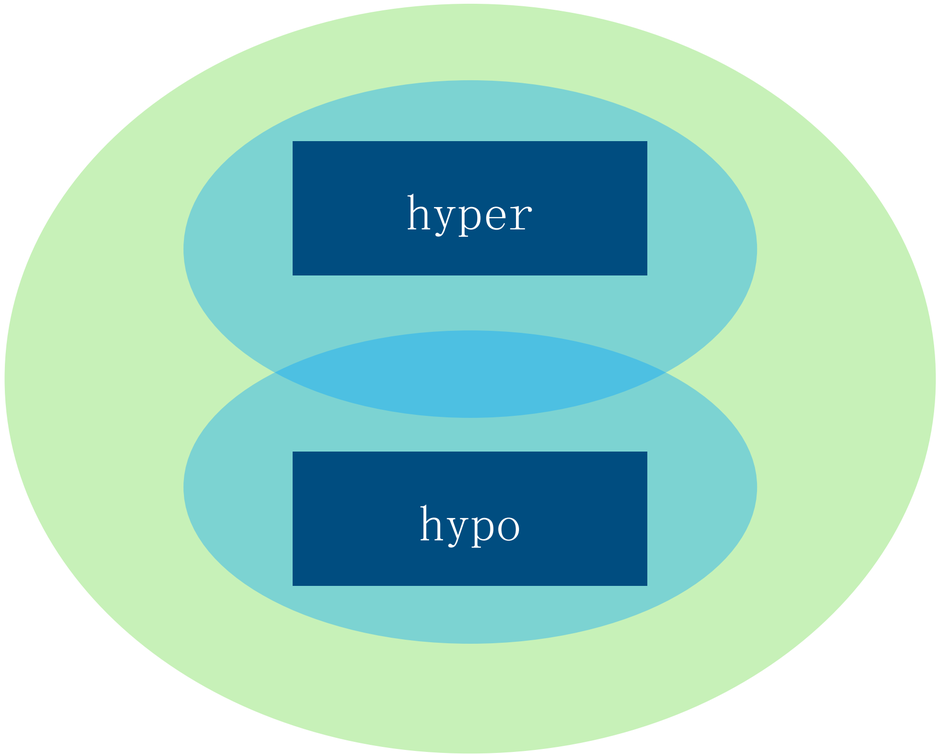
Hypernymy Detection / Discovery

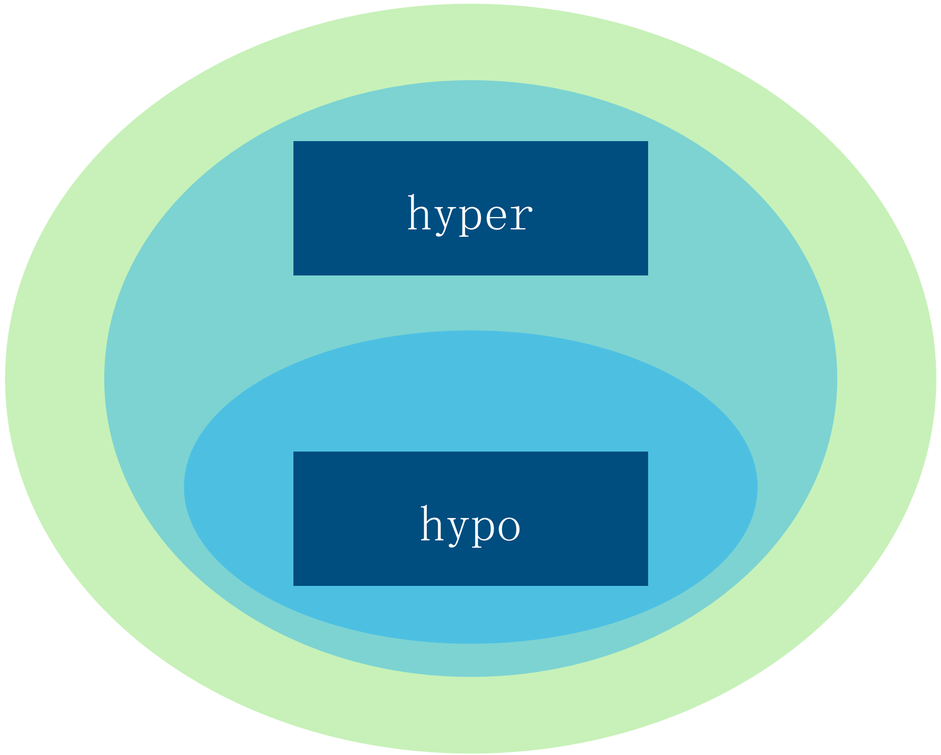
Research Focus

- statistical and linguistic
- distributional inclusion hypothesis
- distributional
- problem of lexical memorization
- domain adaptation
- categories resources
- weakly supervised
- generation

Hypernymy Detection

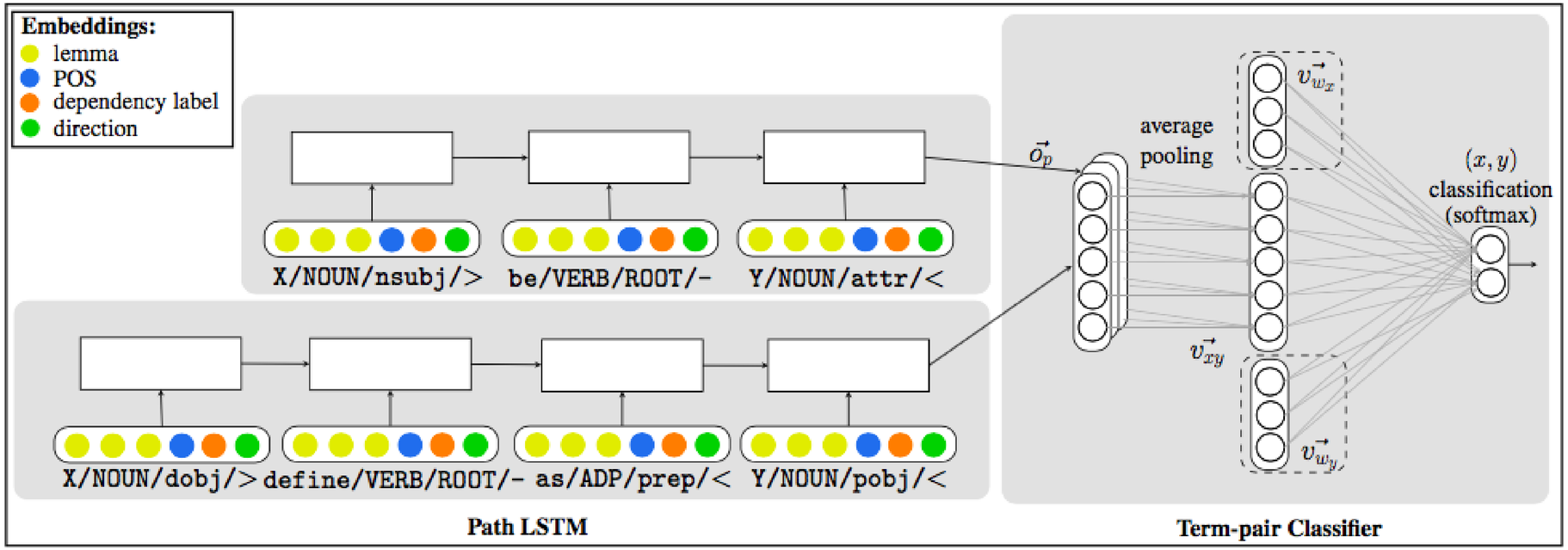
Given a pair of words, detecting whether there is a hypernymy between them.





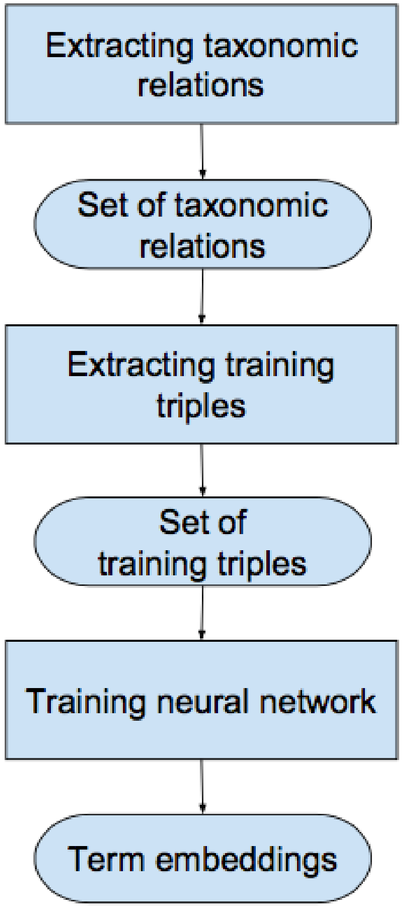
Hypernymy Detection . distributional

integrated path-based and distributional method

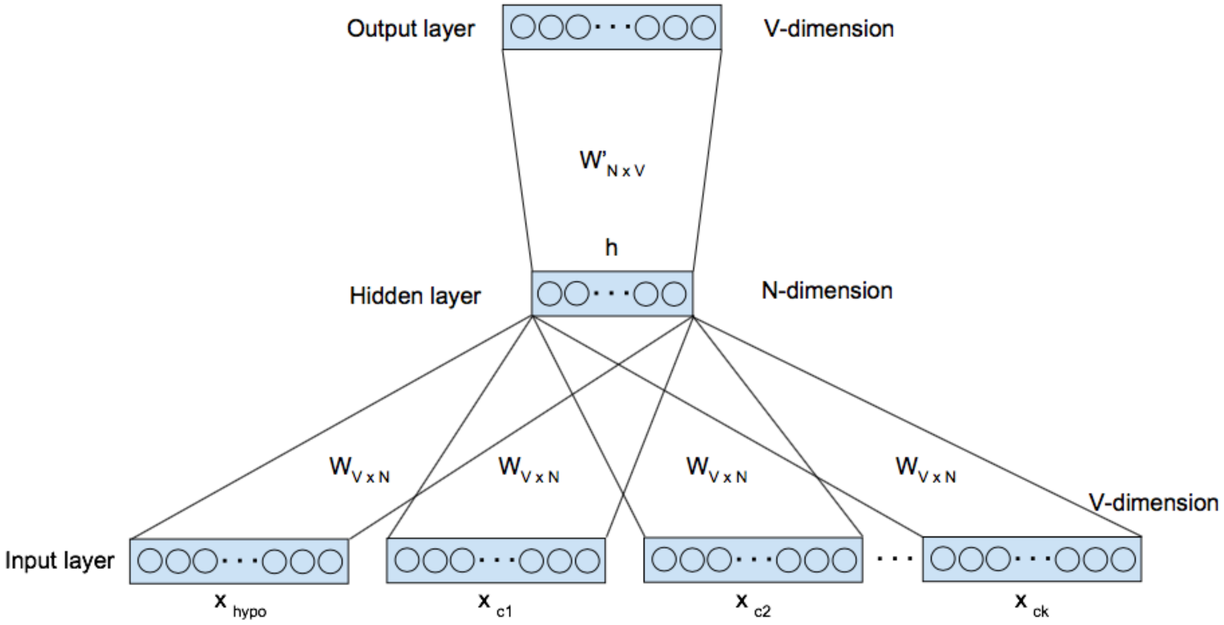


Vered Shwartz et al., 2016

Hypernymy Detection . distributional



dynamic weighting neural network model



Luu Anh Tuan et al., 2016

Hypernymy Detection . distributional

domain adaptation

Domain Clustering

Training Data Expansion

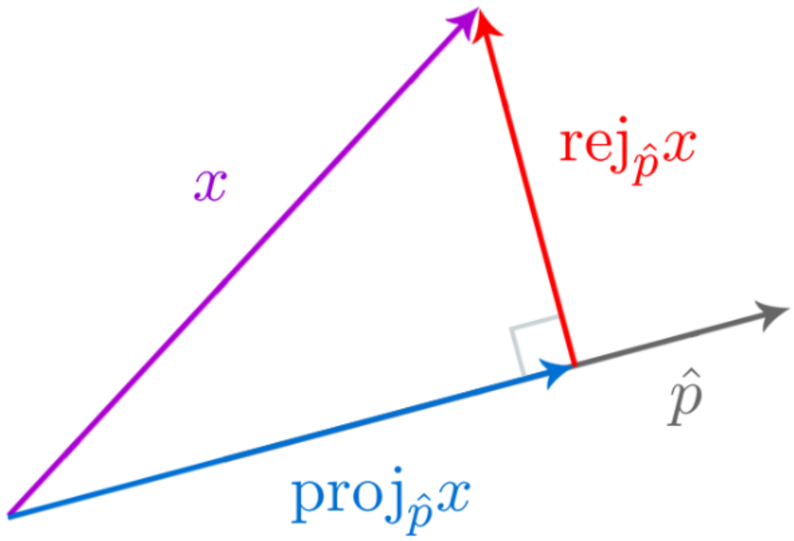
Learning a Hypernym Detection Matrix

$$\hat{d}(b) = \max_{d \in D} WO(\vec{d}, \vec{b})$$

$$\min_{\Psi^C} \sum_{i=1}^{|T^d|} ||\Psi^C \vec{x}_i^d - \vec{y}_i^d||^2$$

Luis Espinosa-Anke et al., 2016

Hypernymy Detection . problem of lexical memorization



Stephen Roller et al., 2016

Relation Extraction

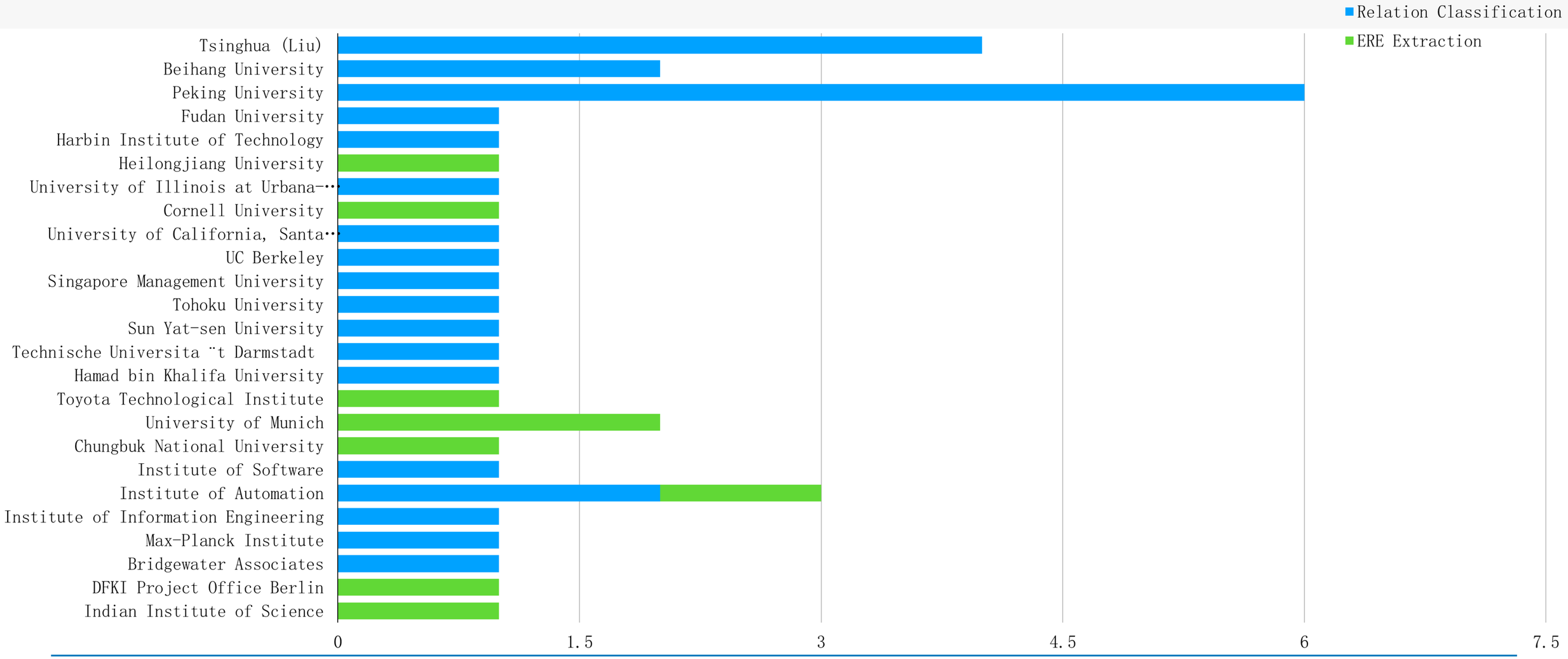
37 papers from 25 organization

Tsinghua (Liu)
Beihang University
Peking University
Fudan University
Harbin Institute of Technology
Heilongjiang University
Sun Yat-sen University

University of Illinois at Urbana-Champaign
Cornell University
University of California, Santa Barbara
UC Berkeley
Singapore Management University
Tohoku University
Chungbuk National University
Technische Universita ¨t Darmstadt
Hamad bin Khalifa University
Toyota Technological Institute
University of Munich

Institute of Software
Institute of Automation
Institute of Information Engineering
Max-Planck Institute
Bridgewater Associates
DFKI Project Office Berlin
Indian Institute of Science

Relation Extraction



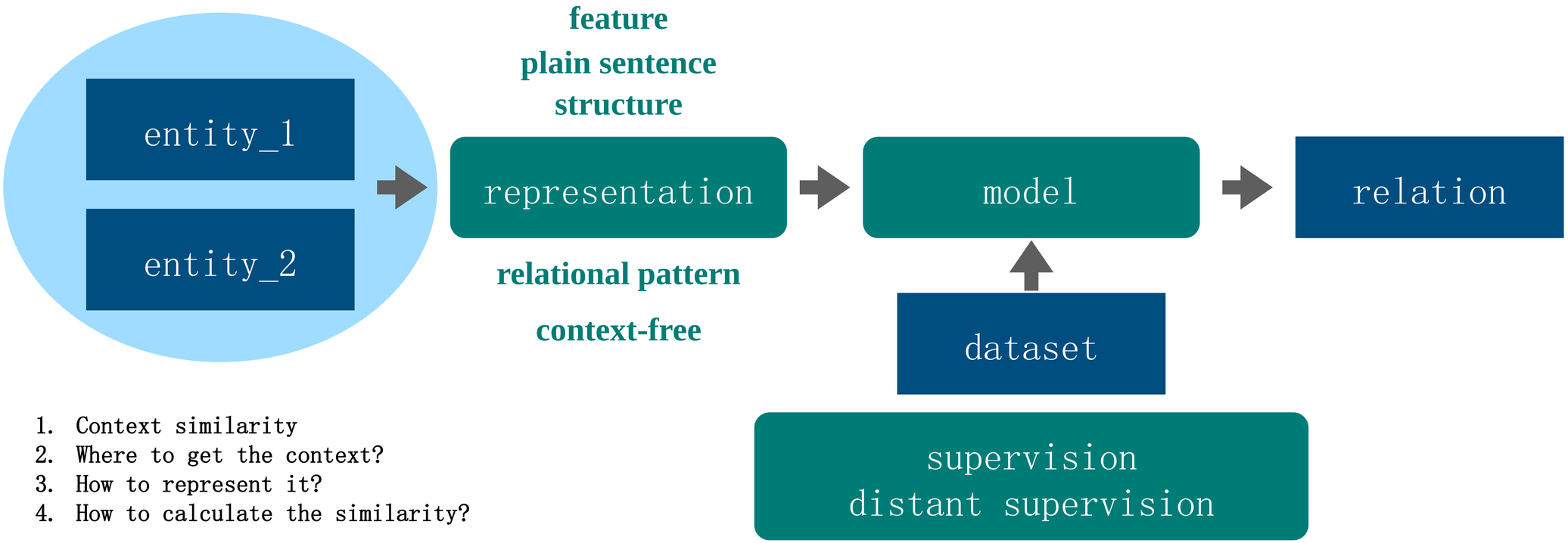
Relation Extraction

Relation extraction? Relation classification?

1. Pipeline?
2. Need classification?
3. Focus precision or recall?

Relation Classification

Given a pair of entities, classifying the relation type between them, usually, with sentence instances containing the two entities.



1. Context similarity
2. Where to get the context?
3. How to represent it?
4. How to calculate the similarity?

Relation Classification

ISSUES:

1. Where is entity pair? (POSITION) how?
2. The direction of relation? (DUAL) cons?
3. RNN or CNN?

Relation Classification

the task of extracting relational facts from text

- learning that a person is employed by a particular organization,
- learning that a geographic entity is located in a particular region.
- …

supervised approaches

Sentences in a corpus are first hand-labeled for the presence of entities and the relations between them

unsupervised approaches

Extracting strings of words between entities in large amounts of text, and clusters and simplifies these word strings to produce relation-strings

bootstrap learning

Using a very small number of seed instances or patterns to do bootstrap learning

distant supervision

Any sentence that contains a pair of entities that participate in a known Freebase relation is likely to express that relation in some way.

Relation Classification

	# relation	# in train	# in test
SemEval-2010 Task 8 dataset	9+1	8,000 (sentences)	2,717 (sentences)
NYT 2010	52+1	522,611 (sentences) 281,270 (entity-pairs) 18,252 (relational facts)	172,448 (sentences) 96,678 (entity-pairs) 1,950 (relational facts)

SemEval-2010 Task 8 dataset

- ✓ **Cause-Effect**
- ✓ **Component-Whole**
- ✓ **Content-Container**
- ✓ **Entity-Destination**
- ✓ **Entity-Origin**
- ✓ **Message-Topic**
- ✓ **Member-Collection**
- ✓ **Instrument-Agency**
- ✓ **Product-Agency**
- ✓ **Other**

Hendrickx I, Kim S N, Kozareva Z, et al. Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals[C]//Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions. Association for Computational Linguistics, 2009: 94-99.

Riedel S, Yao L, McCallum A. Modeling relations and their mentions without labeled text[J]. Machine learning and knowledge discovery in databases, 2010: 148-163.

Relation Classification

Research Focus

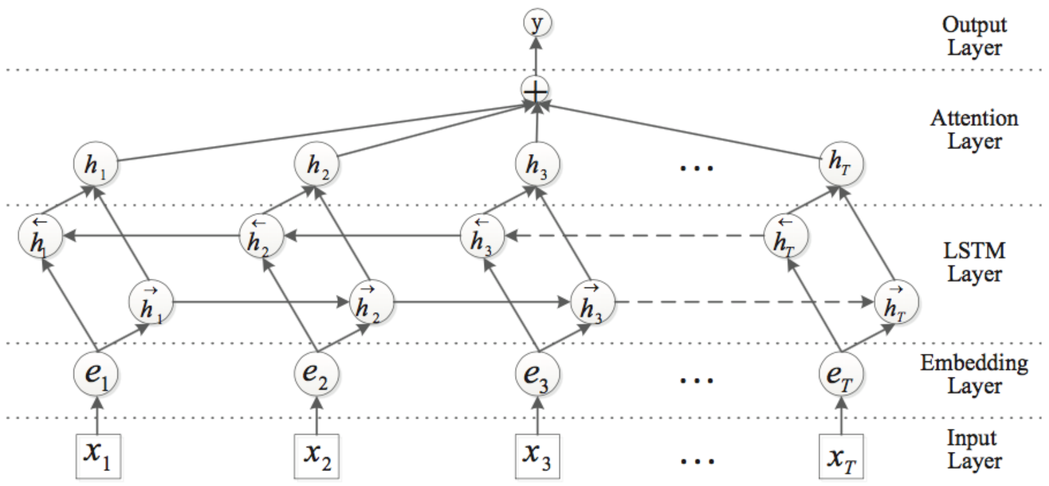
traditional classification

relational phrase

distant supervision

multi-lingual

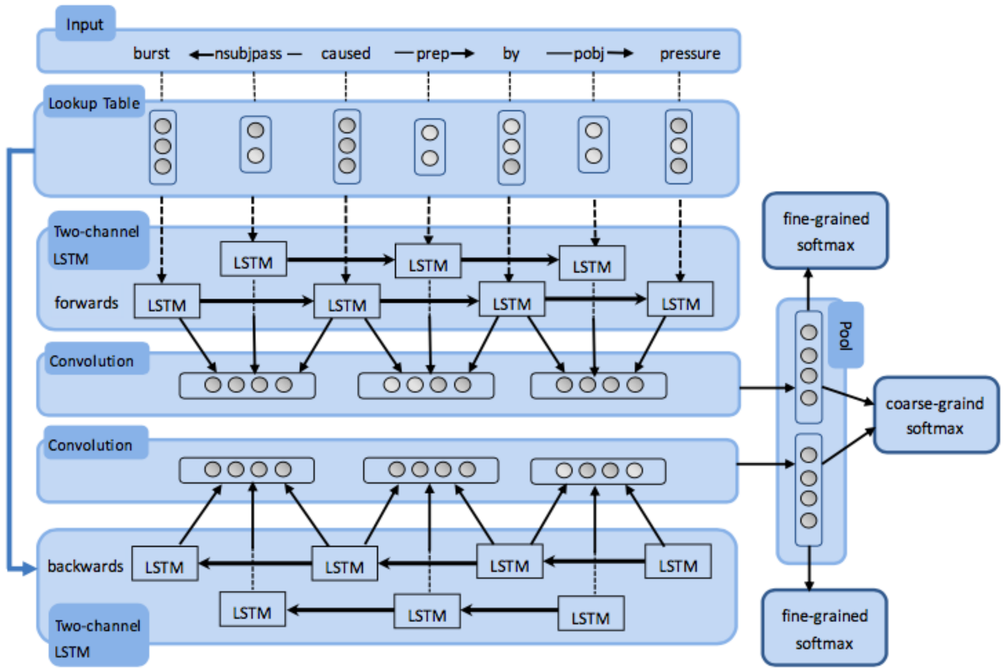
Relation Classification . traditional classification



Peng Zhou et al., 2016

Institute of Automation

SemEval-2010
F1=84.0



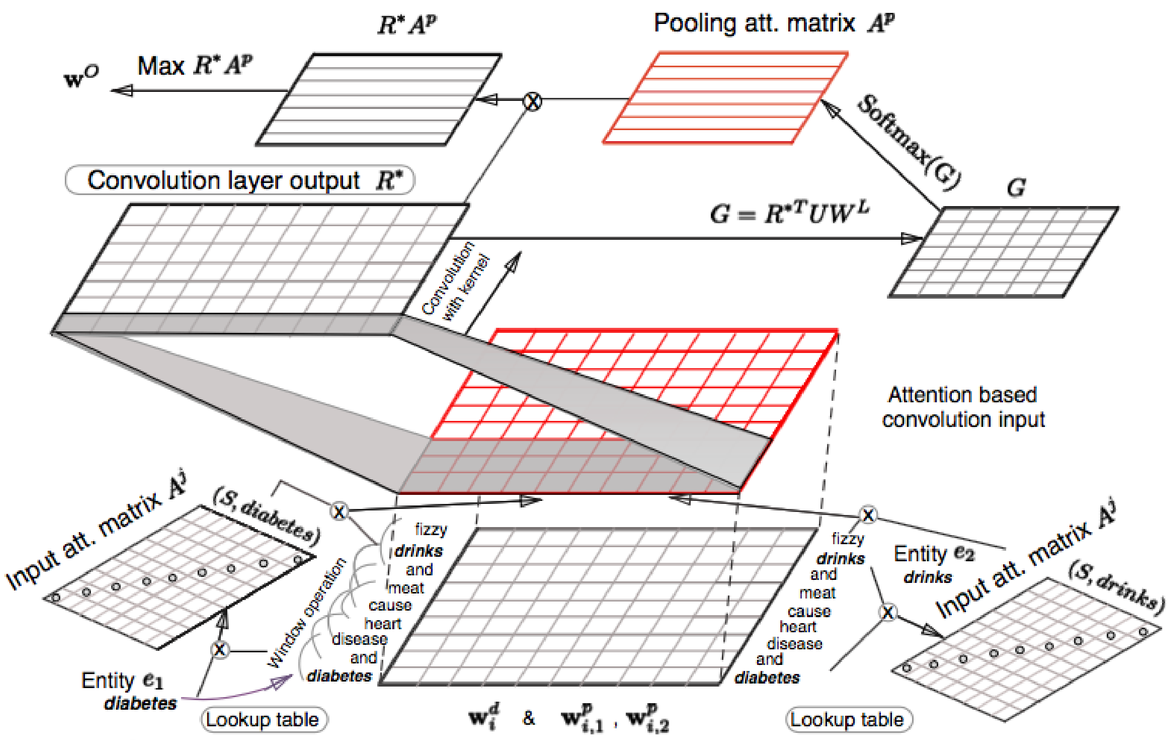
Rui Cai et al., 2016

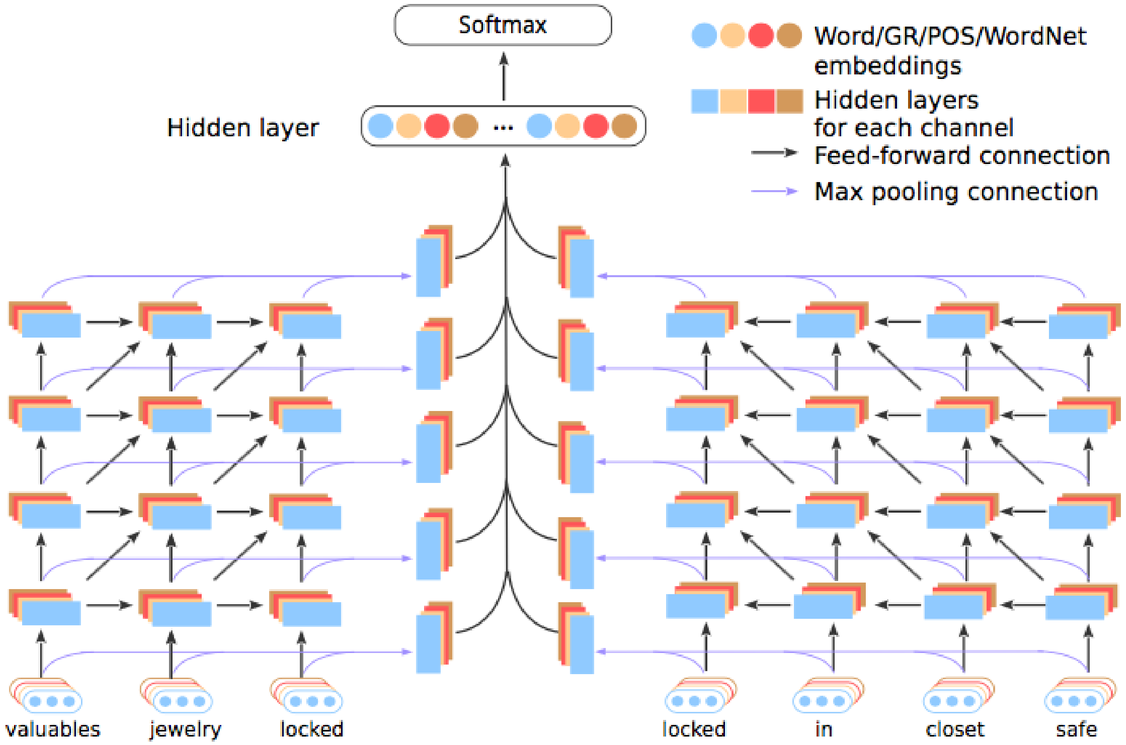
Peking University

SemEval-2010
F1=85.4

Model	F_1
CNN	81.8
LSTM	76.6
Two-channel LSTM	81.5
RCNN	**82.4**

Relation Classification . traditional classification





Linlin Wang et al., 2016

Tsinghua University

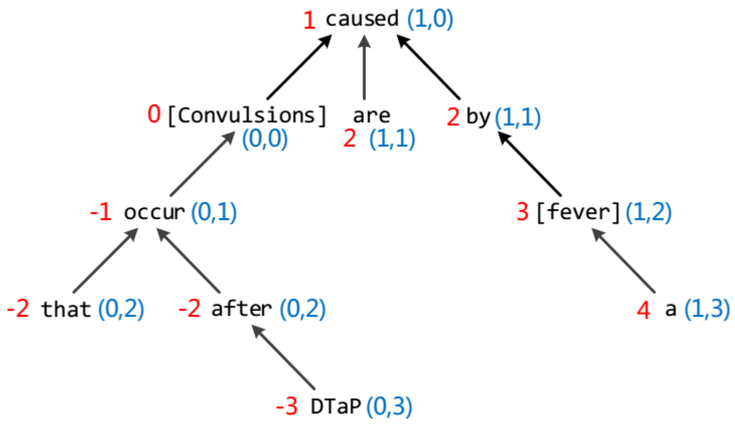
SemEval-2010
F1=88.0

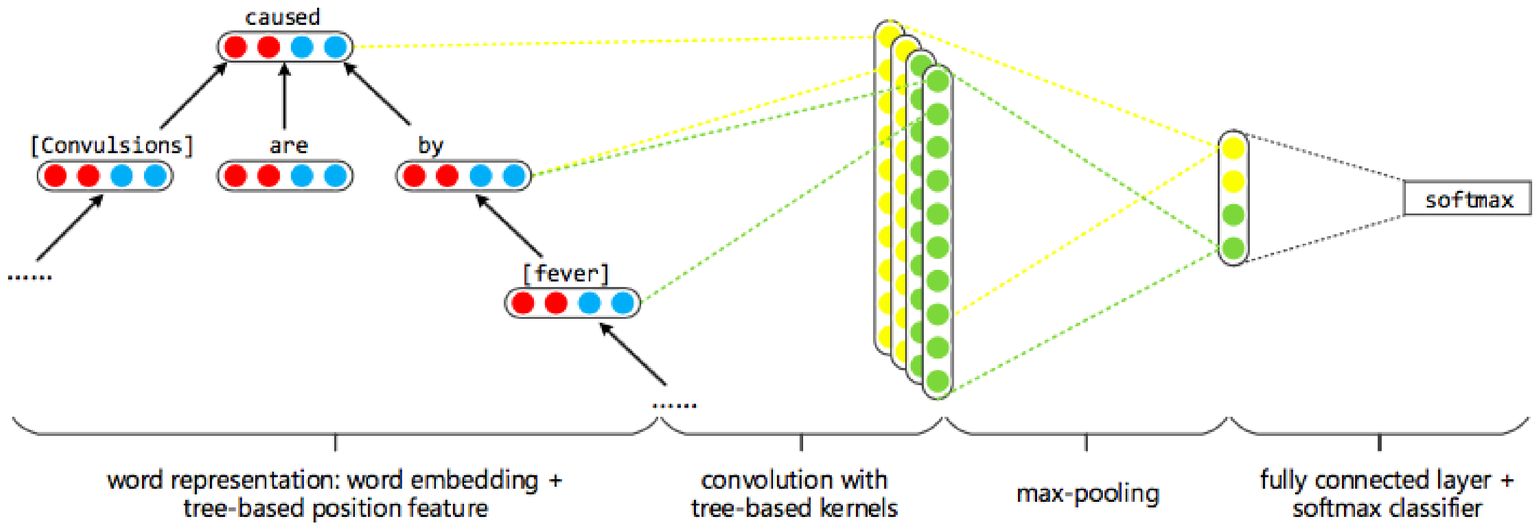
Yan Xu et al., 2016

Peking University

SemEval-2010
F1=86.1

Relation Classification . traditional classification



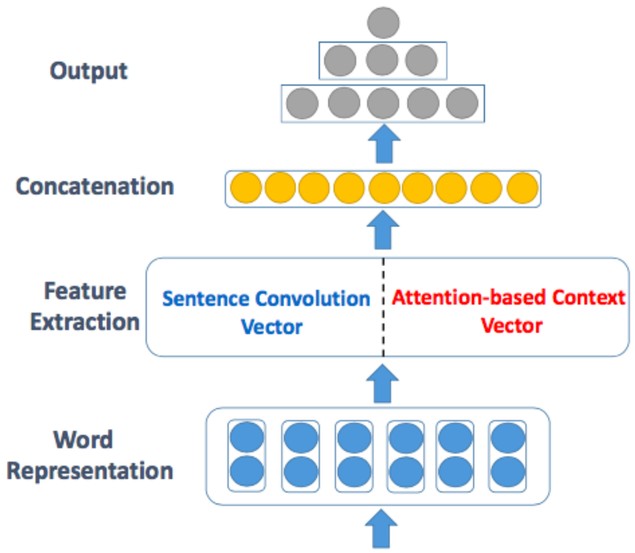


Yunlun Yang et al., 2016

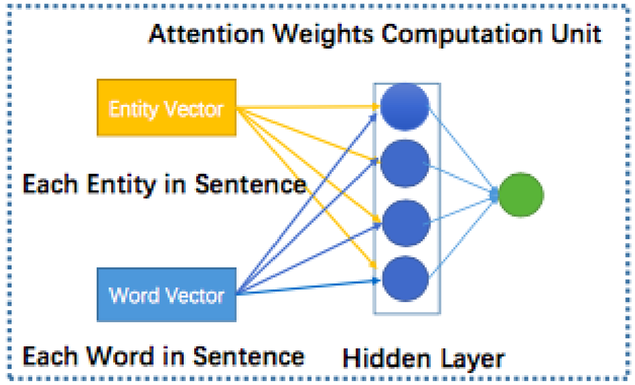
Peking University

SemEval-2010
F1=84.6

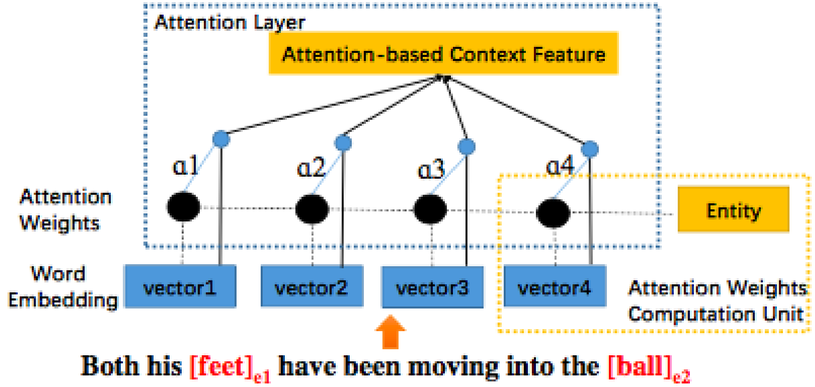
Relation Classification . traditional classification



Both his [feet]$_{e1}$ have been moving into the [ball]$_{e2}$



(a) Attention weights computation unit.



Both his [feet]$_{e1}$ have been moving into the [ball]$_{e2}$

(b) Attention layer network.

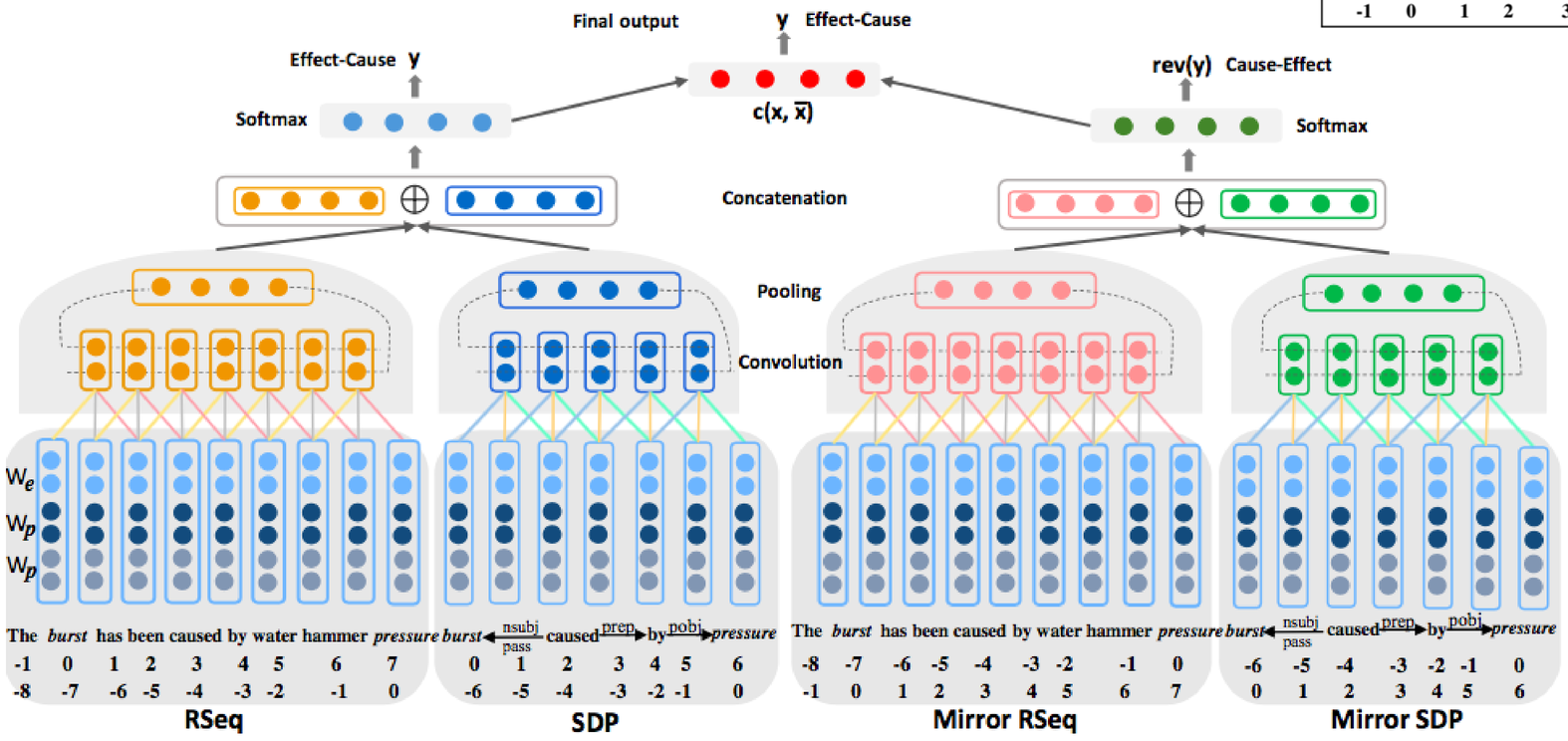
Yatian Shen et al., 2016

Fudan University

SemEval-2010
F1=84.3

Relation Classification . traditional classification

Raw Sequence (RSeq)	Shortest Dependency Path (SDP)	Label
The $[burst]_{e_1}$ has been caused by water hammer $[pressure]_{e_2}$ -1 0 1 2 3 4 5 6 7 -8 -7 -6 -5 -4 -3 -2 -1 0	$[burst]_{e_1} \xleftarrow{nsubjpass}$ caused $\xrightarrow{prep}$ by $\xrightarrow{pobj} [pressure]_{e_2}$ 0 1 2 3 4 5 6 -6 -5 -4 -3 -2 -1 0	Effect-Cause
The $[burst]_{e_2}$ has been caused by water hammer $[pressure]_{e_1}$ -8 -7 -6 -5 -4 -3 -2 -1 0 -1 0 1 2 3 4 5 6 7	$[burst]_{e_2} \xleftarrow{nsubjpass}$ caused $\xrightarrow{prep}$ by $\xrightarrow{pobj} [pressure]_{e_1}$ -6 -5 -4 -3 -2 -1 0 0 1 2 3 4 5 6	Cause-Effect

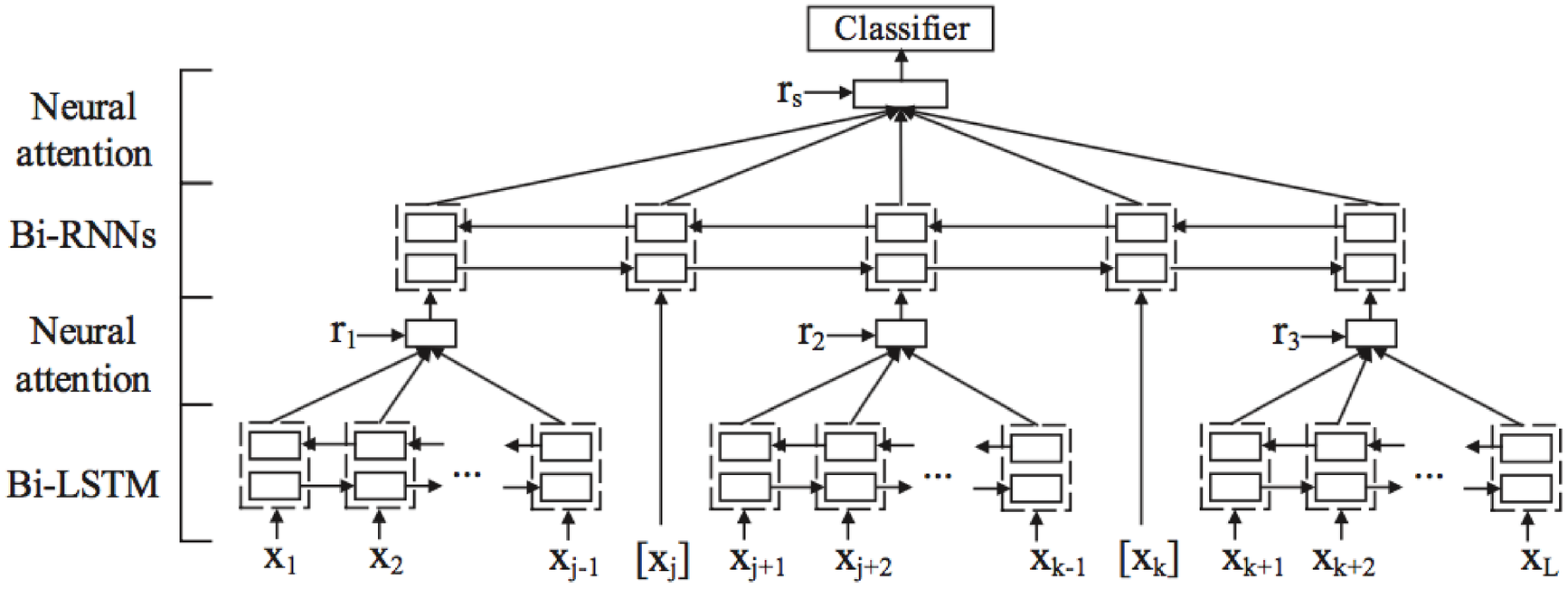


Jianfei Yu et al., 2016

Singapore Management University

SemEval-2010
F1=85.0

Relation Classification . traditional classification



Minguang Xiao et al., 2016

Sun Yat-sen University

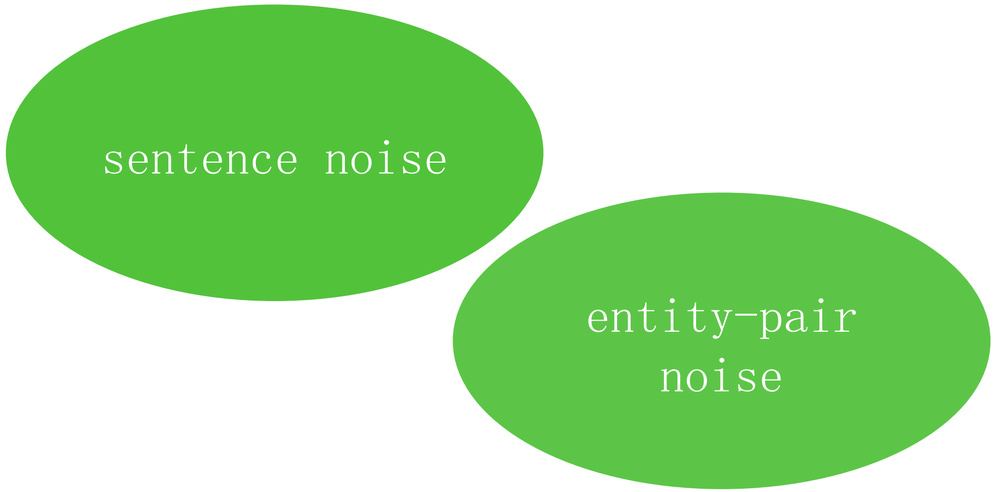
SemEval-2010
F1=83.9

Relation Classification . distant supervision

automatically generate training data via aligning KBs and texts

- a sentence that mentions two entities participating in a relation in knowledge bases would not necessarily express this relation explicitly
- there may be multiple relations holding for the same pair of entities
- Knowledge base incompletion
- entity-level learning (ignoring detailed sentence-level information, and could not deal with the multi-label problem.)
- sentence-level learning (ignoring relations also can be mentioned across sentences)

- all sentences express that relation.
- at-least-one multi-instance learning
- multi-instance multi-label learning



Relation Classification . distant supervision

IDEAS for MIML:

1. Scoring for relations
2. Attention mechanism
3. Pooling mechanism
4. Noise modeling
5. Class-tie
6. Heterogeneous supervision

Relation Classification . distant supervision

Algorithm 1 The learning algorithm for RankRE-local

Input:
The training dataset $D = \{(X^q, S^q) | q = 1, \ldots, M\}$ consisting of (instance set, relation set) pairs.
A parameter T specifying the number of iterations over the training set.

Output:
The parameter vector $\mathbf{w}$.

For convenience, we define the summed feature vector $\mathbf{\Psi}_r^q = \sum_{x \in X^q} \mathbf{\Phi}(x, r)$
initialize parameter vector $\mathbf{w} \leftarrow \mathbf{0}$
for $t = 1, \ldots, T$ **do**
  for $q = 1, \ldots, M$ **do**
    $m \leftarrow |S^q|$
    $S_\star^q \leftarrow \text{top\_}m(X^q)$ [1]
    for each $r \in S^q$ **do**
      $\mathbf{w} \leftarrow \mathbf{w} + \mathbf{\Psi}_r^q$
    end for
    for each $r \in S_\star^q$ **do**
      $\mathbf{w} \leftarrow \mathbf{w} - \mathbf{\Psi}_r^q$
    end for
  end for
end for
return $\mathbf{w}$

Hao Zheng et al., 2016

Beihang University

multi-instance multi-label learning framework

$$f_i = \sum_{j=1}^{n} f(x_j, r_i) \qquad f(x_j, r_i) = \mathbf{w}^T \mathbf{\Phi}(x_j, r_i)$$

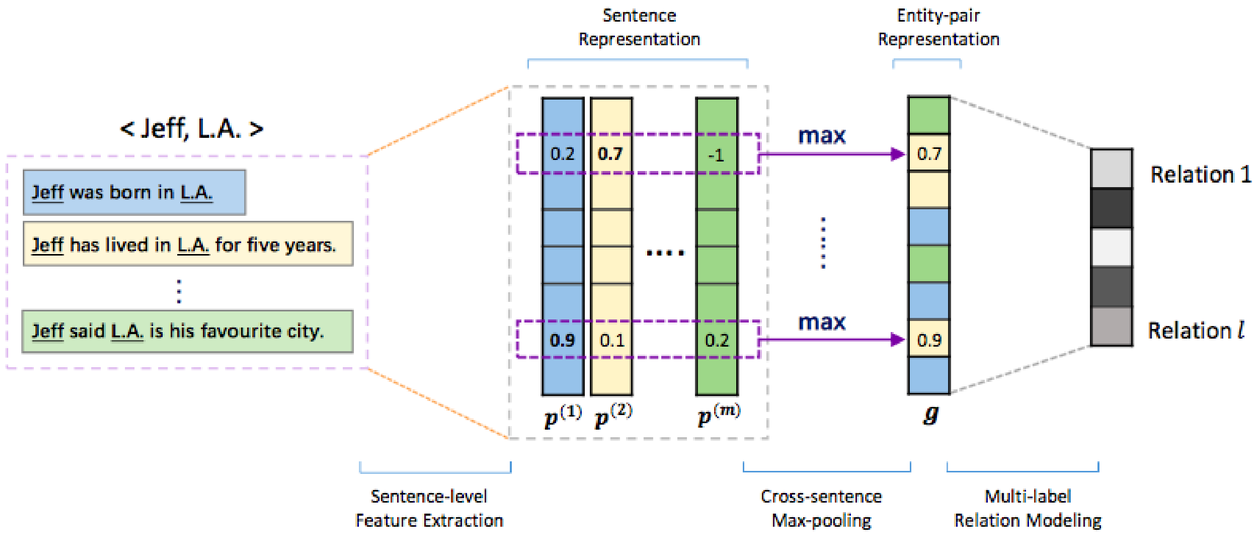
$$P_m(r_{l_1}, \ldots, r_{l_m}) = \frac{e^{\sum_{i=1}^{m} f_{l_i}}}{\sum_{r_{s_1}, r_{s_2}, \ldots, r_{s_m} \in R} e^{\sum_{i=1}^{m} f_{s_i}}}$$

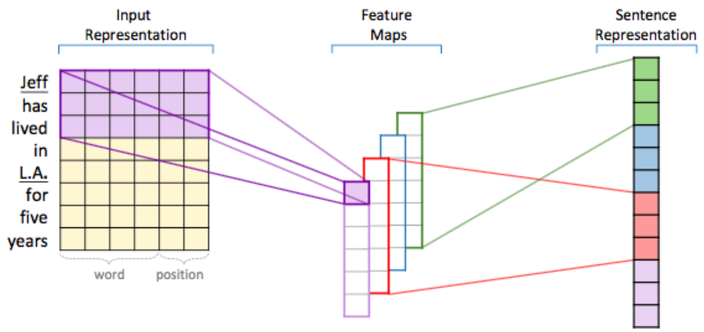
KBP

Top-N	Top-50	Top-100	Top-150
Mintz++	0.540	0.410	0.333
MIML-RE	0.500	0.470	0.406
RankRE-local	**0.580**	**0.490**	**0.426**
RankRE-global	**0.560**	**0.490**	**0.433**

Relation Classification . distant supervision

multi-instance multi-label learning framework





	Mintz++	MultiR	MIMLRE	PCNN	MIMLCNN
P@10	0.70	0.80	0.60	0.84	**0.90**
P@20	0.65	0.65	0.70	0.80	**0.83**
P@30	0.60	0.63	0.63	0.76	**0.80**
P@50	0.54	0.62	0.68	0.72	**0.75**
P@100	0.53	0.62	0.68	0.68	**0.69**
P@200	0.51	0.63	**0.64**	0.62	**0.64**
P@300	0.49	**0.63**	0.62	0.58	0.59
P@500	0.42	0.48	0.51	**0.53**	**0.53**
Mean	0.56	0.63	0.63	0.69	**0.72**

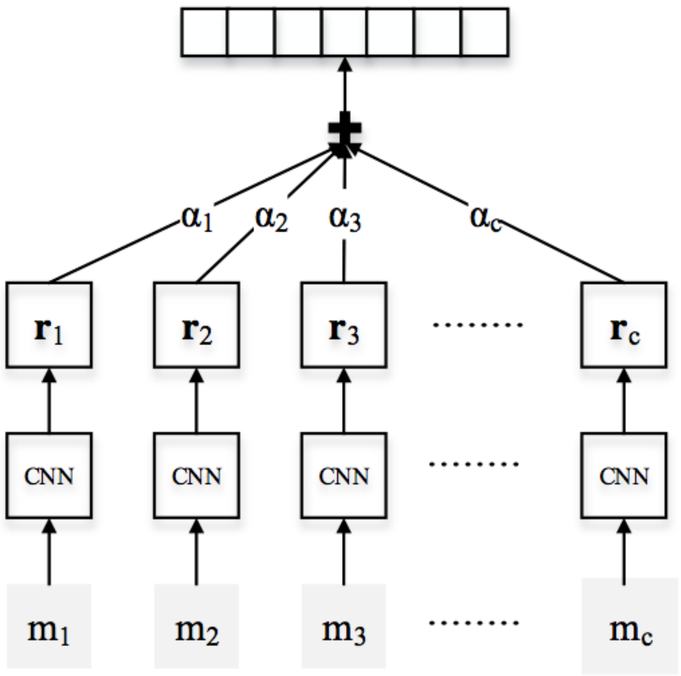
NYT

Xiaotian Jiang et al., 2016

Institute of Information Engineering

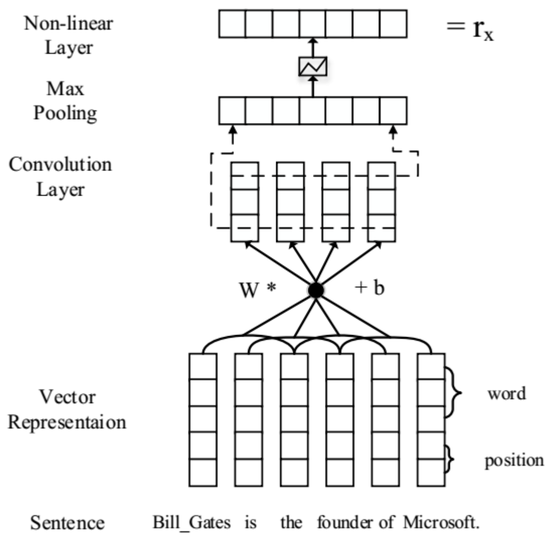
Relation Classification . distant supervision

multi-instance multi-label learning framework

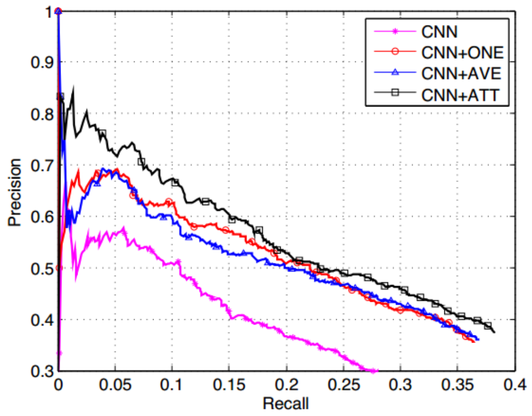


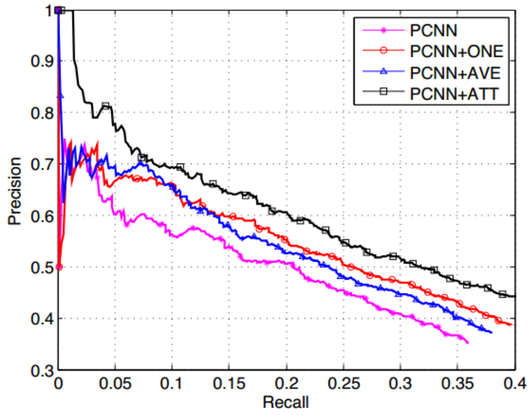
Yankai Lin et al., 2016

Tsinghua University



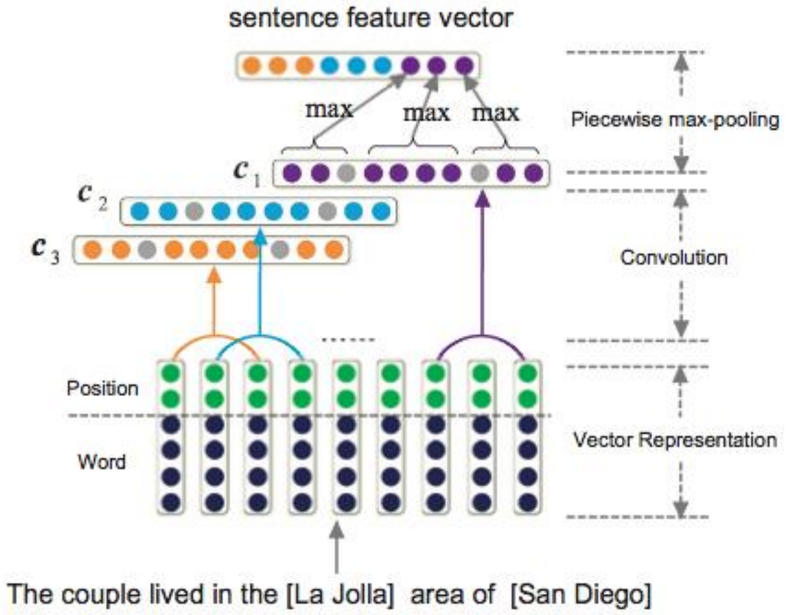
P@N(%)	100	200	300	Mean
CNN+ONE	67.3	64.7	58.1	63.4
+AVE	64.4	60.2	60.1	60.4
+ATT	**76.2**	68.6	59.8	68.2
PCNN+ONE	72.3	69.7	64.1	68.7
+AVE	73.3	66.7	62.8	67.6
+ATT	**76.2**	**73.1**	**67.4**	**72.2**



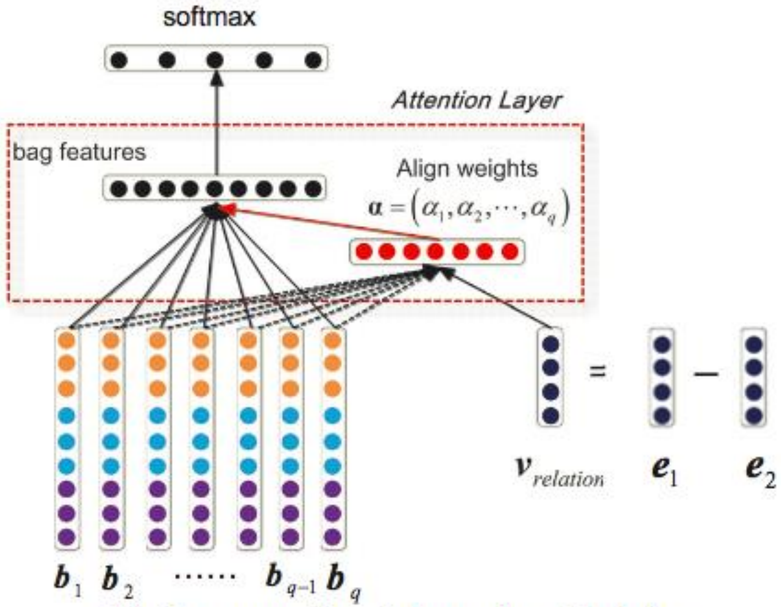


Relation Classification . distant supervision

multi-instance multi-label learning framework



(a) PCNNs Module



(b) Sentence-level Attention Module

$$\mathcal{L}_e = \sum_{i=1}^{|\mathcal{D}|} \| e_i - d_i \|_2^2$$

NYT

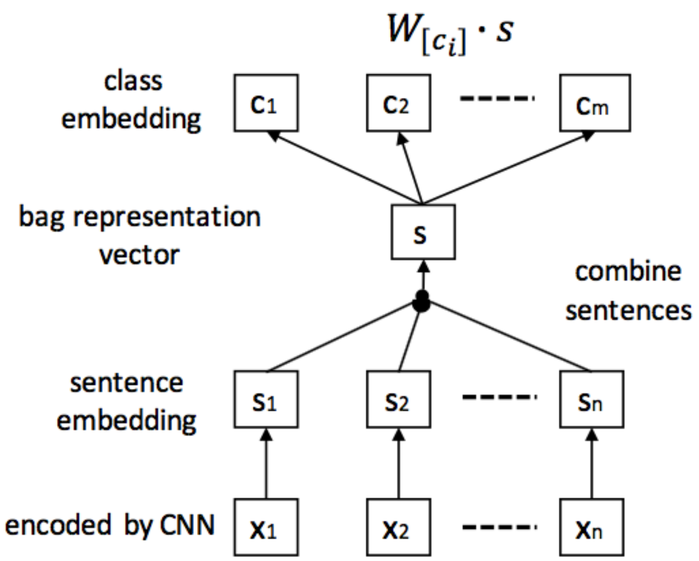
Accuracy (%)	Top 100	Top 200	Top 500	Average
Mintz	0.77	0.71	0.55	0.676
MultiR	0.83	0.74	0.59	0.720
MIML	0.85	0.75	0.61	0.737
PCNNs+MIL	0.86	0.80	0.69	0.783
APCNNs	**0.87**	0.82	0.72	0.803
PCNNs+MIL+D	0.86	0.82	0.71	0.797
APCNNs+D	**0.87**	**0.83**	**0.74**	**0.813**

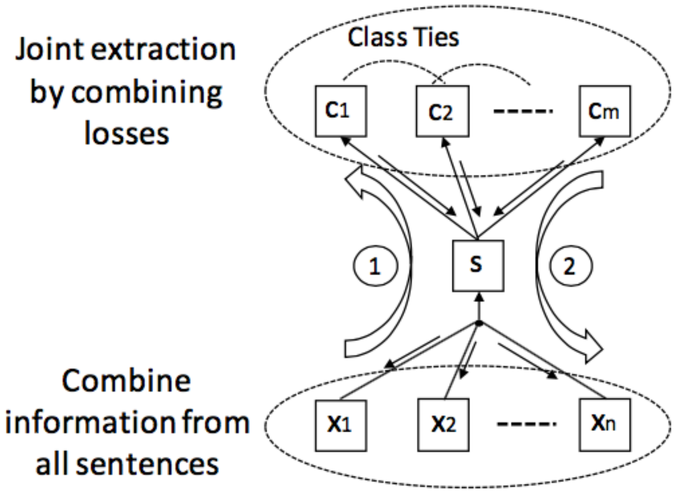
Guoliang Ji et al., 2017

Institute of Automation

Relation Classification . distant supervision

MIML; class ties

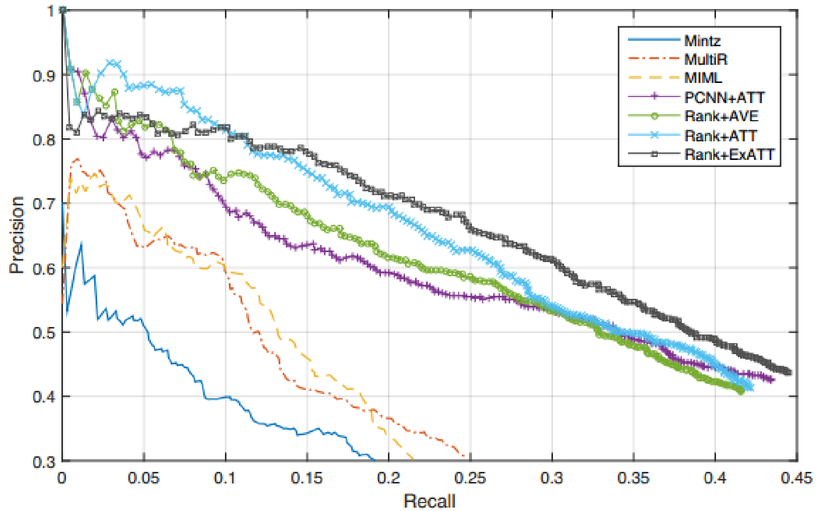




$$\mathcal{F}(s,c) = W_{[c]} \cdot s$$

$$\forall c^+ \in L_k, \forall c^- \in \mathcal{L} - L_k : \mathcal{F}(s,c^+) > \mathcal{F}(s,c^-) + \beta$$

$$G_{[att]} = \sum_{c^+ \in L_k} (\rho[0, \sigma^+ - \mathcal{F}(s^{c^+}, c^+)]_+ + \rho[0, \sigma^- + \mathcal{F}(s^{c^+}, c^-)]_+)$$



Hai Ye et al., 2017

Beihang University

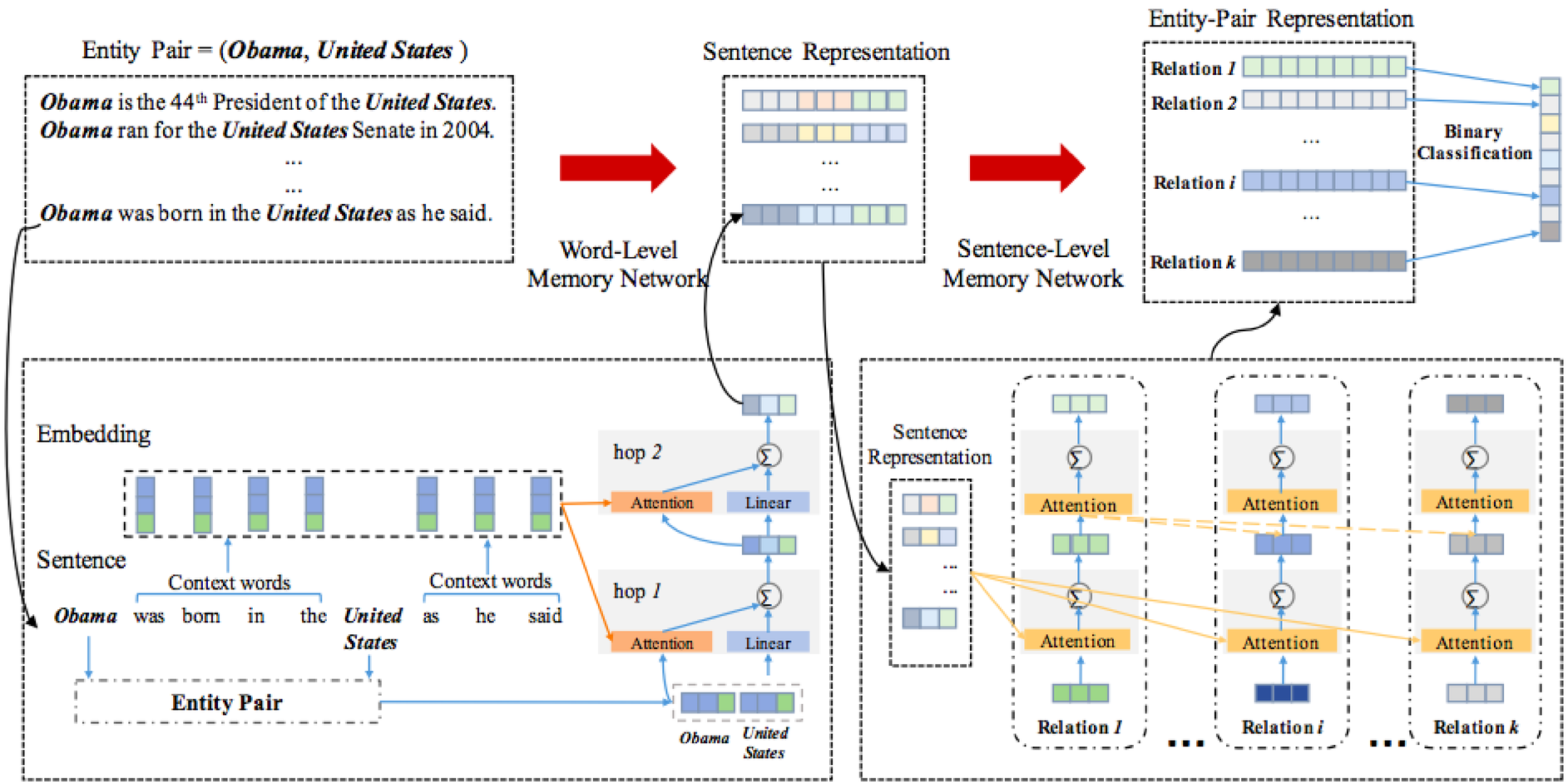
NYT

Relation Classification . distant supervision

MIML

NYT

	Top 100	Top 200	Top 500	Average
Mintz	0.77	0.71	0.55	0.676
Multir	0.83	0.74	0.59	0.720
MIML	0.85	0.75	0.61	0.737
PCNN	0.84	0.77	0.64	0.750
ATT	0.86	0.80	0.68	0.780
DMN	**0.89**	**0.82**	0.68	**0.797**



Xiaocheng Feng et al., 2017

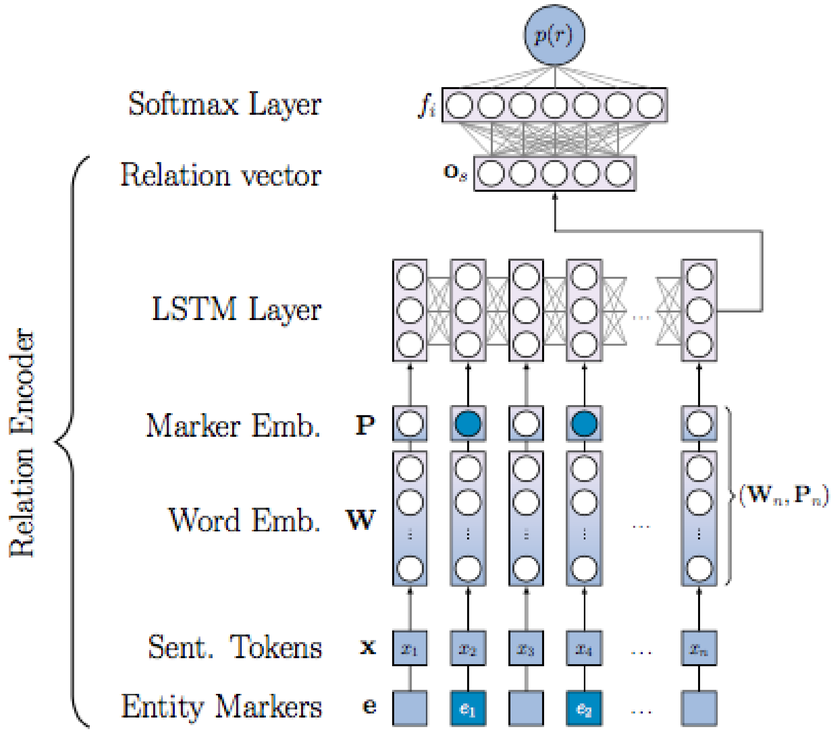
Harbin Institute of Technology

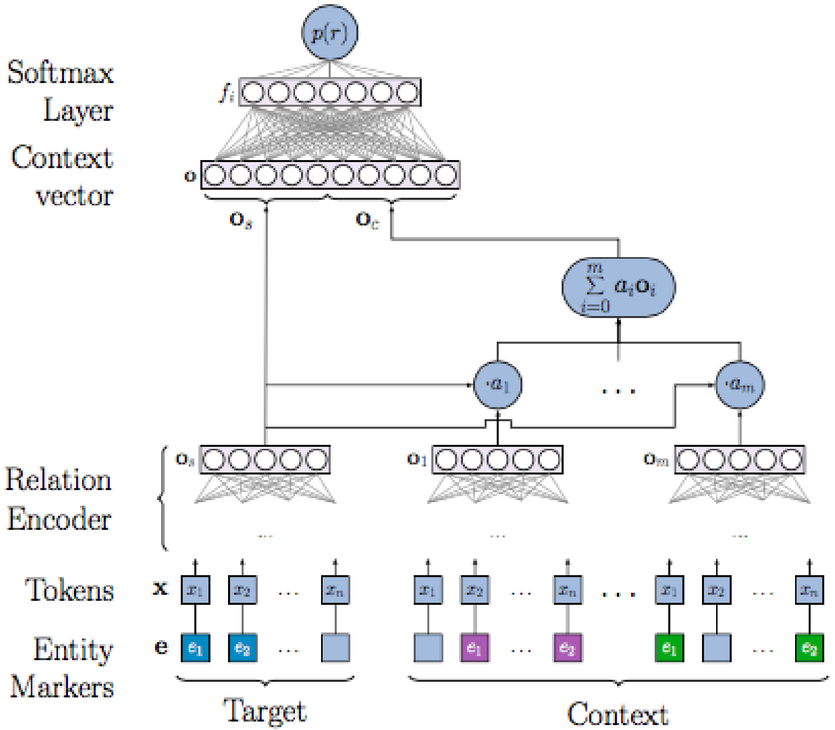
Relation Classification . distant supervision

MIML; multiple relation instances per sentence.

[Swag It Out] is the official [debut single] by [American singer] [e_1 Zendaya], known for starring in the series [e_2 Shake It Up].

Relation type	LSTM-Baseline P	LSTM-Baseline R	ContextAtt P	ContextAtt R
COUNTRY	0.8899	0.9344	0.9130	0.9382
LOCATED IN	0.8329	0.8832	0.8655	0.8994
SHARES BORDER	0.7579	0.7078	0.7962	0.8075
INSTANCE OF	0.7864	0.8568	0.8478	0.8401
SPORT	0.9753	0.9828	0.9822	0.9823
CITIZENSHIP	0.9001	0.9448	0.9041	0.9417
PART OF	0.5623	0.4854	0.6269	0.5113
SUBCLASS OF	0.5230	0.4390	0.5272	0.5908



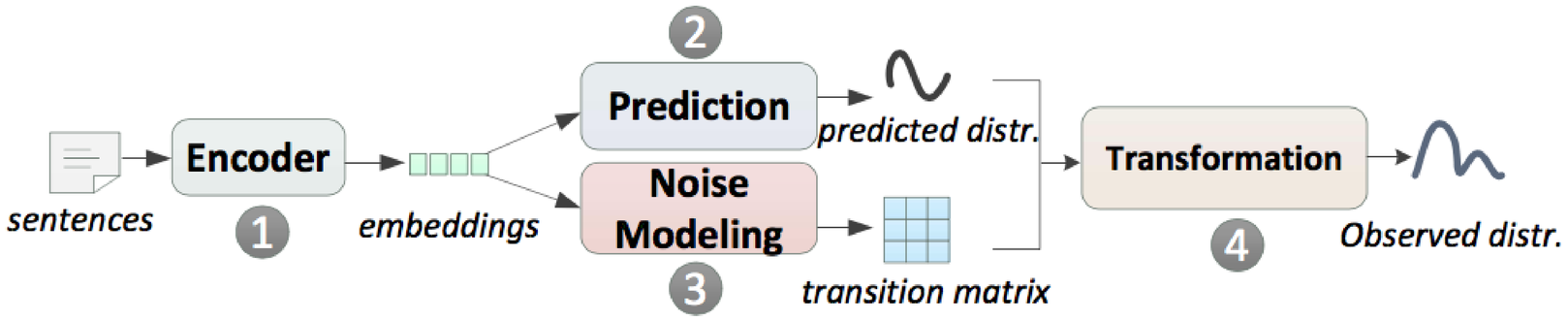


Daniil Sorokin et al., 2017

Technische Universita¨t Darmstadt

Relation Classification . distant supervision

sentence-level



$$L = \sum_{i=1}^{N} -((1-\alpha)log(o_{iy_i}) + \alpha log(p_{iy_i})) - \beta trace(\mathbf{T}^i)$$

Bingfeng Luo et al., 2017

Peking University

Method	P@R_10	P@R_20	P@R_30
Mintz	39.88	28.55	16.81
MultiR	60.94	36.41	-
MIML	60.75	33.82	-
avg	58.04	51.25	42.45
avg_TM	58.56	52.35	43.59
att	61.51	56.36	**45.63**
att_TM	**67.24**	**57.61**	44.90

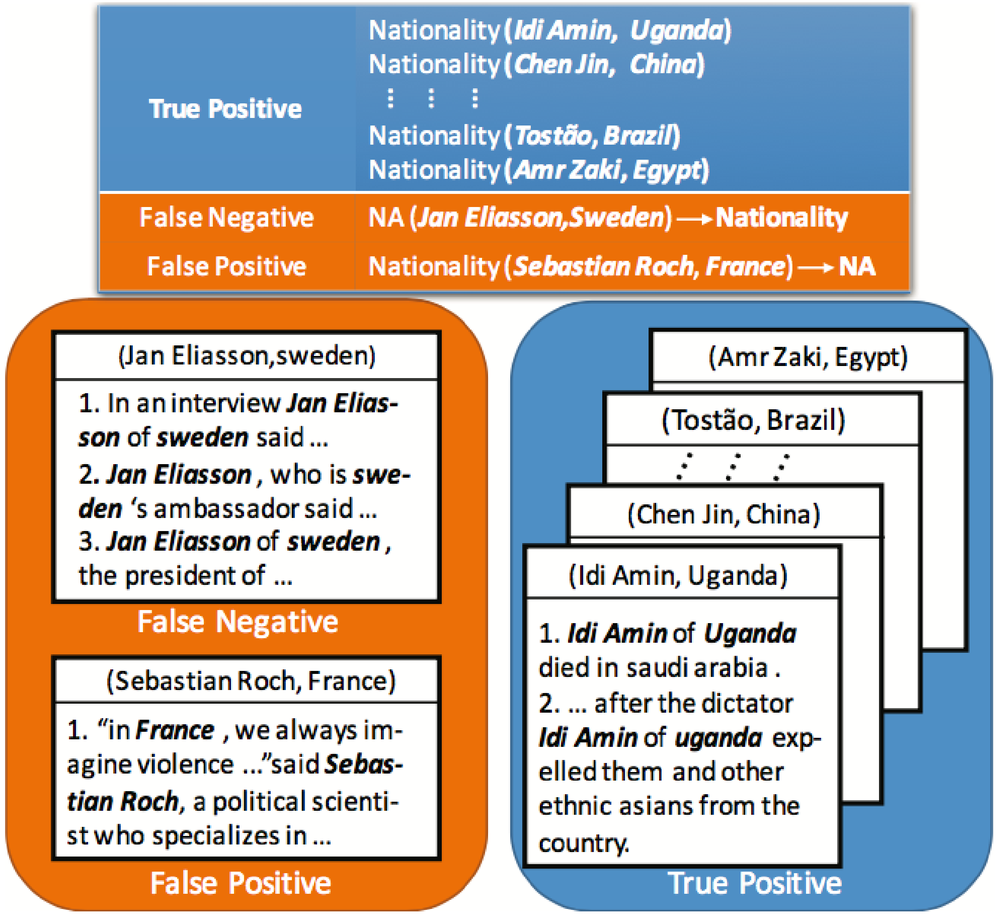
Relation Classification . distant supervision

MIML; entity-pair noise

hard-label VS soft-label

$$r_i = \arg\max(\mathbf{o} + \max(\mathbf{o})\mathbf{A} \odot L_i)$$

$$J(\theta) = \sum_{i=1}^{n} \log p(r_i|\mathbf{s}_i;\theta)$$

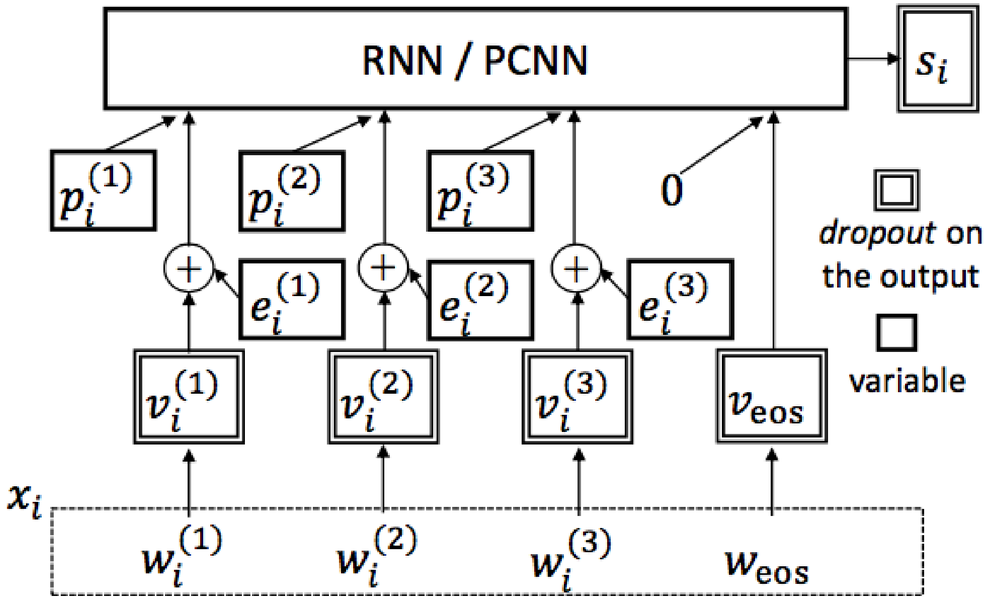


Tianyu Liu et al., 2017

Peking University

Relation Classification . distant supervision

MIML; adversarial training



$$L_{\text{adv}}(X;\theta) = L(X + e_{\text{adv}};\theta), \text{ where} \quad (2)$$

$$e_{\text{adv}} = \arg\max_{\|e\|\le\epsilon} L(X + e;\hat{\theta}) \quad (3)$$

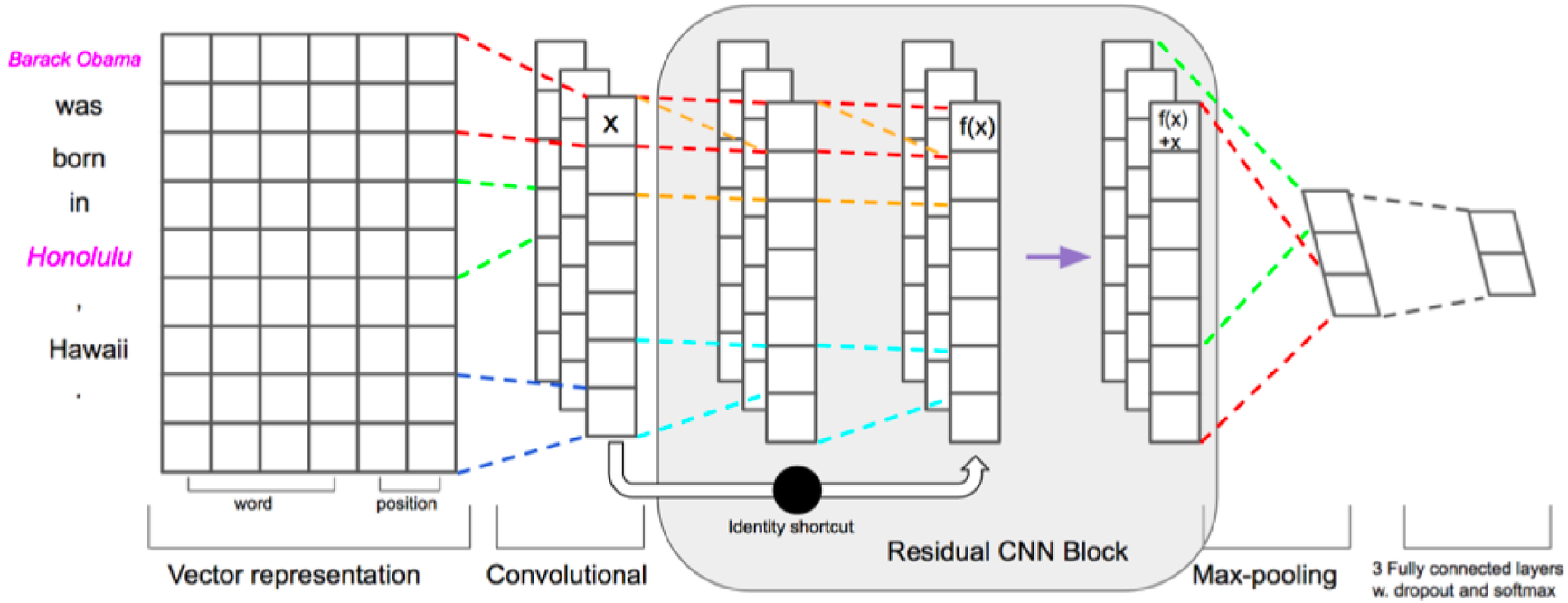
$$e_{\text{adv}} = \epsilon g/\|g\|, \text{ where } g = \nabla_V L(X;\hat{\theta}). \quad (4)$$

Yi Wu et al., 2017

UC Berkeley

Relation Classification . distant supervision

MIML; residual

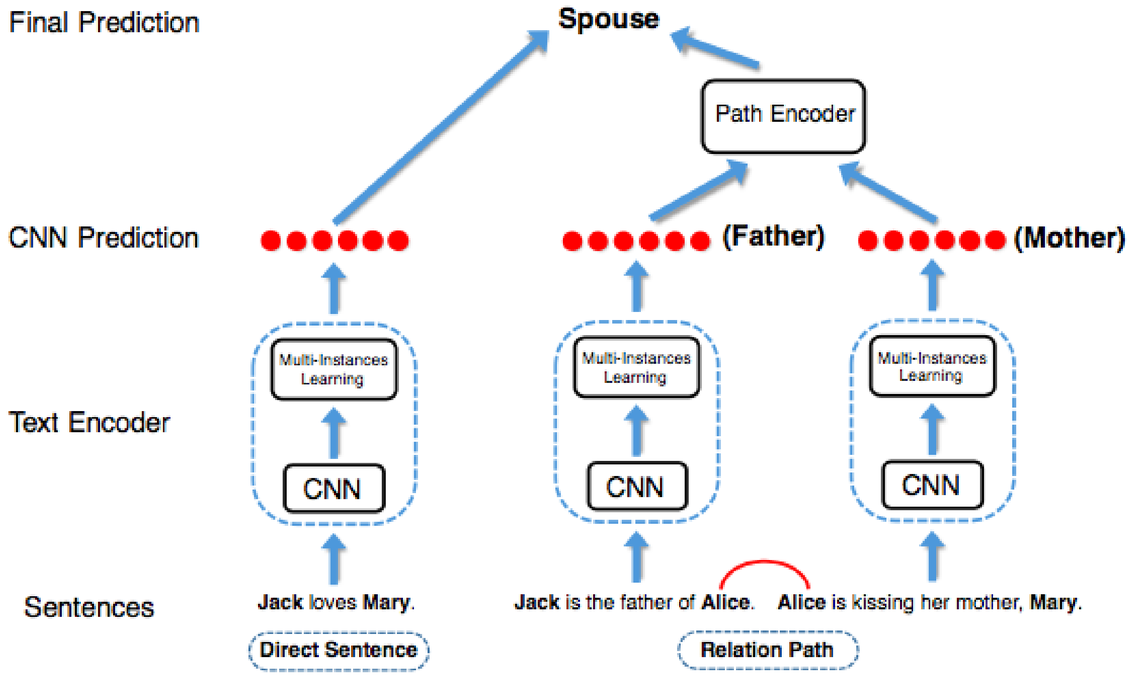


Yi Yao Huang et al., 2017

National Taiwan University

Relation Classification . distant supervision

MIML; relation paths



Wenyuan Zeng et al., 2017

Tsinghua University

Relation Classification . distant supervision

1.**Barack Obama** is the 44th and current President of the **U.S.**.(**President of**)
2.**Barack Obama** ended **U.S.**military involvement in the Iraq War.(-)
3.**Barack Obama** was born in honolulu, Hawaii, **U.S.**.(**Born in**)
4.**Barack Obama** ran for the **U.S.**Senate in 2004.(**Senate of**)

MIML; nonparametric Bayesian

$$Z = \begin{bmatrix} X_{train} & Y_{train} \\ X_{test} & Y_{test} \end{bmatrix}$$

$$Z = Z^* + E$$

$$Z^* = \begin{bmatrix} X^*_{train} & Y^*_{train} \\ X^*_{test} & Y^*_{test} \end{bmatrix}$$

$$E = \begin{bmatrix} E_{X_{train}} & E_{Y_{train}} \\ E_{X_{test}} & 0 \end{bmatrix}$$

$$p(\varepsilon_{i,j}) = \sum_{k=1}^{\infty} \theta_k N(\varepsilon_{i,j}|0, \sigma_k).$$

$$p(\mathbf{U}|\lambda) = \prod_{r=1}^{R} N(u_{.r}|0, \lambda_r \mathbf{I_U}),$$

$$p(\mathbf{V}|\lambda) = \prod_{r=1}^{R} N(v_{.r}|0, \lambda_r \mathbf{I_V}),$$

$$\lambda_r \sim IG(a_1, b_1),$$

$$p(y_{i,j}) = N(y_{i,j}| \underbrace{u_{i.}v^T_{j.}}_{low-rank\ component}, \underbrace{\varepsilon_{i,j}}_{noise\ component}),$$

$$\varepsilon_{i,j} = \sigma_{z_{ij}},$$

$$\sigma_{z_{ij}} \sim IG(a_0, b_0),$$

$$z_{ij} = k \sim Mult(\theta_k),$$

$$p(\mathbf{U}, \mathbf{V}, \lambda, \sigma, \mathbf{z}, \beta | \mathbf{X}_{observed}, \mathbf{Y}_{observed})$$

Qing Zhang et al., 2017

Peking University

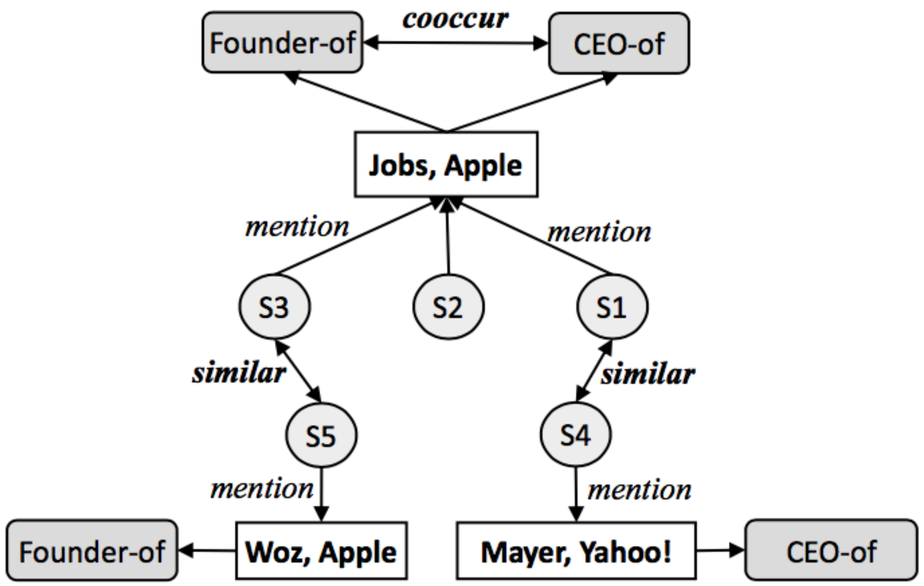
Relation Classification . distant supervision

dealing with noise; global distant supervision

Knowledge Base
CEO-of(*Steve Jobs, Apple*)
Founder-of(*Steve Jobs, Apple*)
CEO-of(*Marissa Mayer, Yahoo!*)
Founder-of(*Steve Wozniak, Apple*)

Heuristically labeling

Relation instances	Possible labels by DS	Golden labels
S1: ***Jobs**, the CEO of **Apple***	**CEO-of, Founder-of, NR**	**CEO-of**
S2: ***Jobs** left **Apple***	**CEO-of, Founder-of, NR**	**NR**
S3: ***Jobs** is the cofounder of **Apple***	**CEO-of, Founder-of, NR**	**Founder-of**
S4: ***Mayer** is the new CEO of **Yahoo!***	**CEO-of, NR**	**CEO-of**
S5: ***Woz** is the cofounder of **Apple***	**Founder-of, NR**	**Founder-of**

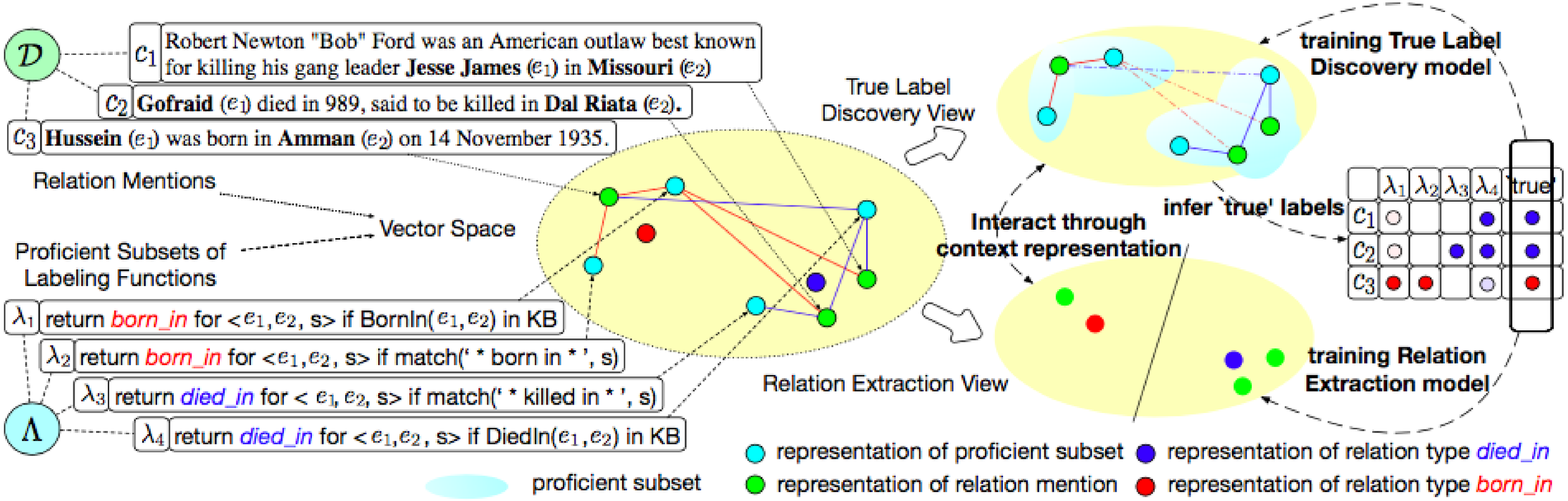


Xianpei Han et al., 2016

Institute of Software

Relation Classification . distant supervision

dealing with noise; heterogeneous supervision

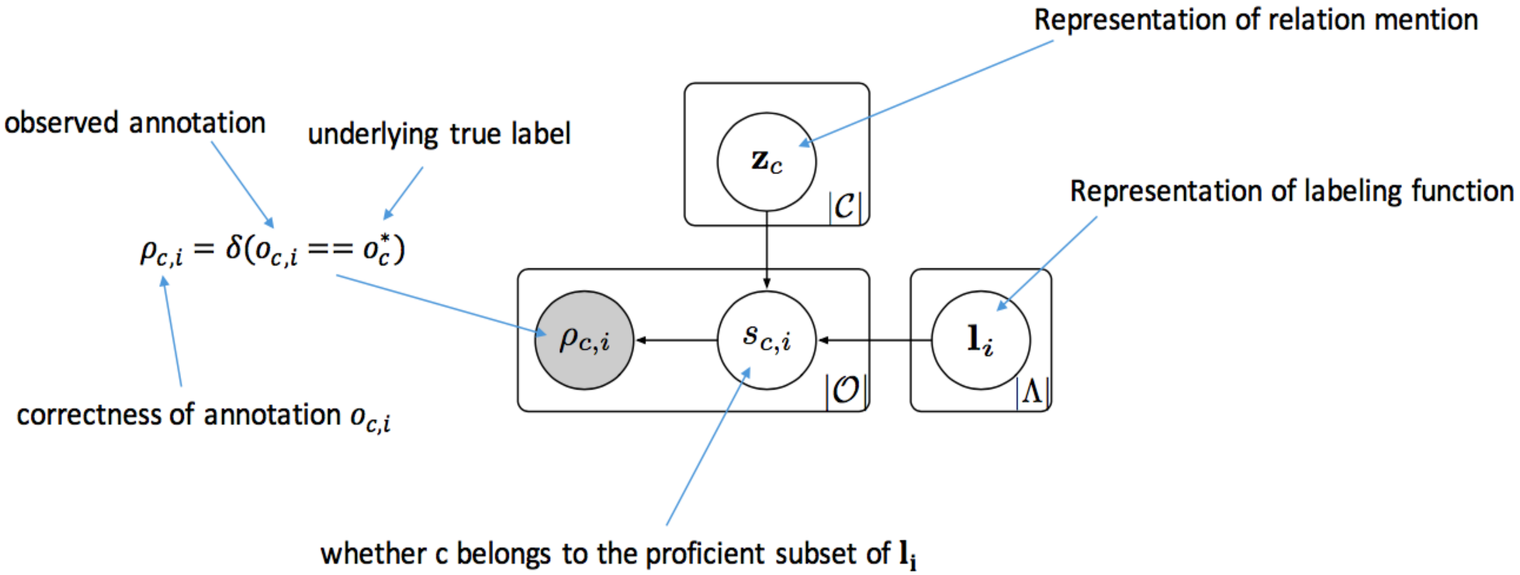


Liyuan Liu et al., 2017

University of Illinois at Urbana-Champaign

Relation Classification . distant supervision

dealing with noise; heterogeneous supervision



Method	Relation Extraction						Relation Classification	
	NYT			Wiki-KBP			NYT	Wiki-KBP
	Prec	Rec	F1	Prec	Rec	F1	Accuracy	Accuracy
NL+FIGER	0.2364	0.2914	0.2606	0.2048	0.4489	0.2810	0.6598	0.6226
NL+BFK	0.1520	0.0508	0.0749	0.1504	0.3543	0.2101	0.6905	0.5000
NL+DSL	0.4150	0.5414	0.4690	0.3301	0.5446	0.4067	0.7954	0.6355
NL+MultiR	0.5196	0.2755	0.3594	0.3012	0.5296	0.3804	0.7059	0.6484
NL+FCM	0.4170	0.2890	0.3414	0.2523	0.5258	0.3410	0.7033	0.5419
NL+CoType-RM	0.3967	0.4049	0.3977	**0.3701**	0.4767	0.4122	0.6485	0.6935
TD+FIGER	0.3664	0.3350	0.3495	0.2650	**0.5666**	0.3582	0.7059	0.6355
TD+BFK	0.1011	0.0504	0.0670	0.1432	0.1935	0.1646	0.6292	0.5032
TD+DSL	0.3704	0.5025	0.4257	0.2950	0.5757	0.3849	0.7570	0.6452
TD+MultiR	**0.5232**	0.2736	0.3586	0.3045	0.5277	0.3810	0.6061	0.6613
TD+FCM	0.3394	0.3325	0.3360	0.1964	0.5645	0.2914	0.6803	0.5645
TD+CoType-RM	0.4516	0.3499	0.3923	0.3107	0.5368	0.3879	0.6409	0.6890
REHESSION	0.4122	**0.5726**	**0.4792**	0.3677	0.4933	**0.4208**	**0.8381**	**0.7277**

Liyuan Liu et al., 2017

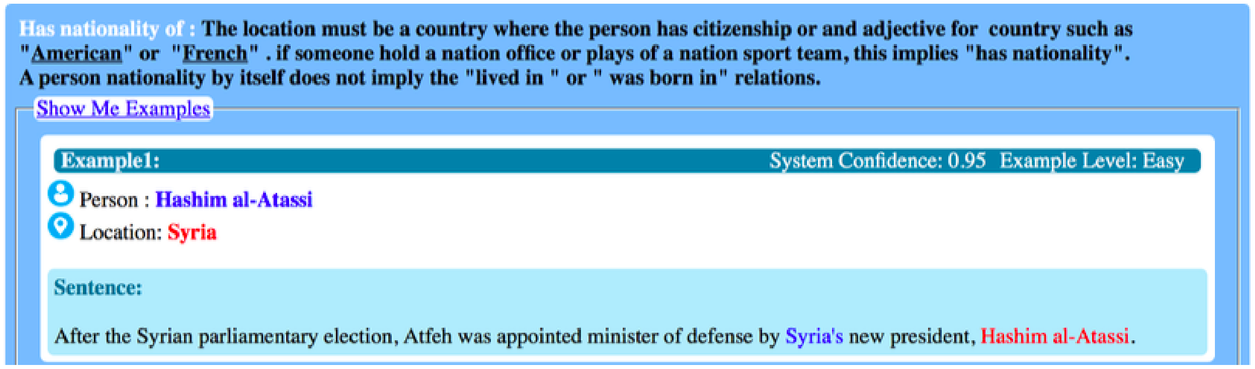
University of Illinois at Urbana-Champaign

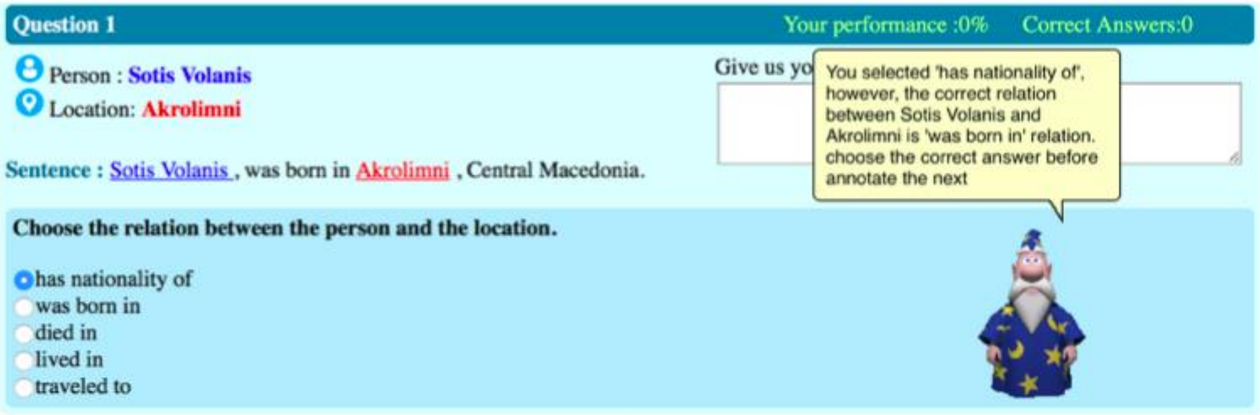
Relation Classification . distant supervision

dealing with noise;
adding high-quality annotated data;
crowdsourcing

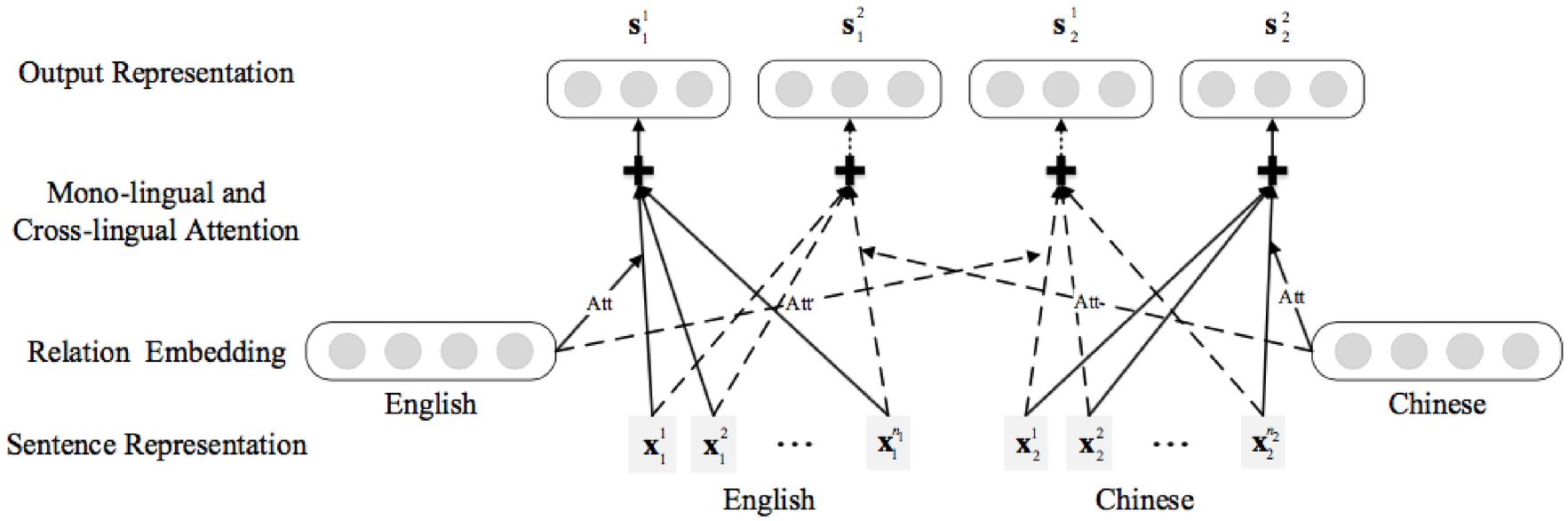
Azad Abad et al., 2017

University of Trento





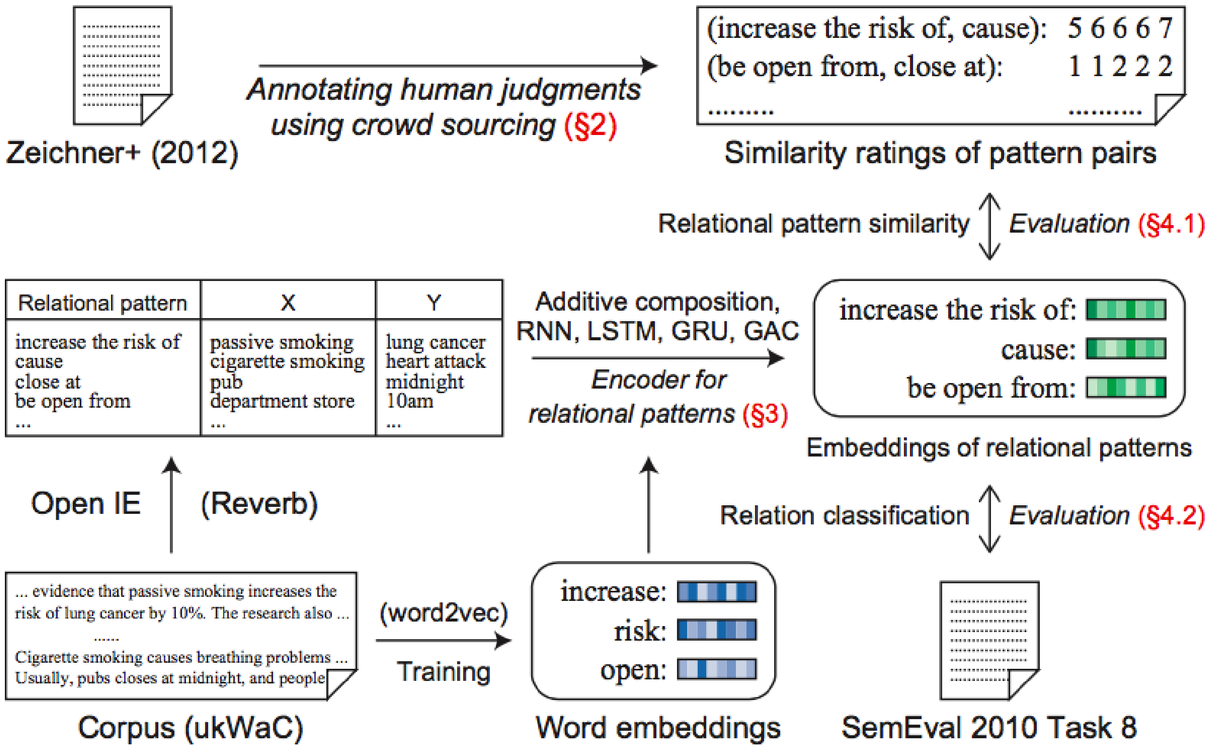
Relation Classification . multi-lingual

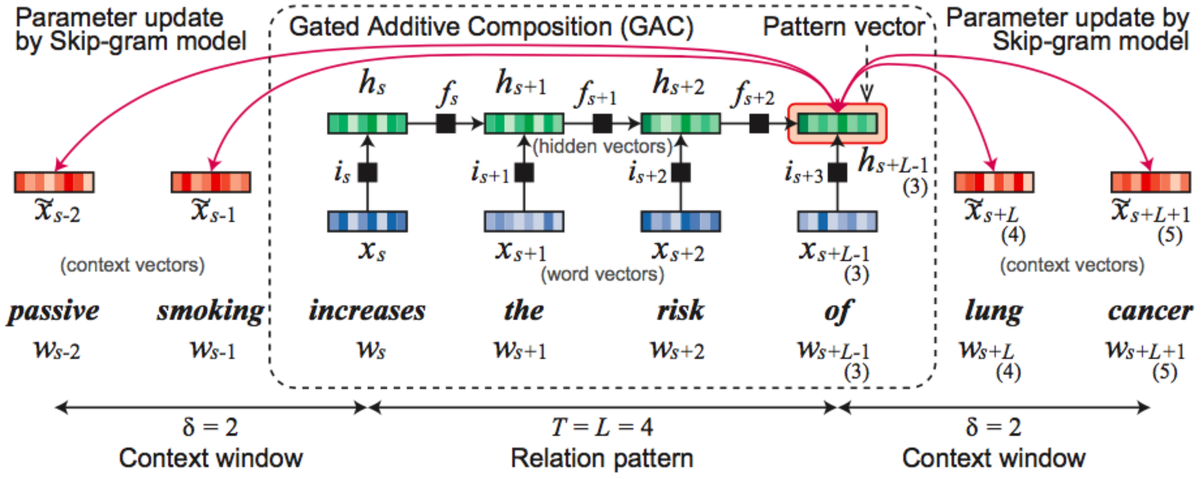


Yankai Lin et al., 2017

Tsinghua University

Relation Classification . relational phrases





Sho Takase et al., 2016

Tohoku University

Relation Classification . relational phrases

Relational phrase typing

↓

Relational phrase clustering

Cluster of relational phrases
<location> is the heart of <location>
<location> is situated in <location>
<location> is enclosed by <location>
<location> is located amidst <location>
<location> is surrounded by <location>

Table 1: Example of a cluster of relational phrases

Id	Relation phrase	Synonymous Relational Phrase
1	<location> is surrounded by <region>	<location> is the heart of <region>
2	<artifact> is reminiscent of <time_period>	<artifact> recalls <time_period>
3	<painter> was a participant in <show>	<painter> has participated in <show>
4	<group> maintains a partnership with <district>	<group> has partnered with <district>
5	<movie> was shot at <location>	<movie> was filmed in <location>
6	<person> was shot by <group>	<person> was shot dead by <group>
7	<movie> was shot by <film_director>	<movie> was directed by <film_director>

Table 3: Examples of synonyms of semantically typed relational phrases

Adam Grycner et al., 2016

Max-Planck Institute for Informatics

Relation Classification . maybe future work

- Collective learning (entity-level, entity-pair level, sentence-level, relation-level)
- Exploration in noise
- Low resource (language, domain)

ERE Extraction

Joint extracting entities and relations.

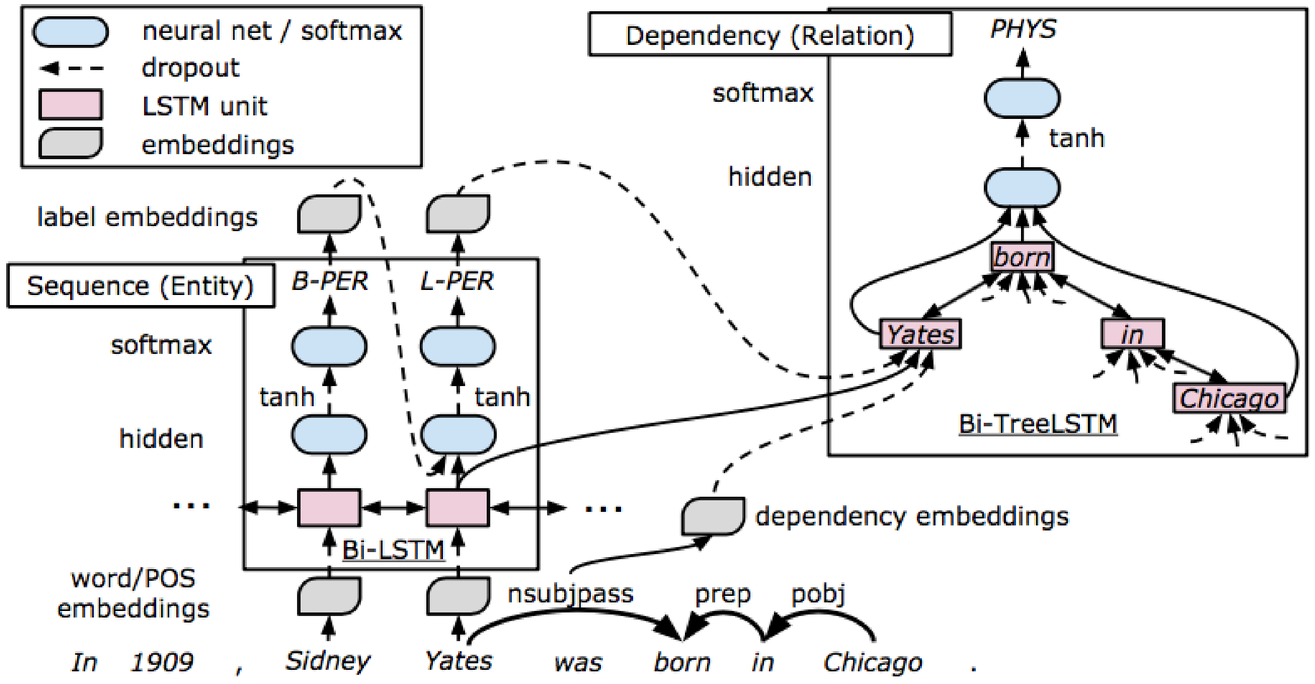
Research Focus

multi-task

joint learning

relation schema induction

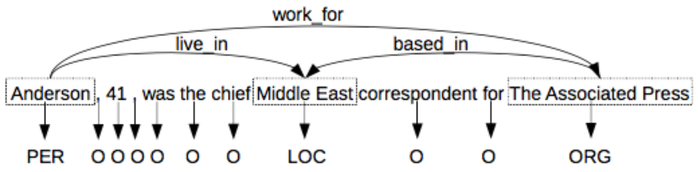
ERE Extraction . multi-task



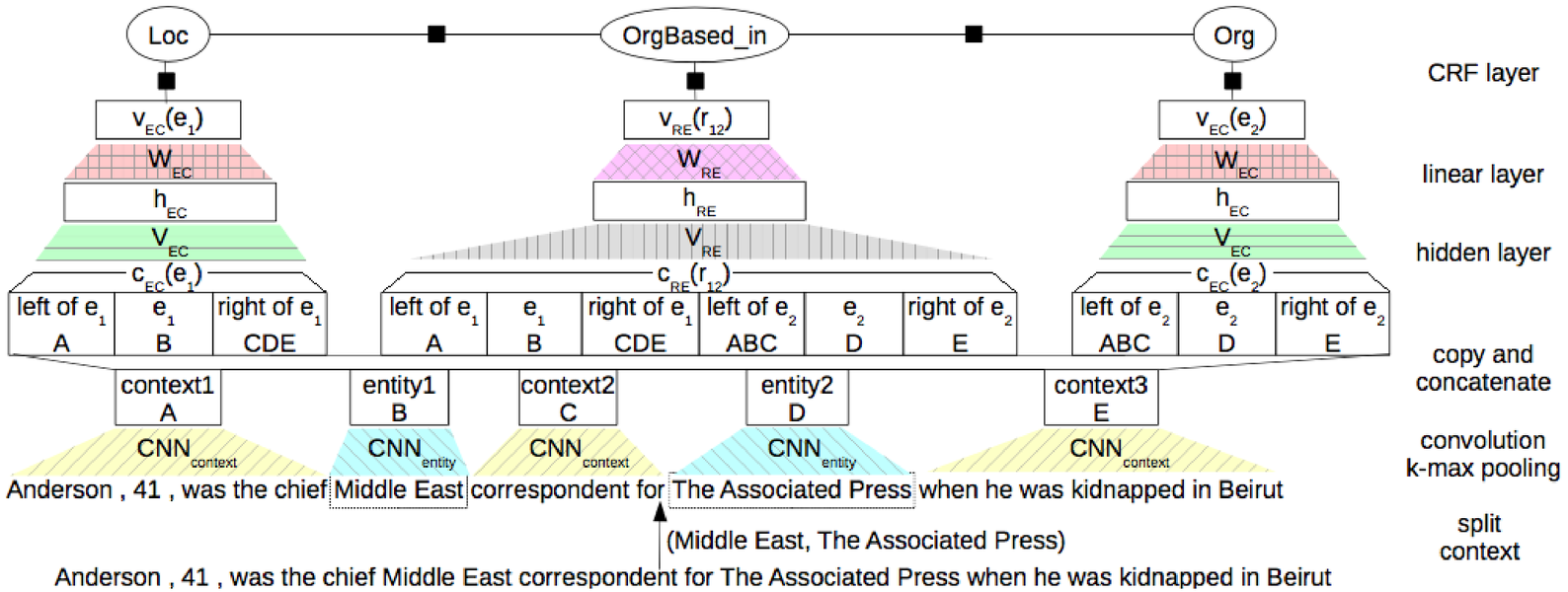
complementary linguistic structures:
word sequence and dependency tree structures

Makoto Miwa et al., 2016

ERE Extraction . multi-task



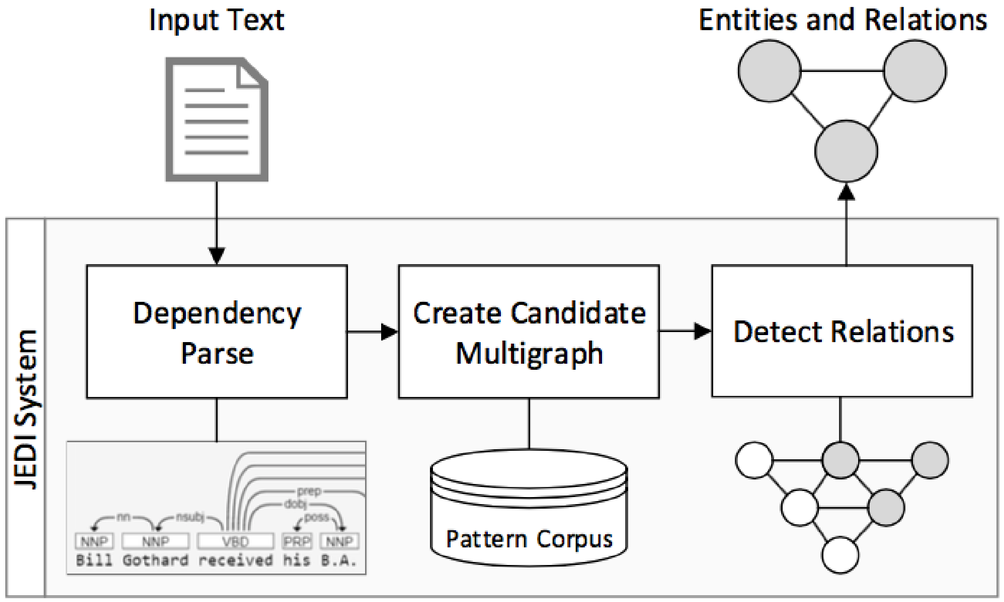
Model inputs (query entity pairs) ⇒ Model outputs:
(“Anderson”, “,”) ⇒ PER - N - O
(“Anderson”, “41”) ⇒ PER - N - O
...
(“Anderson”, “chief”) ⇒ PER - N - O
(“Anderson”, “Middle East”) ⇒ PER - live_in - LOC
...
(“was”, “for”) ⇒ O - N - O
...
(“for”, “The Associated Press”) ⇒ O - N - ORG



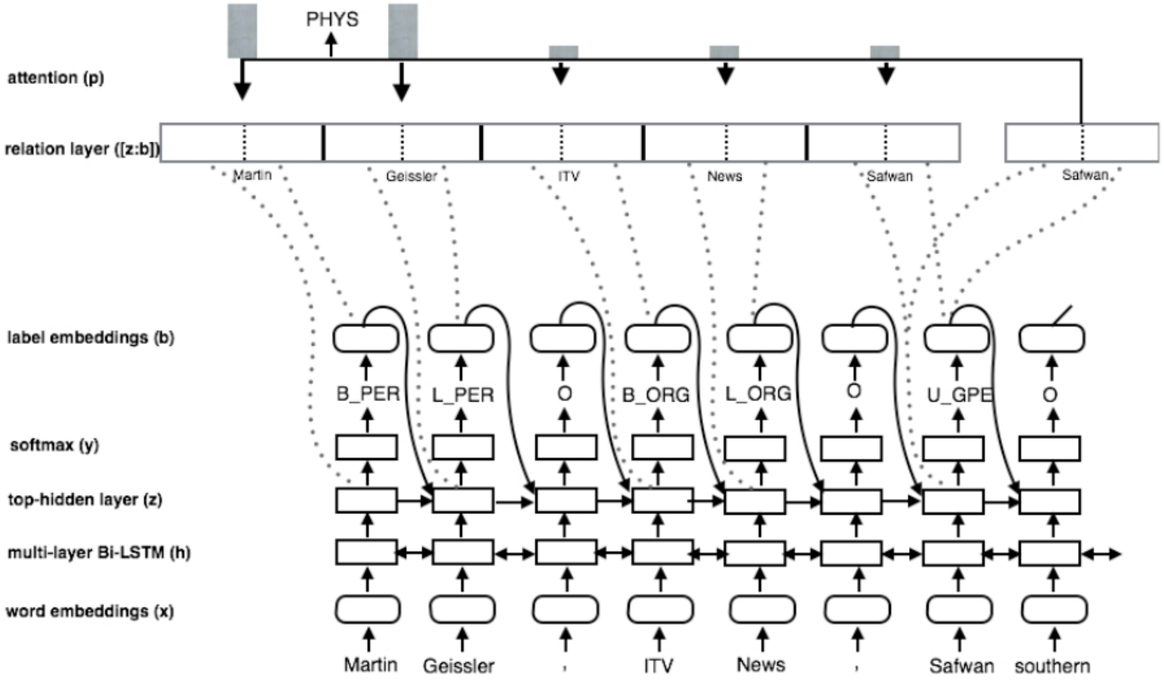
Heike Adel et al., 2017

ERE Extraction . joint learning

a graph perspective

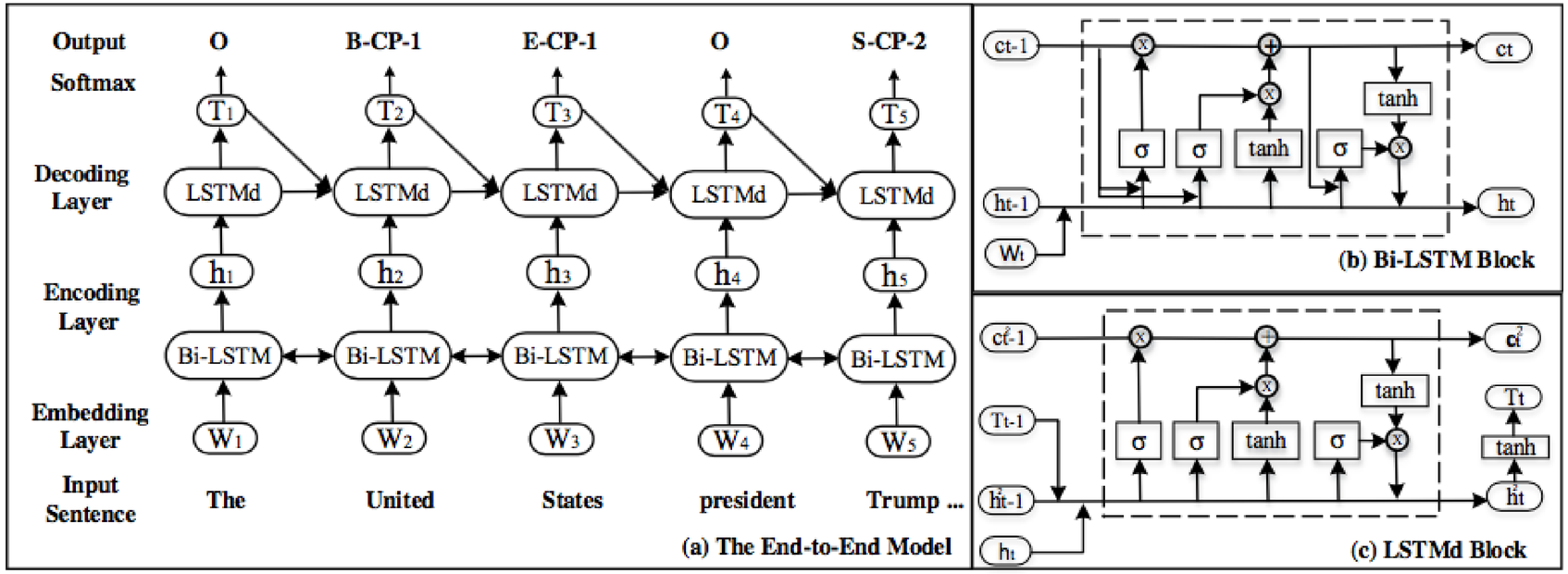


Johannes Kirschnick et al., 2016



Arzoo Katiyar et al., 2017

ERE Extraction . joint learning



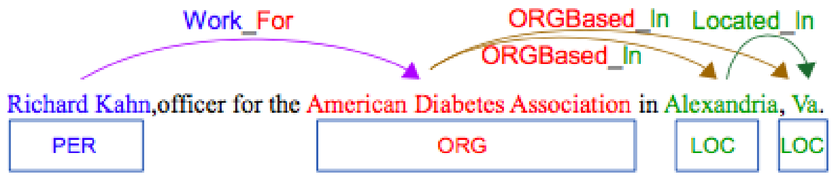
unified tagging

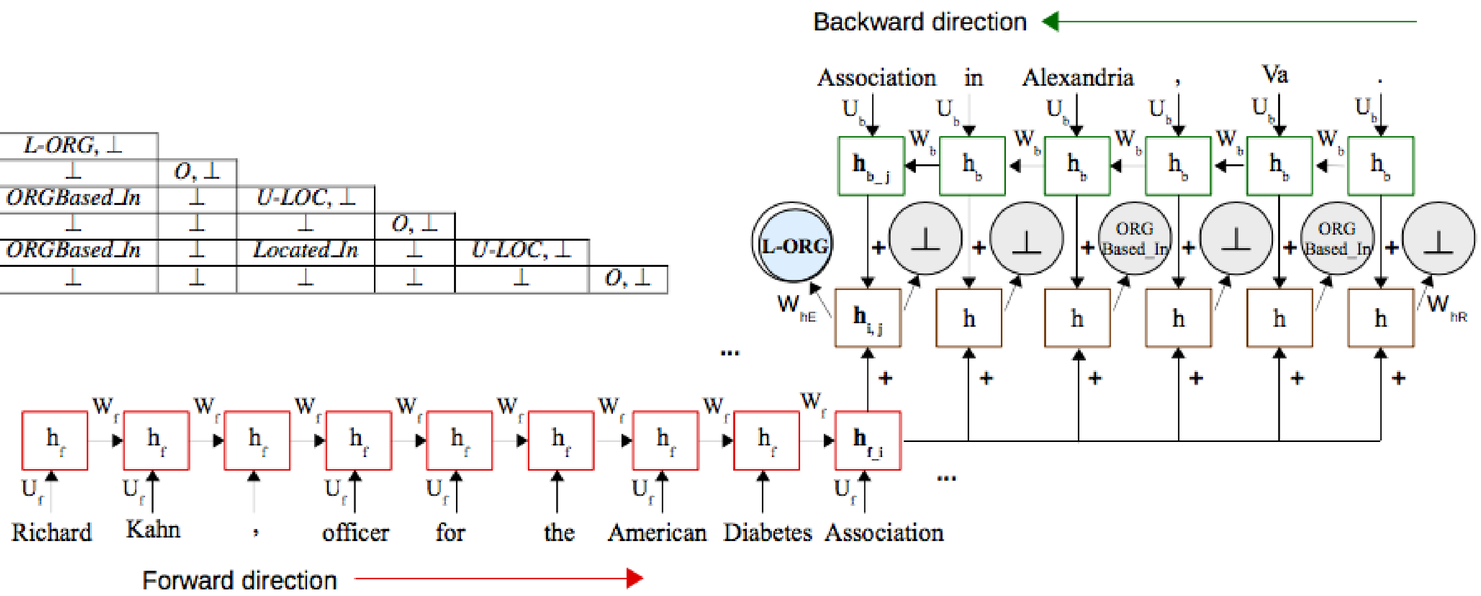
Suncong Zheng et al., 2017

ERE Extraction . joint learning

table filling

	Richard	Kahn	,	officer	for	the	American	Diabetes	Association	in	Alexandria	,	Va	.
Richard	*B-PER*, ⊥													
Kahn	⊥	*L-PER*, ⊥												
,	⊥	⊥	*O*, ⊥											
officer	⊥	⊥	⊥	*O*, ⊥										
for	⊥	⊥	⊥	⊥	*O*, ⊥									
the	⊥	⊥	⊥	⊥	⊥	*O*, ⊥								
Ameriacan	⊥	⊥	⊥	⊥	⊥	⊥	*B-ORG*, ⊥							
Diabetes	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*I-ORG*, ⊥						
Association	⊥	*Work_For*	⊥	⊥	⊥	⊥	⊥	⊥	*L-ORG*, ⊥					
in	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*O*, ⊥				
Alexandria	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*ORGBased_In*	⊥	*U-LOC*, ⊥			
,	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*O*, ⊥		
Va	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*ORGBased_In*	⊥	*Located_In*	⊥	*U-LOC*, ⊥	
.	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	*O*, ⊥



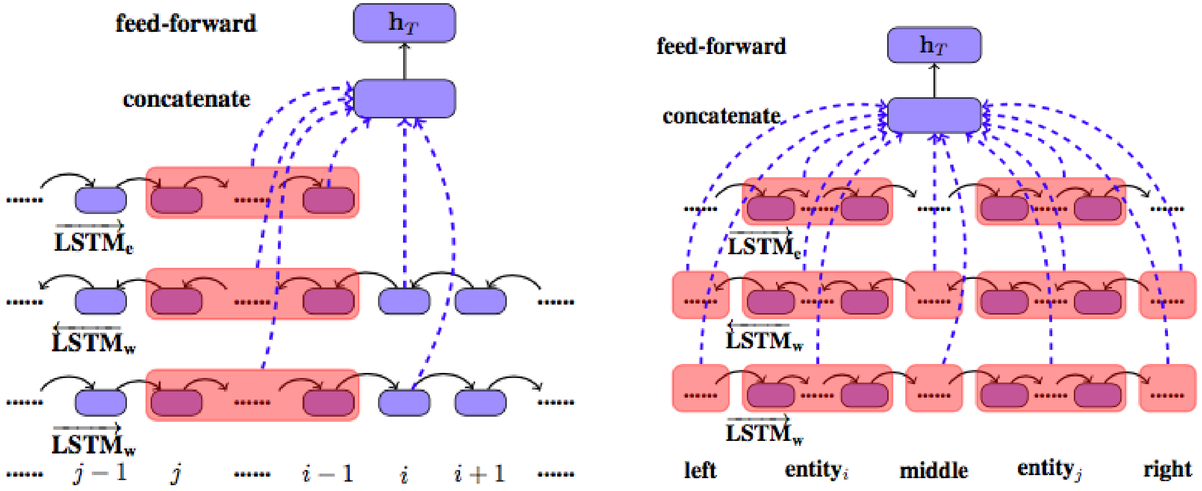


Pankaj Gupta et al., 2016

ERE Extraction . joint learning

table filling

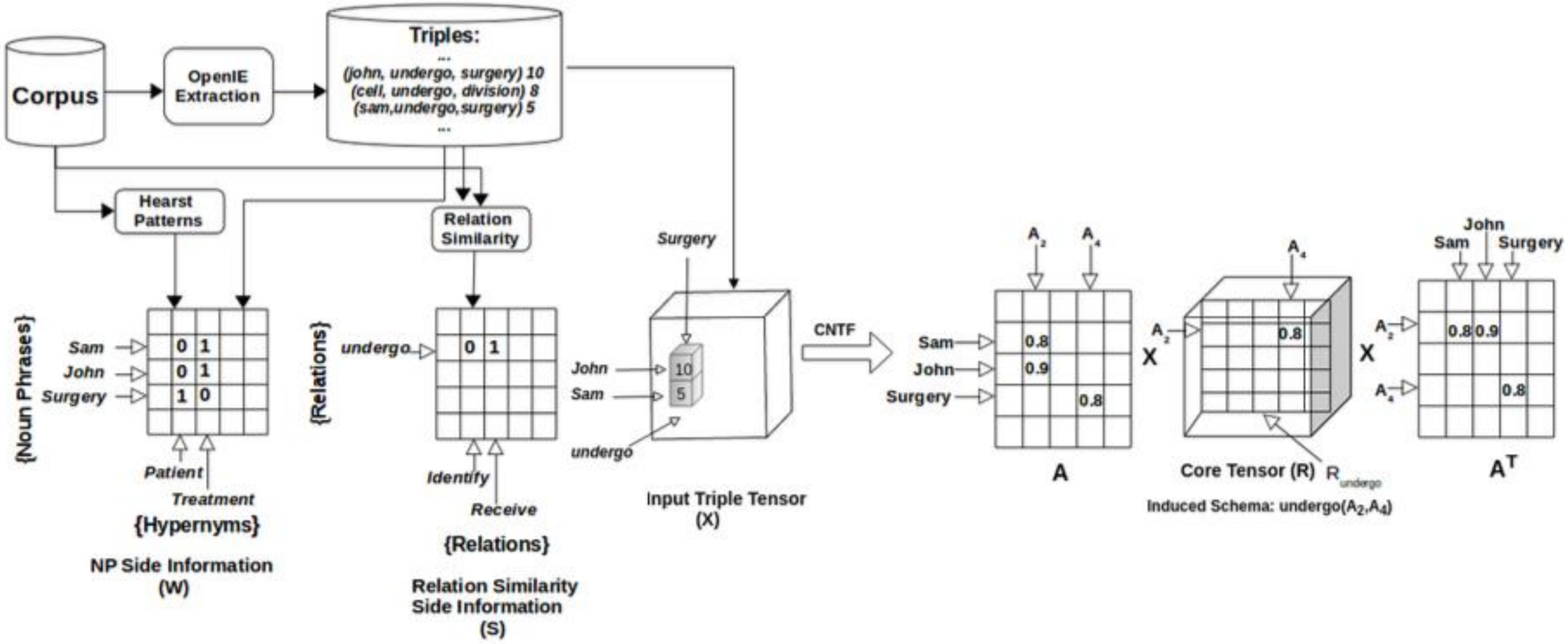
	Associated	Press	writer	Patrick	McDowell	in	Kuwait	City
Associated	1 B-ORG	9 ⊥	16 ⊥	22 ⊥	27 ⊥	31 ⊥	34 ⊥	36 ⊥
Press		2 L-ORG	10 $\overleftarrow{\text{ORG-AFF}}$	17 ⊥	23 ⊥	28 ⊥	32 ⊥	35 ⊥
writer			3 U-PER	11 ⊥	18 ⊥	24 ⊥	29 ⊥	33 ⊥
Patrick				4 B-PER	12 ⊥	19 ⊥	25 ⊥	30 ⊥
McDowell					5 L-PER	13 ⊥	20 ⊥	26 $\overrightarrow{\text{PHYS}}$
in						6 O	14 ⊥	21 ⊥
Kuwait							7 B-GPE	15 ⊥
City								8 L-GPE



(a) entity detection (b) relation classification

Meishan Zhang et al., 2017

ERE Extraction . relation schema induction



Madhav Nimishakavi et al., 2016

KB Reasoning (Embeddings & Completion)

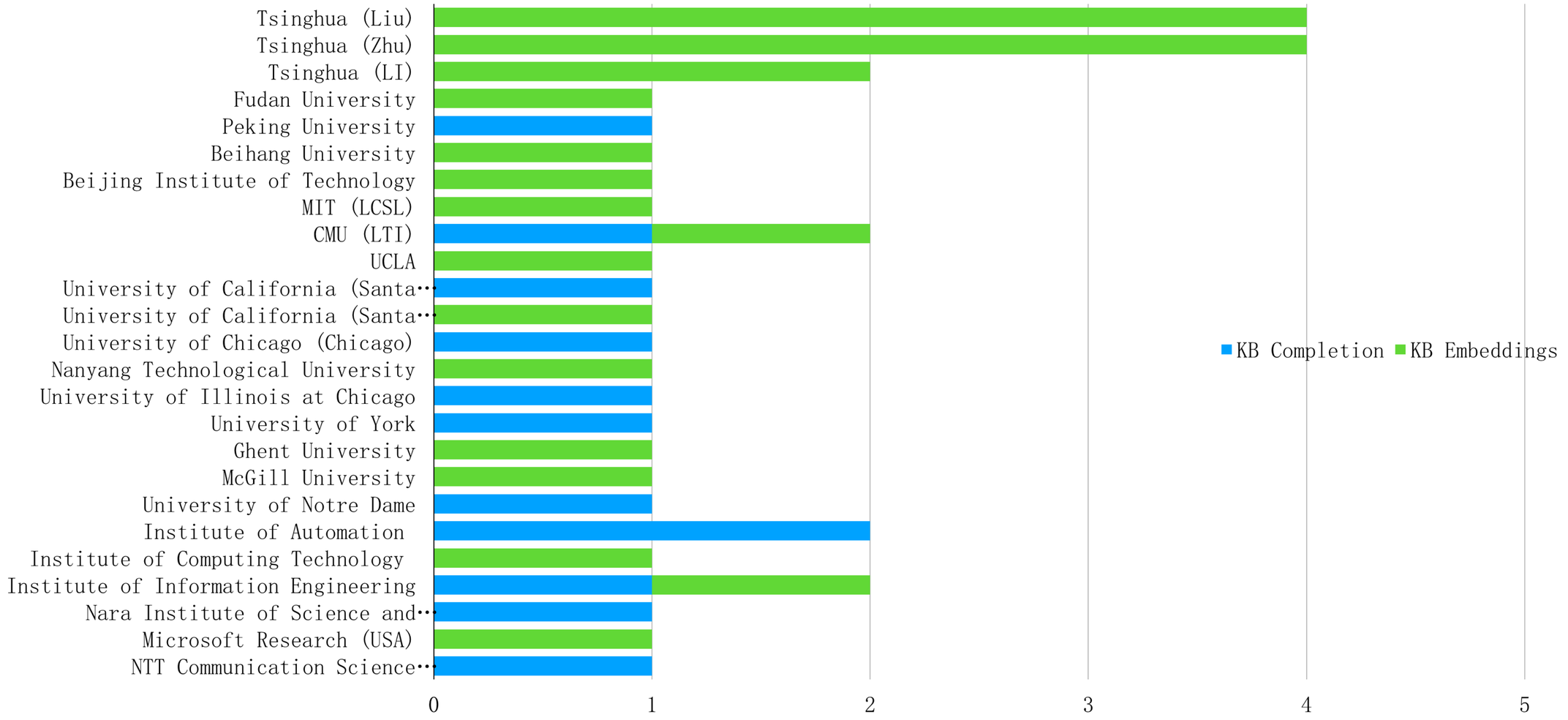
45 papers from 25 organization

Tsinghua (Liu)
Tsinghua (Zhu)
Tsinghua (LI)
Fudan University
Peking University
Beihang University
Beijing Institute of Technology

MIT (LCSL)
CMU (LTI)
UCLA
University of California (Santa Barbara)
University of California (Santa Cruz)
University of Chicago (Chicago)
Nanyang Technological University
University of Illinois at Chicago
University of York
Ghent University
McGill University
University of Notre Dame

Institute of Automation
Institute of Computing Technology
Institute of Information Engineering
Nara Institute of Science and Technology
Microsoft Research (USA)
NTT Communication Science Laboratories

KB Reasoning (Embeddings & Completion)



KB Embeddings

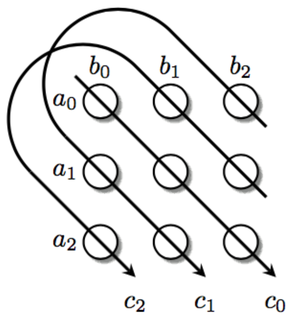
KB Embeddings aims at offering a continuous knowledge representation paradigm by transforming the entities and relations into continuous vector space.

Research Focus

- efficiency and simplicity, expressivity
- margin loss function training
- text information, lexical resources, logical rules
- multi-fold relations
- knowledge fusion
- bridge text and KB

KB Embeddings . efficiency and simplicity, expressivity

A mathematic perspective

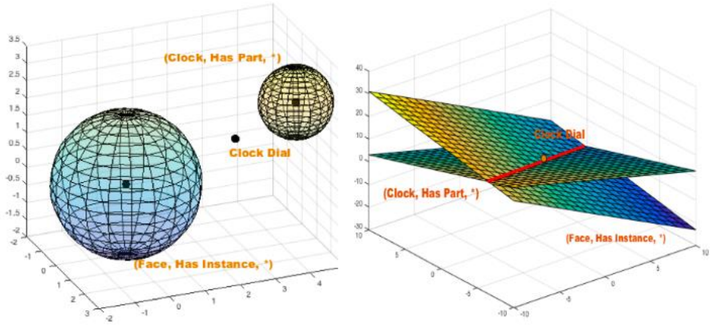


$$\boldsymbol{c} = \boldsymbol{a} \star \boldsymbol{b}$$

$$c_0 = a_0 b_0 + a_1 b_1 + a_2 b_2$$
$$c_1 = a_0 b_2 + a_1 b_0 + a_2 b_1$$
$$c_2 = a_0 b_1 + a_1 b_2 + a_2 b_0$$

	WN18					FB15k				
	MRR		Hits at			MRR		Hits at		
Method	Filter	Raw	1	3	10	Filter	Raw	1	3	10
TransE	0.495	0.351	11.3	88.8	94.3	0.463	0.222	29.7	57.8	**74.9**
TransR	0.605	0.427	33.5	87.6	94.0	0.346	0.198	21.8	40.4	58.2
ER-MLP	0.712	0.528	62.6	77.5	86.3	0.288	0.155	17.3	31.7	50.1
RESCAL	0.890	0.603	84.2	90.4	92.8	0.354	0.189	23.5	40.9	58.7
HOLE	**0.938**	**0.616**	**93.0**	**94.5**	**94.9**	**0.524**	**0.232**	**40.2**	**61.3**	73.9

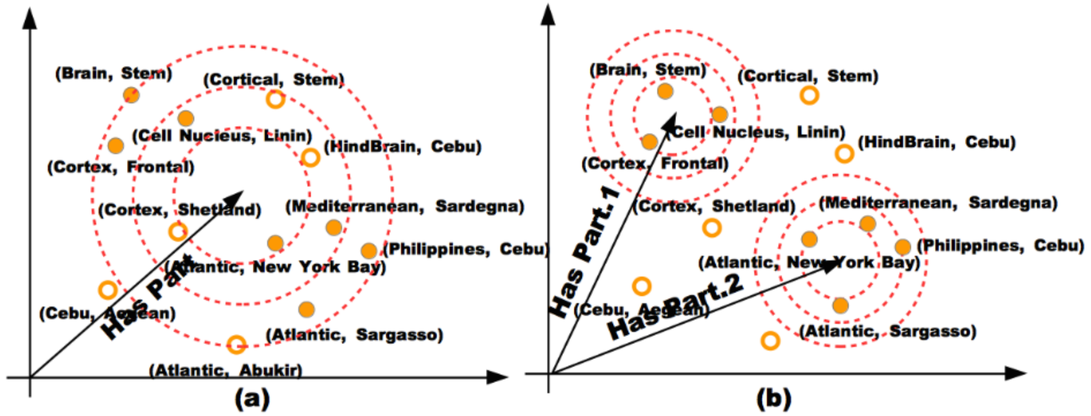
Nickel et al., 2016



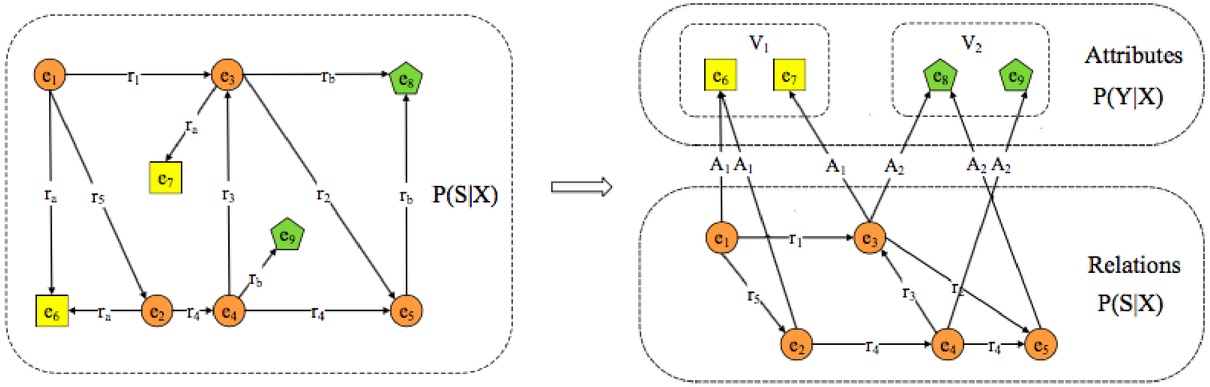
Han Xiao et al., 2016

Datasets	WN18				FB15K			
Metric	HITS@10(%)		HITS@1(%)	Time(s)	HITS@10(%)		HITS@1(%)	Time(s)
	Raw	Filter	Filter	One Epos	Raw	Filter	Filter	One Epos
SE[Bordes *et al.*, 2011]	68.5	80.5	-	-	28.8	39.8	-	-
TransE [Bordes *et al.*, 2013]	75.4	89.2	29.5	**0.4**	34.9	47.1	24.4	**0.7**
TransH [Wang *et al.*, 2014]	73.0	82.3	31.3	1.4	48.2	64.4	24.8	4.8
TransR [Lin *et al.*, 2015b]	79.8	92.0	33.5	9.8	48.4	68.7	20.0	29.1
PTransE [Lin *et al.*, 2015a]	-	-	-	-	51.4	84.6	63.3	266.0
KG2E [He *et al.*, 2015]	80.2	92.8	54.1	10.7	48.9	74.0	40.4	44.2
ManifoldE Sphere	80.7	92.8	55.8	**0.4**	**55.7**	86.2	64.1	**0.7**
ManifoldE Hyperplane	**84.2**	**94.9**	**93.2**	0.5	55.2	**88.1**	**70.5**	0.8

KB Embeddings . efficiency and simplicity, expressivity



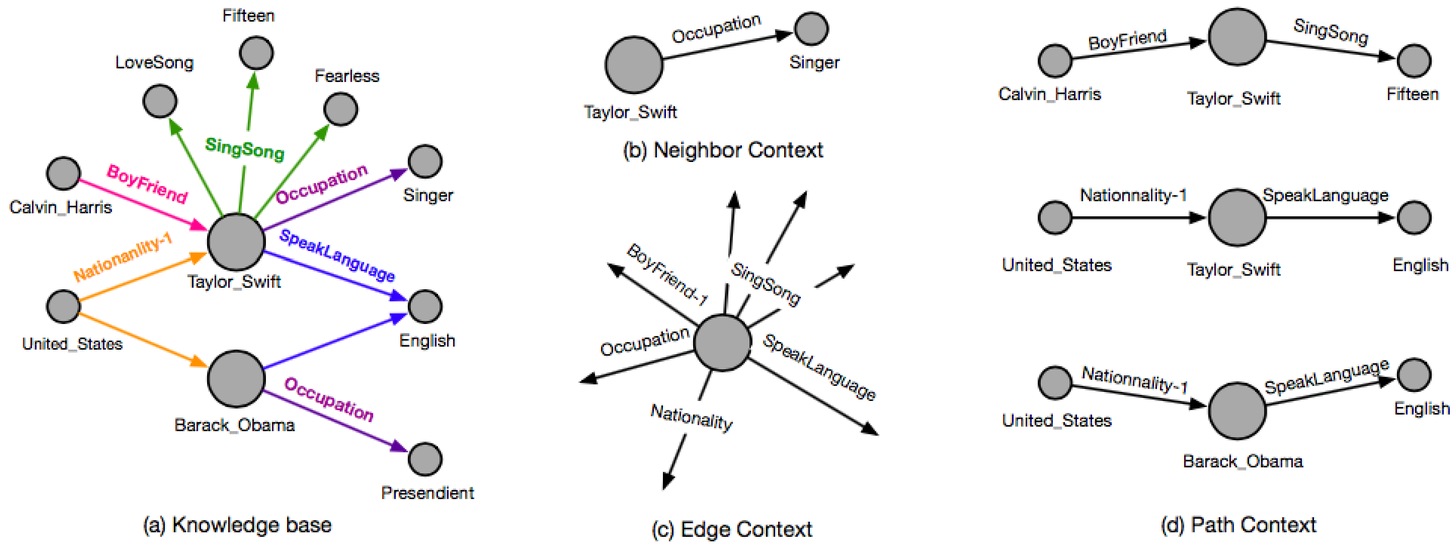
Han Xiao et al., 2016



Yankai Lin et al., 2016

KB Embeddings . efficiency and simplicity, expressivity

A graph perspective



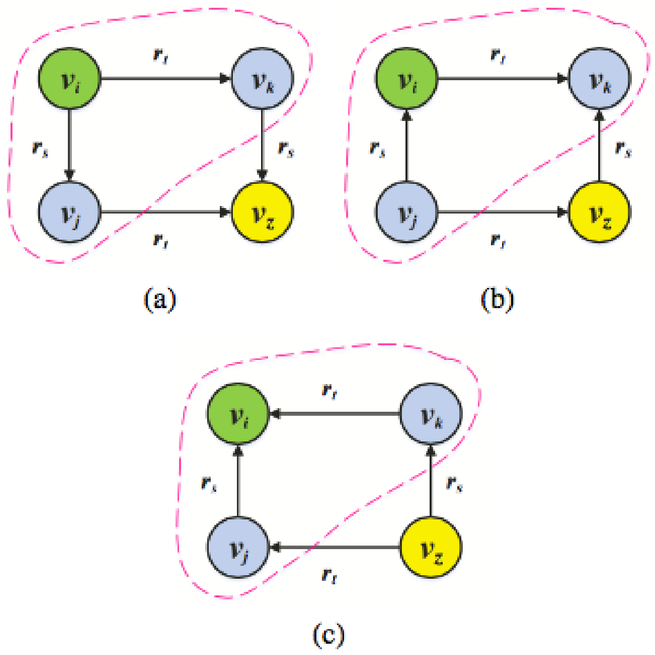
$$P(s_i|c(s_i)) = \frac{exp(\phi(s_i)^\top \pi(c(s_i))}{\sum_{j=1}^{|S|} exp(\phi(s_j)^\top \pi(c(s_i)))}$$

$$O_N = \sum_{s_i \in S} \sum_{c_N(s_i) \in C_N(s_i)} \log p(s_i|c_N(s_i))$$

Jun Feng et al., 2016

KB Embeddings . efficiency and simplicity, expressivity

A graph perspective



Lin Liu et al., 2017

$$O = \sum_{i \in V} \lambda_i KL\left(\hat{p}\left(\cdot|v_i\right)||p\left(\cdot|v_i\right)\right)$$

$$O_1 = -\sum_{\substack{(v_i, r_s, v_j) \in E \\ \wedge (v_i, r_t, v_k) \in E}} \omega_{ij} * \omega_{ik} * \log p_1(v_j^{r_s}, v_k^{r_t}|v_i)$$

$$O_2 = -\sum_{\substack{(v_j, r_s, v_i) \in E \\ \wedge (v_i, r_t, v_k) \in E}} \omega_{ji} * \omega_{ik} * \log p_2(v_j^{r_s}, v_k^{r_t}|v_i)$$

$$O_3 = -\sum_{\substack{(v_j, r_s, v_i) \in E \\ \wedge (v_k, r_t, v_i) \in E}} \omega_{ji} * \omega_{ki} * \log p_3(v_j^{r_s}, v_k^{r_t}|v_i)$$

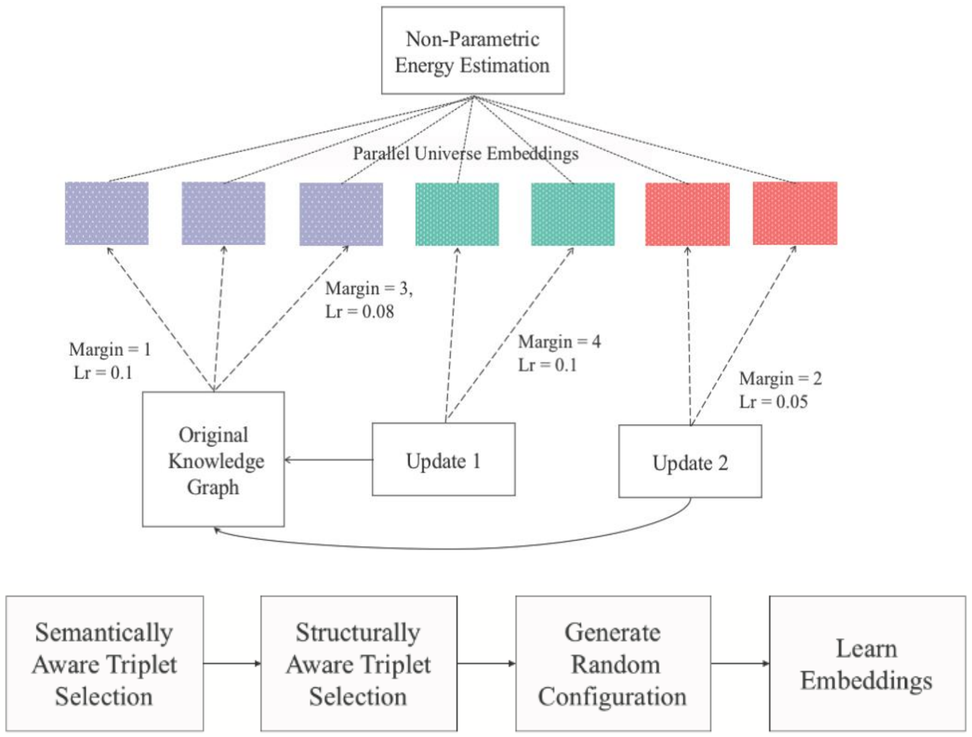
KB Embeddings . margin loss function, training

A mathematic perspective

$$\sum_{(h,r,t)\in\Delta} \sum_{(h',r,t')\in Delta'} \max(0, f_r(h,t) + M - f_r(h',t')))$$

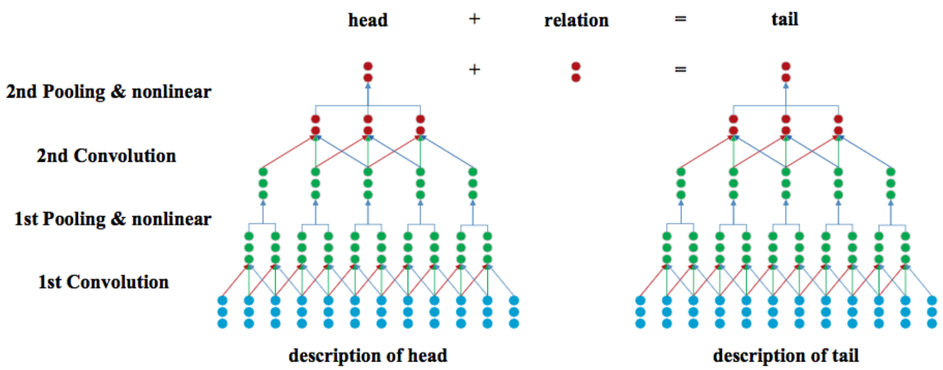
$$\sum_{(h,r,t)\in\Delta} \sum_{(h',r,t')\in Delta'} \max(0, f_r(h,t) + M_{opt} - f_r(h',t')))$$

Yantao Jia et al., 2016

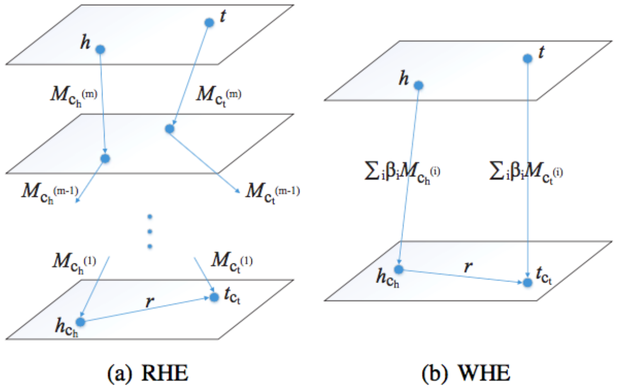


Yi Tay et al., 2017

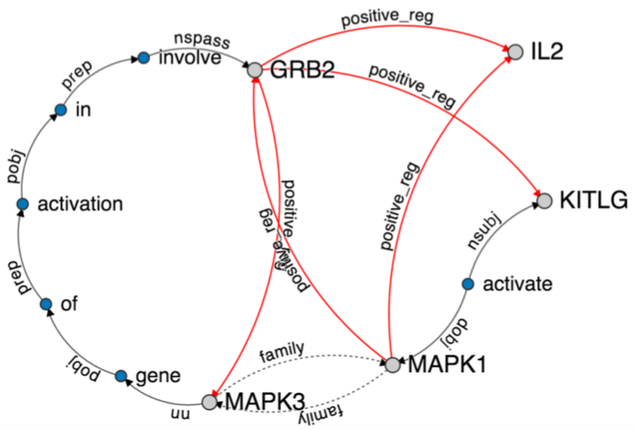
KB Embeddings . text information, lexical resources



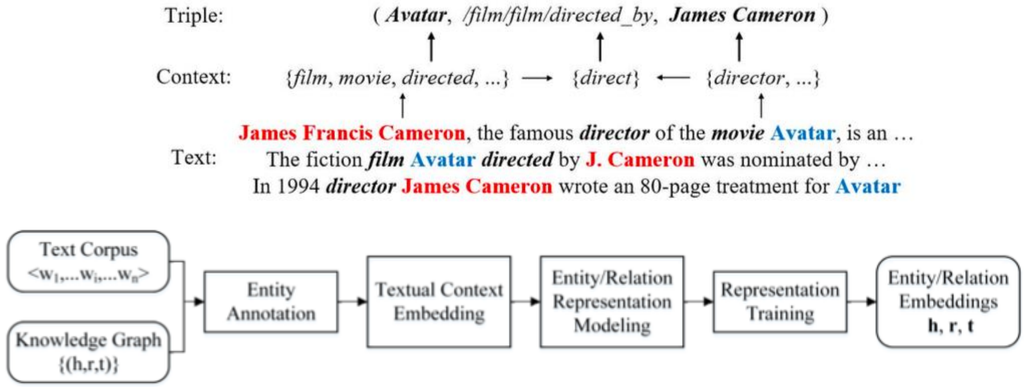
Ruobing Xie et al., 2016



Ruobing Xie et al., 2016

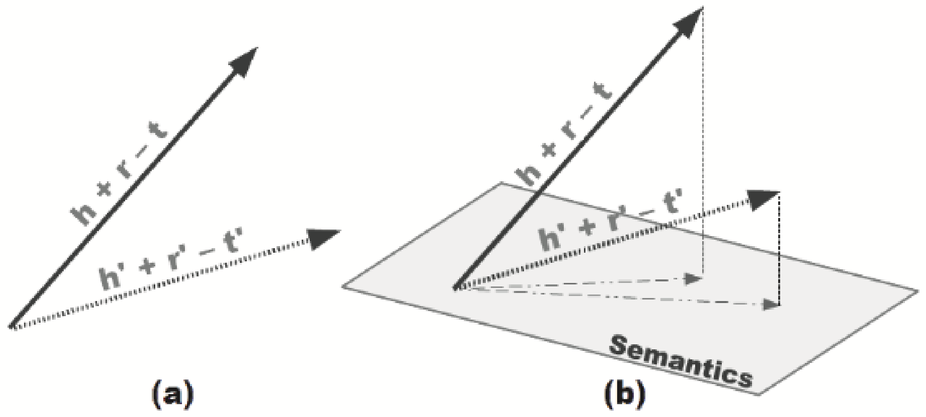


Toutanova et al., 2016

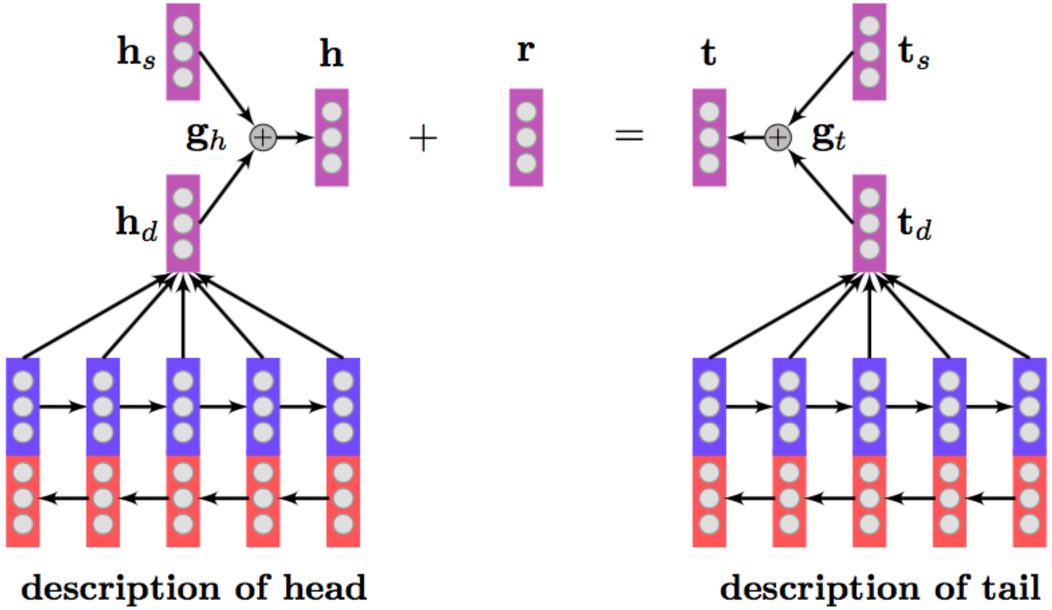


Zhigang Wang et al., 2016

KB Embeddings . text information, lexical resources



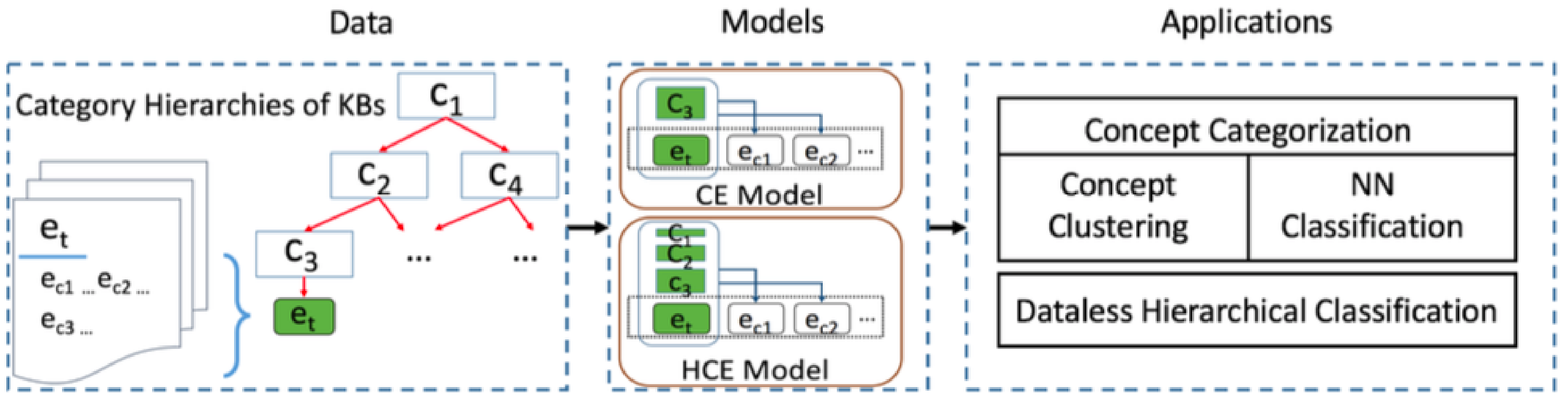
Han Xiao et al., 2017



Jiacheng Xu et al., 2017

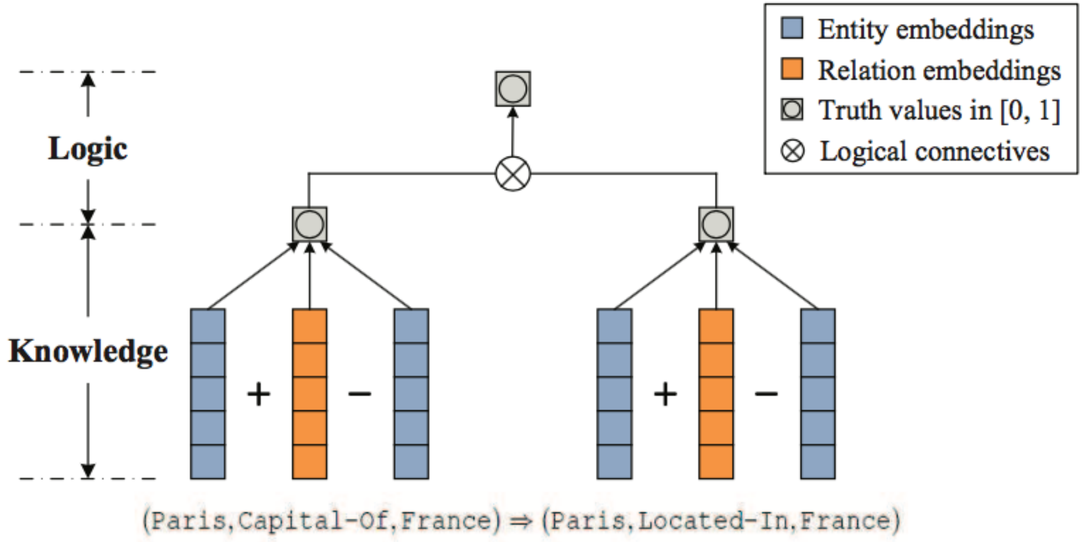
KB Embeddings . text information, lexical resources

- Initializing representations with entity descriptions (Teng Long et al., 2016)
- Learning distributed representations of both entities and categories (Yuezhang Li et al., 2016)



Yuezhang li et al., 2016

KB Embeddings . logical rules



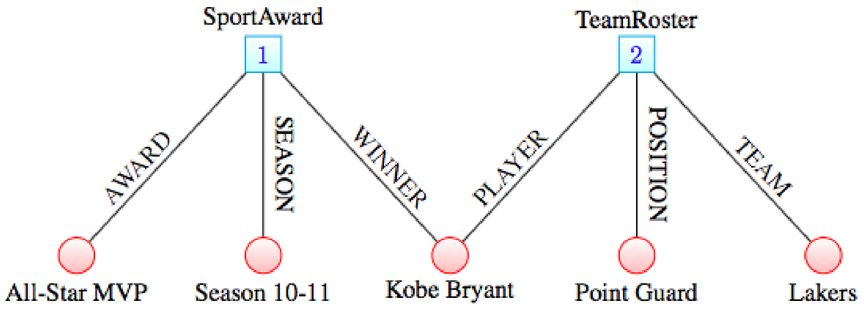
Shu Guo et al., 2016

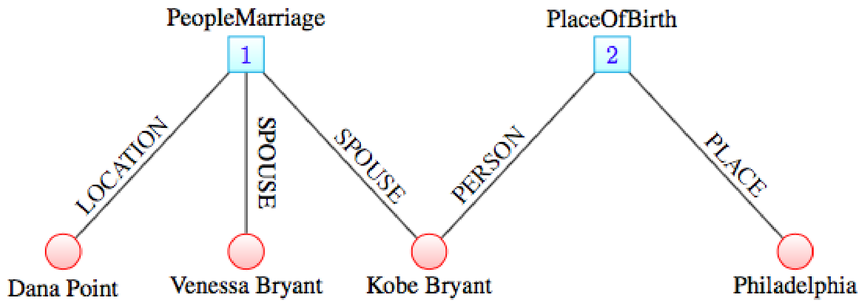
$$\forall t \in \mathcal{T} : \langle r_p, t \rangle \Rightarrow \langle r_q, t \rangle,$$

$$\forall t \in \mathcal{T} : \boldsymbol{r_p}^\top \boldsymbol{t} \leq \boldsymbol{r_q}^\top \boldsymbol{t}$$

Thomas Demeester et al., 2016

KB Embeddings . multi-fold relations





$$f_r(\mathbf{x}, \mathbf{y}) = \|\mathbb{P}_{\mathbf{n}_r}(\mathbf{x}) + \mathbf{d}_r - \mathbb{P}_{\mathbf{n}_r}(\mathbf{y})\|^2,$$

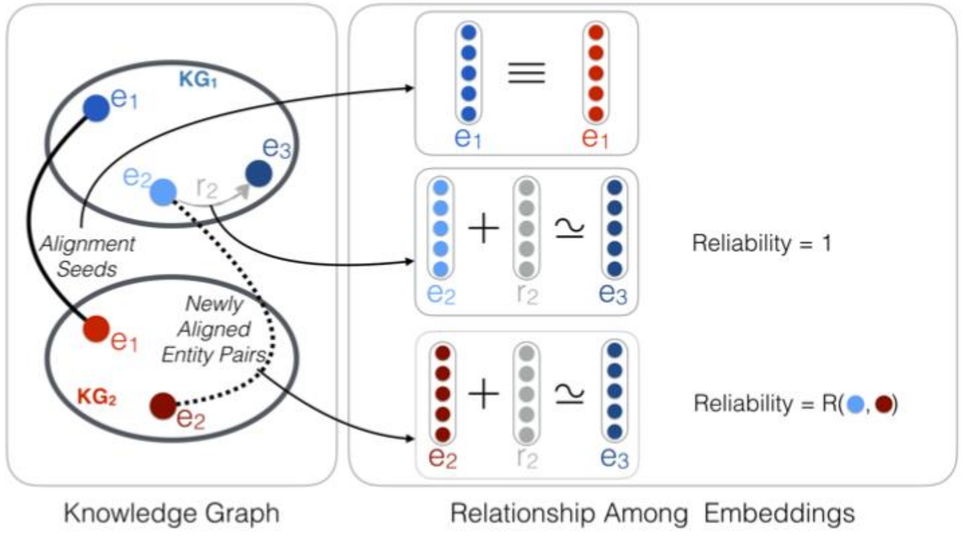
$$f_r(\mathbf{t}) := \left\| \sum_{\rho \in \mathcal{M}(R_r)} \mathbf{a}_r(\rho) \mathbb{P}_{\mathbf{n}_r}(\mathbf{t}(\rho)) + \mathbf{b}_r \right\|^2, \ \mathbf{t} \in \mathcal{N}^{\mathcal{M}(R_r)}.$$

Jianfeng Wen et al., 2017

KB Embeddings . knowledge fusion

$$S_K = \sum_{L \in \{L_i, L_j\}} \sum_{(h,r,t) \in G_L} \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|$$

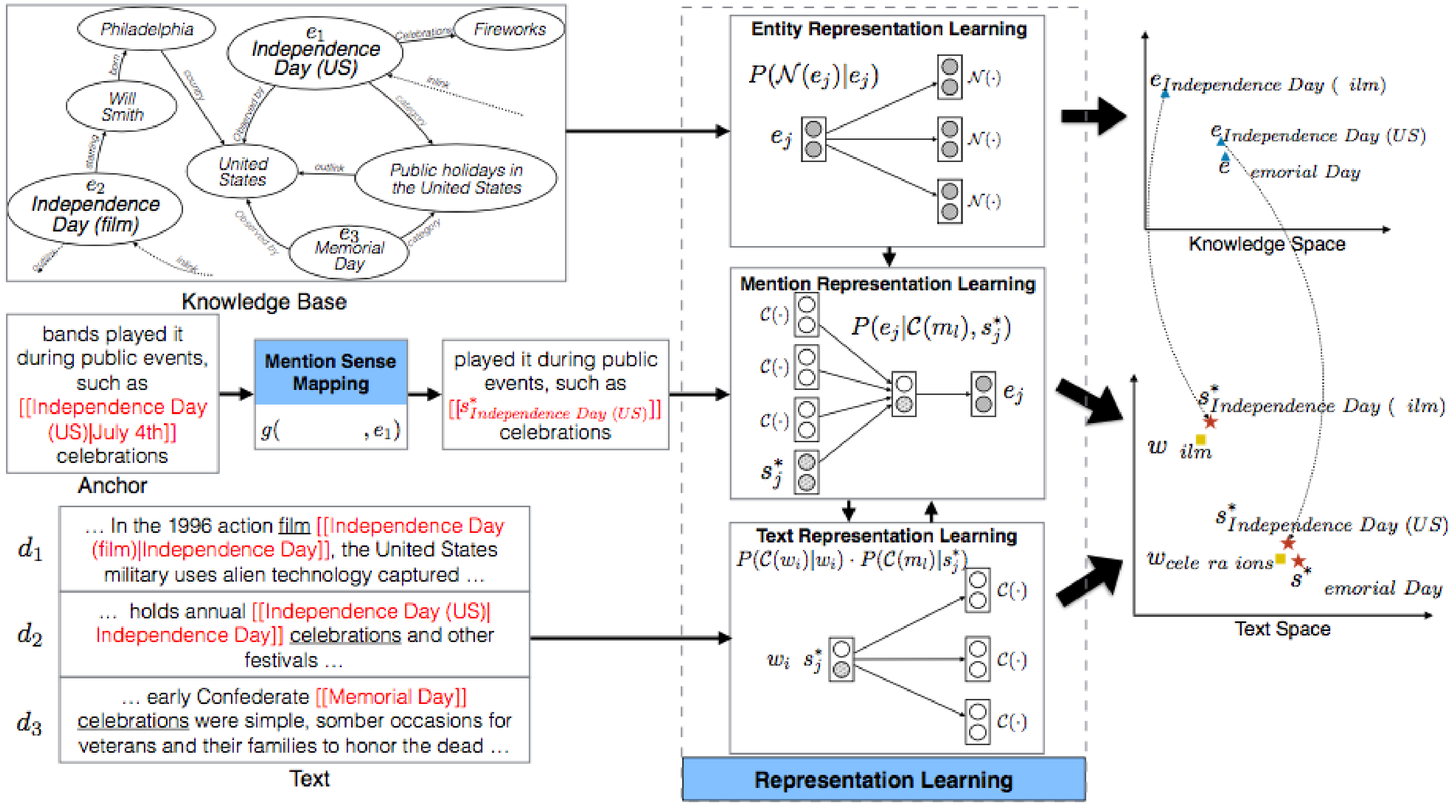
$$S_A = \sum_{(T,T') \in \delta(L_i, L_j)} S_a(T, T')$$



Muhao Chen et al., 2017

Hao Zhu et al., 2017

KB Embeddings . bridge text and KB



Yixin cao et al., 2017

KB Completion

Using Bayesian method to learn the latent structure of relational data. Including class-based and feature-based latent variable.

Including tensor and collective matrix factorization. Suitable for large-scale and sparse relational data, learning distributed representation of data.

Trained to assign low energies to plausible triplets of a multi- relational graph.

Non-Parametric Bayesian method	Kemp et al., 2006 Miller et al., 2009
Factorization method	Singh et al., 2008 Sutskever et al., 2009 Nickel et al., 2011
Energy-Based Frameworks	Bordes et al., 2011 Bordes et al., 2013 Socher et al., 2013

KB Completion

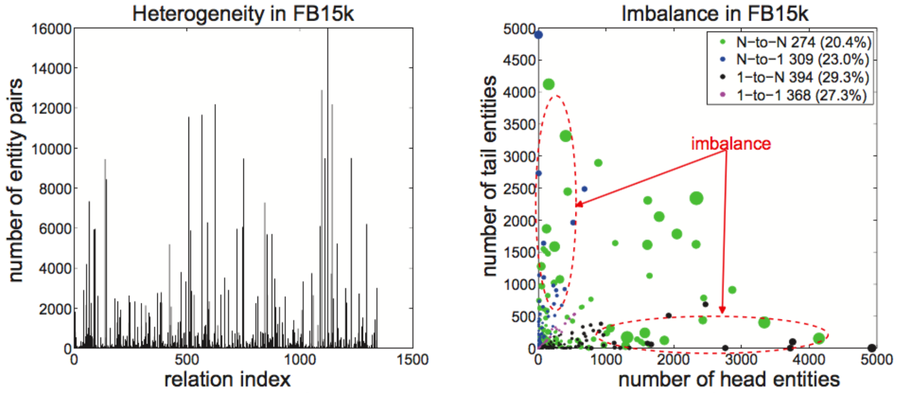
Automatically inferring missing facts by examining existing ones in KB.

Research Focus

- KB embeddings
- path ranking algorithms
  random walks
- Reinforcement Learning
- time-aware
- OOKB

KB Completion . KB embeddings

sparsity transfer model



(a) Heterogeneity (b) Imbalance

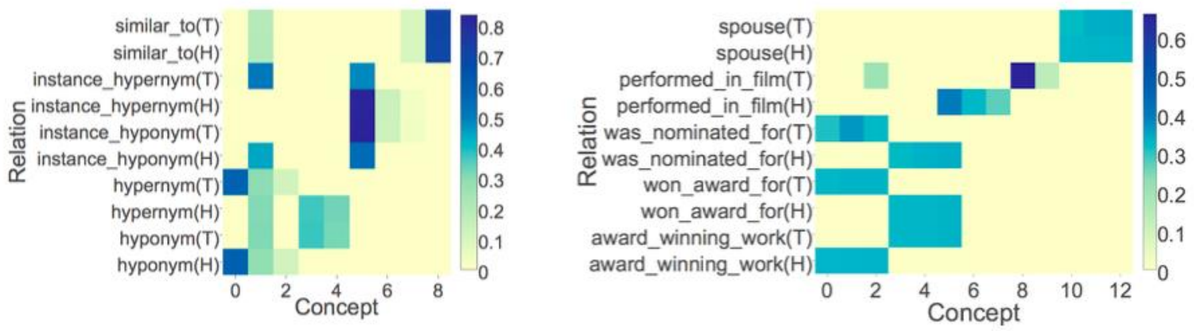
$$\mathbf{h}_p = \mathbf{M}_r^h(\theta_r^h)\mathbf{h}, \quad \mathbf{t}_p = \mathbf{M}_r^t(\theta_r^t)\mathbf{t}.$$

$$f_r(\mathbf{h},\mathbf{t}) = \|\mathbf{h}_p + \mathbf{r} - \mathbf{t}_p\|_{\ell_{1/2}}^2.$$

Guoliang Ji et al., 2016

$$f_r(h,t) = \|\boldsymbol{\alpha}_r^H \cdot \mathbf{D} \cdot \mathbf{h} + \mathbf{r} - \boldsymbol{\alpha}_r^T \cdot \mathbf{D} \cdot \mathbf{t}\|_\ell$$

$$\mathcal{L} = \sum_{\substack{(h,r,t)\sim P,\\ (h',r,t')\sim N}} \left[\gamma + f_r(h,t) - f_r(h',t')\right]_+$$

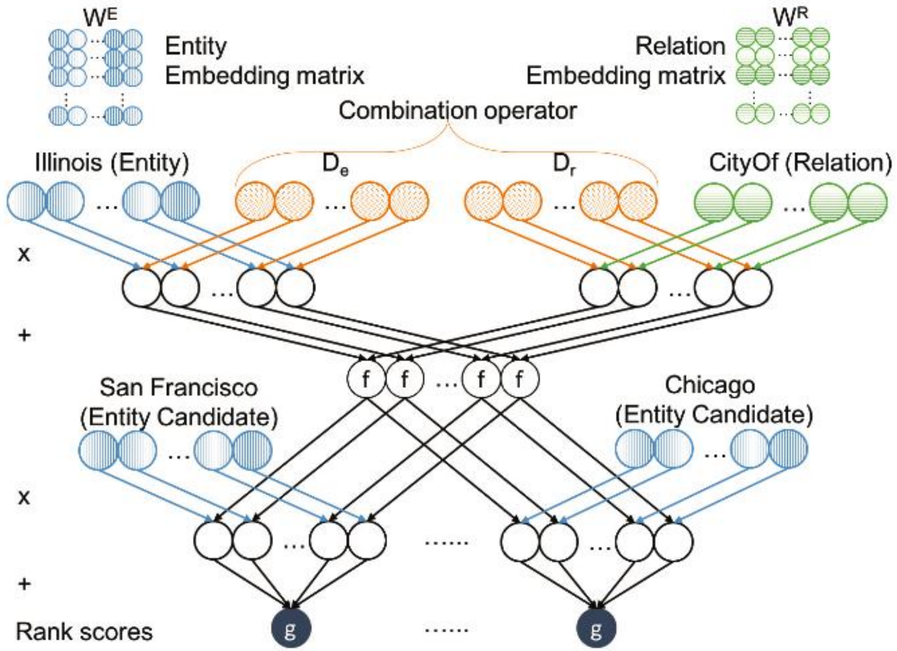


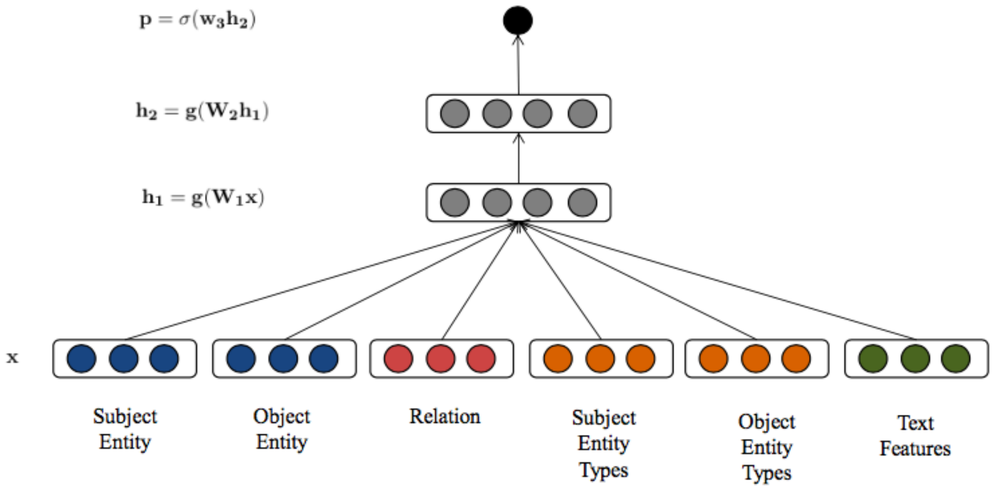
(a) WN18 (b) FB15k

Qizhe Xie et al., 2017

KB Completion . KB embeddings

neural model





Baoxu Shi et al., 2017

Alexandros Komninos et al., 2017

KB Completion . path ranking algorithms

Modeling certain relations in a collective way

Path Ranking Algorithm

$$\mathrm{Sim}(\mathcal{C}_i, \mathcal{C}_j) = \frac{|\mathbf{\Pi}_{\mathcal{C}_i} \cap \mathbf{\Pi}_{\mathcal{C}_j}|}{\min(|\mathbf{\Pi}_{\mathcal{C}_i}|, |\mathbf{\Pi}_{\mathcal{C}_j}|)},$$

Coupled Path Ranking Algorithm

Quan Wang et al., 2016

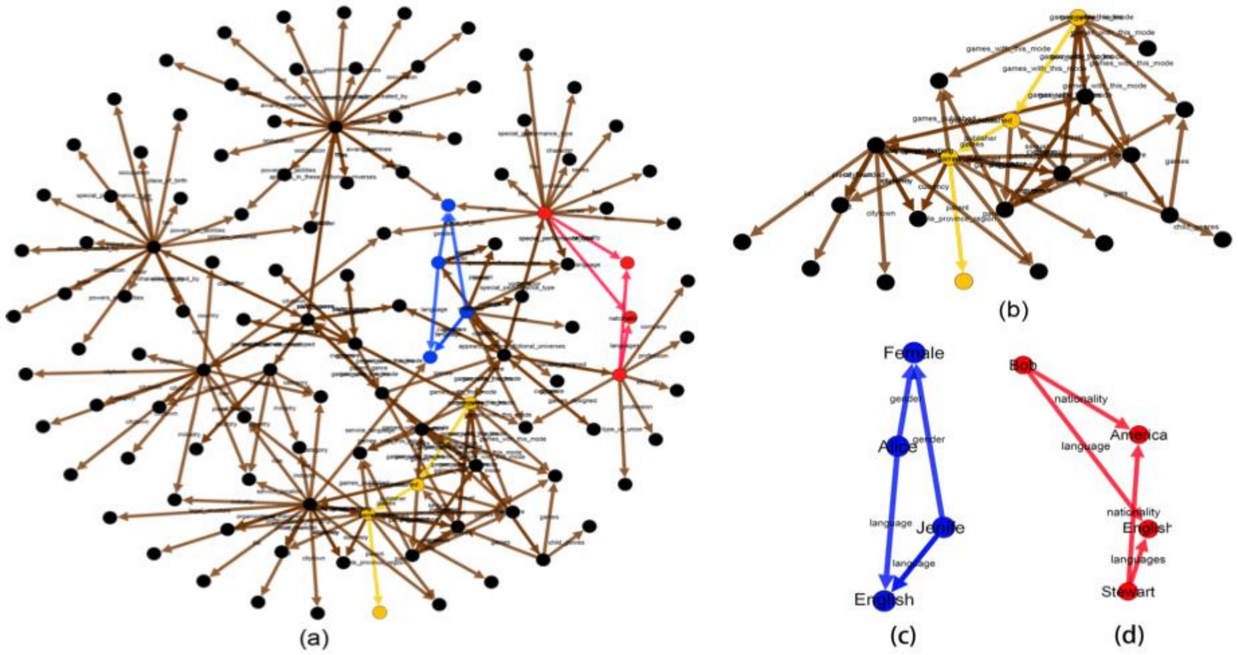
- scalability (high RAM consumption)
- feature explosion (large number of features)

Path Ranking Algorithm

Context-aware Path Ranking

Sahisnu Mazumder et al., 2017

KB Completion . path ranking algorithms



goal-directed random walks

$$P_{r_{i,j}} = \begin{cases} \frac{\Phi(r(i,j),\rho)}{\sum_{k \in Adj(i)} \Phi(r(i,j),\rho)}, & j \in Adj(i) \\ 0, & j \notin Adj(i) \end{cases}$$

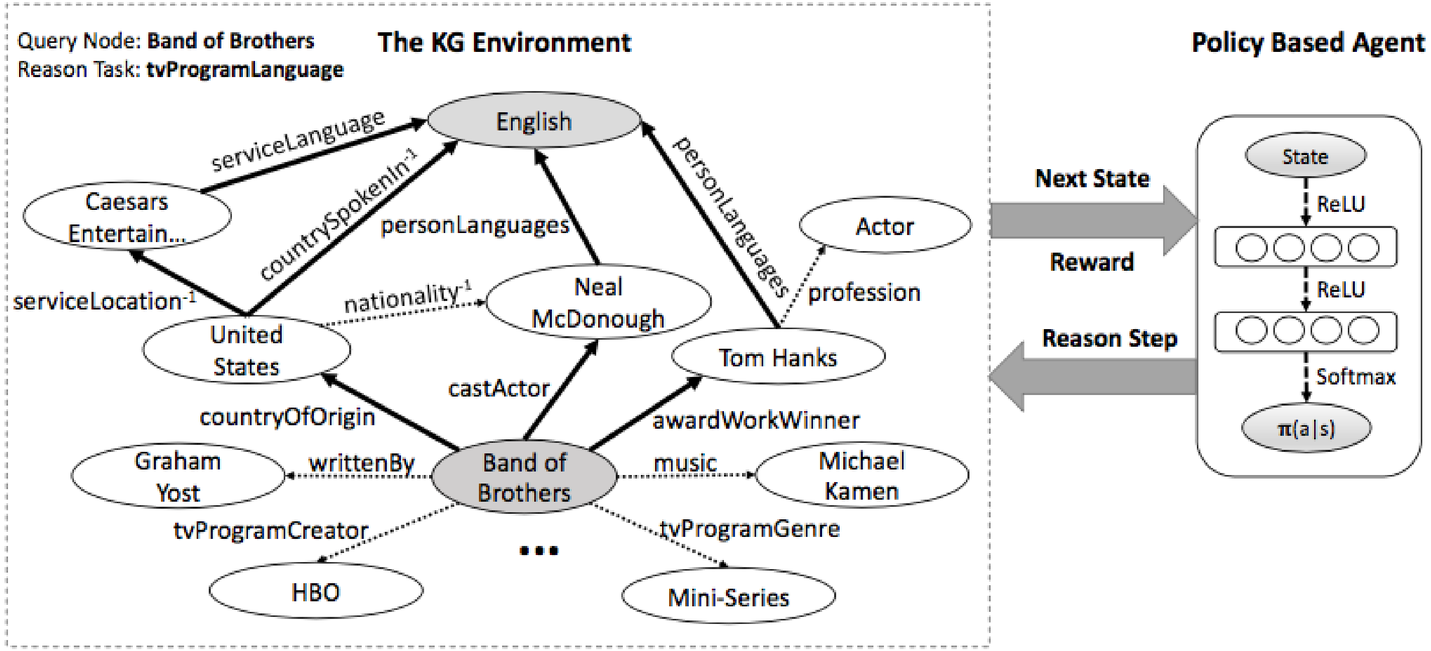
$$\mathrm{maxP}_{\mathbb{P}} = \prod_{\mathrm{pHt} \in \mathbb{P}} \mathrm{P}_{\mathrm{pHt}}^{\mathrm{a}} (1 - \mathrm{P}_{\mathrm{pHt}})^{\mathrm{b+c}}$$

Zhuoyu Wei et al., 2016

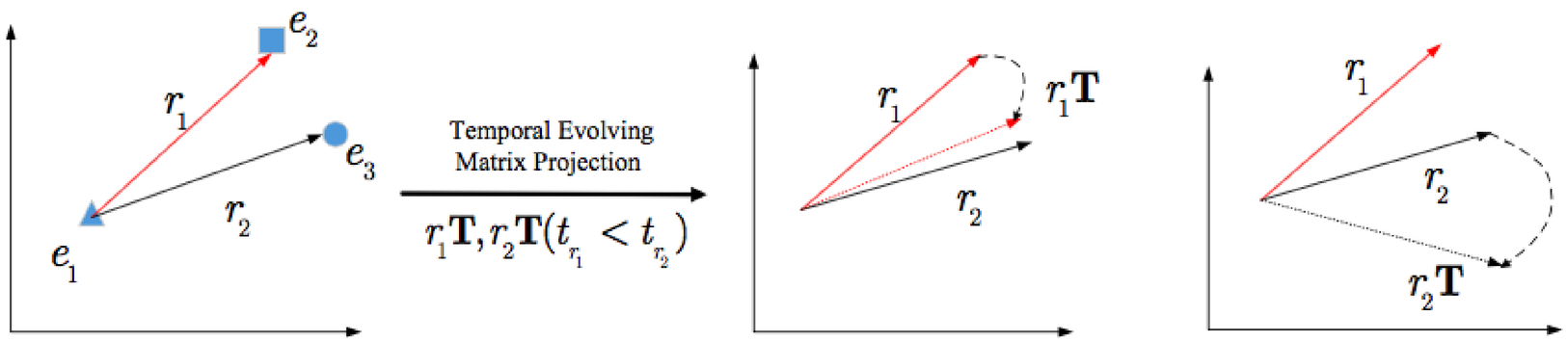
KB Completion . Reinforcement Learning

framing the path learning process as reinforcement learning

Wenhan Xiong et al., 2017



KB Completion . time-aware



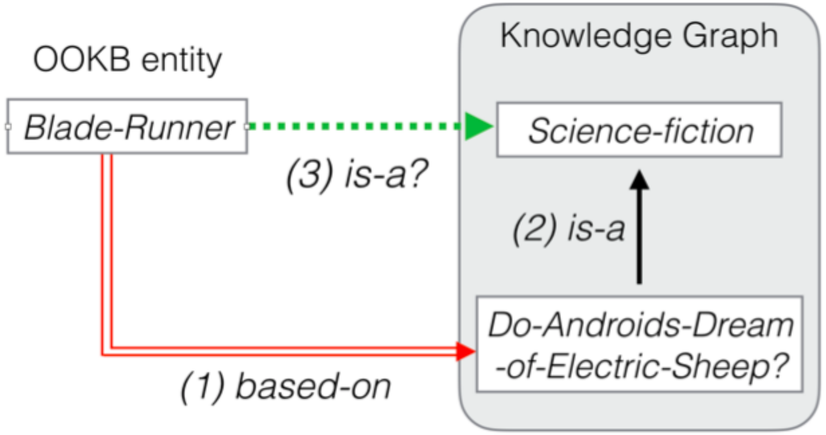
(a)TransE

(b)TransE-TAE

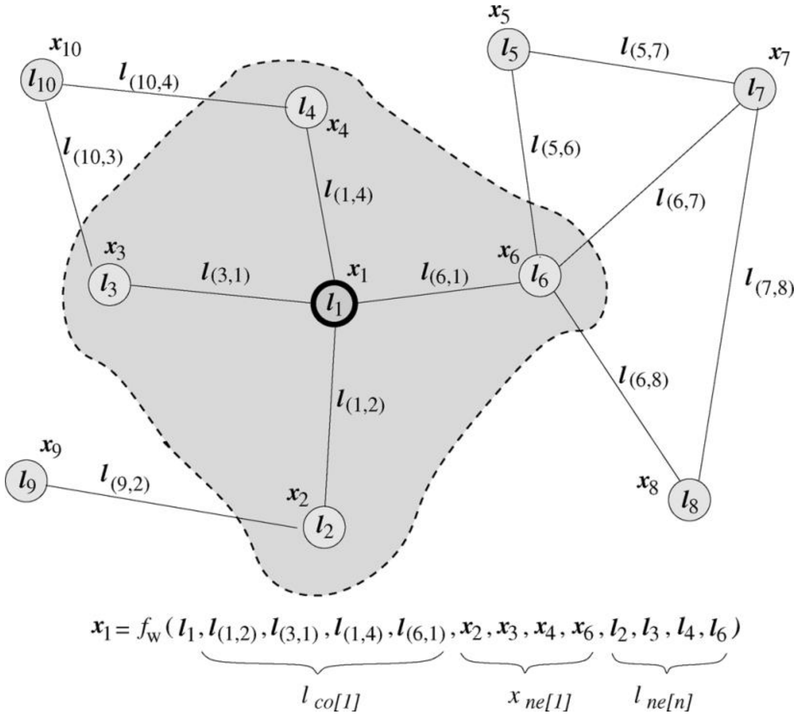
Testing quads	TransE	TransE-TAE	TransE-ILP
(Stanford_Moore,?,New_York_City,[1982,1982])	wasBornIn,**diedIn**	**diedIn**,wasBornIn	**diedIn**,wasBornIn
(John_Schoenherr,?,Caldecott_Medal,[1988,1988])	owns,**hasWonPrize**	**hasWonPrize**,created	**hasWonPrize**,created
(John_G._Thompson,?,University_of_Cambridge,[1968,1994])	graduatedFrom,**worksAt**	**worksAt**,graduatedFrom	**worksAt**,graduatedFrom
(Tommy_Douglas,?,New_Democratic_Party,[1961,1972])	isMarriedTo,**isAffiliatedTo**	**isAffiliatedTo**,isMarriedTo	**isAffiliatedTo**,isMarriedTo
(Carmen Electra,?,Owen Wilson,[2004,2005])	isMarriedTo,**sameAward_winner**	isMarriedTo,**sameAward_winner**	**sameAward_winner**,isMarriedTo

Tingsong Jiang et al., 2016

KB Completion . OOKB



Graph-NNs



Takuo Hamaguchi et al., 2017

Reference . Relation Classification

- Zheng H, Li Z, Wang S, et al. Aggregating Inter-Sentence Information to Enhance Relation Extraction[C]//AAAI. 2016: 3108-3115.
- Han X, Sun L. Global distant supervision for relation extraction[C]//Thirtieth AAAI Conference on Artificial Intelligence. 2016.
- Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2016, 2: 207-212.
- Cai R, Zhang X, Wang H. Bidirectional Recurrent Convolutional Neural Network for Relation Classification[C]//ACL (1). 2016.
- Takase S, Okazaki N, Inui K. Composing distributed representations of relational patterns[J]. arXiv preprint arXiv:1707.07265, 2017.
- Lin Y, Shen S, Liu Z, et al. Neural Relation Extraction with Selective Attention over Instances[C]//ACL (1). 2016.
- Grycner A, Weikum G. POLY: Mining Relational Paraphrases from Multilingual Sentences[C]//EMNLP. 2016: 2183-2192.

Reference . Relation Classification

- Wang L, Cao Z, de Melo G, et al. Relation Classification via Multi-Level Attention CNNs[C]//ACL (1). 2016.
- Wang C, Cao L, Fan J. Building Joint Spaces for Relation Extraction[C]//IJCAI. 2016: 2936-2942.
- Yang Y, Tong Y, Ma S, et al. A Position Encoding Convolutional Neural Network Based on Dependency Tree for Relation Classification[C]//EMNLP. 2016: 65-74.
- Shen Y, Huang X. Attention-Based Convolutional Neural Network for Semantic Relation Extraction[C]//COLING. 2016: 2526-2536.
- Xu Y, Jia R, Mou L, et al. Improved relation classification by deep recurrent neural networks with data augmentation[J]. arXiv preprint arXiv:1601.03651, 2016.
- Yu J, Jiang J. Pairwise relation classification with mirror instances and a combined convolutional neural network[J]. 2016.
- Jiang X, Wang Q, Li P, et al. Relation Extraction with Multi-instance Multi-label Convolutional Neural Networks[C]//COLING. 2016: 1471-1480.

Reference . Relation Classification

- Liu M X C. Semantic Relation Classification via Hierarchical Recurrent Neural Network with Attention[J]. 2016.
- Ji G, Liu K, He S, et al. Distant Supervision for Relation Extraction with Sentence-Level Attention and Entity Descriptions[C]//AAAI. 2017: 3060-3066.
- Ye H, Chao W, Luo Z. Jointly Extracting Relations with Class Ties via Effective Deep Ranking[J]. arXiv preprint arXiv:1612.07602, 2016.
- Luo B, Feng Y, Wang Z, et al. Learning with noise: enhance distantly supervised relation extraction with dynamic transition matrix[J]. arXiv preprint arXiv:1705.03995, 2017.
- Lin Y, Liu Z, Sun M. Neural Relation Extraction with Multi-lingual Attention[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017, 1: 34-43.
- Abad A, Nabi M, Moschitti A. Self-Crowdsourcing Training for Relation Extraction[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2017, 2: 518-523.

Reference . Relation Classification

- Feng X, Guo J, Qin B, et al. Effective Deep Memory Networks for Distant Supervised Relation Extraction[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI. 2017: 19-25.
- Liu T, Wang K, Chang B, et al. A Soft-label Method for Noise-tolerant Distantly Supervised Relation Extraction[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1791-1796.
- Wu Y, Bamman D, Russell S. Adversarial training for relation extraction[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1779-1784.
- Sorokin D, Gurevych I. Context-Aware Representations for Knowledge Base Relation Extraction[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1785-1790.
- Huang Y Y, Wang W Y. Deep Residual Learning for Weakly-Supervised Relation Extraction[J]. arXiv preprint arXiv:1707.08866, 2017.
- Liu L, Ren X, Zhu Q, et al. Heterogeneous Supervision for Relation Extraction: A Representation Learning Approach[J]. arXiv preprint arXiv:1707.00166, 2017.

Reference . Relation Classification

- Zeng W, Lin Y, Liu Z, et al. Incorporating relation paths in neural relation extraction[J]. arXiv preprint arXiv:1609.07479, 2016.
- Zhang Q, Wang H. Noise-Clustered Distant Supervision for Relation Extraction: A Nonparametric Bayesian Perspective[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1809-1814.

Reference . ERE Extraction

- Miwa M, Bansal M. End-to-end relation extraction using lstms on sequences and tree structures[J]. arXiv preprint arXiv:1601.00770, 2016.
- Kirschnick J, Hemsen H, Markl V. JEDI: Joint Entity and Relation Detection using Type Inference[J]. ACL 2016, 2016: 61.
- Nimishakavi M, Saini U S, Talukdar P. Relation schema induction using tensor factorization with side information[J]. arXiv preprint arXiv:1605.04227, 2016.
- Kim B, Lee J S. Extracting Spatial Entities and Relations in Korean Text[C]//COLING. 2016: 2389-2396.
- Gupta P, Schütze H, Andrassy B. Table Filling Multi-Task Recurrent Neural Network for Joint Entity and Relation Extraction[C]//COLING. 2016: 2537-2547.
- Katiyar A, Cardie C. Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017, 1: 917-928.
- Zheng S, Wang F, Bao H, et al. Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme[J]. arXiv preprint arXiv:1706.05075, 2017.

Reference . ERE Extraction

- Zhang M, Zhang Y, Fu G. End-to-End Neural Relation Extraction with Global Optimization[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1731-1741.
- Adel H, Schütze H. Global Normalization of Convolutional Neural Networks for Joint Entity and Relation Classification[J]. arXiv preprint arXiv:1707.07719, 2017.

Reference . KB Embeddings

- Nickel M, Rosasco L, Poggio T A. Holographic Embeddings of Knowledge Graphs[C]//AAAI. 2016: 1955-1961.
- Jia Y, Wang Y, Lin H, et al. Locally Adaptive Translation for Knowledge Graph Embedding[C]//AAAI. 2016: 992-998.
- Xie R, Liu Z, Jia J, et al. Representation Learning of Knowledge Graphs with Entity Descriptions[C]//AAAI. 2016: 2659-2665.
- Toutanova K, Lin V, Yih W, et al. Compositional Learning of Embeddings for Relation Paths in Knowledge Base and Text[C]//ACL (1). 2016.
- Long T, Lowe R, Cheung J C K, et al. Leveraging Lexical Resources for Learning Entity Embeddings in Multi-Relational Data[C]//The 54th Annual Meeting of the Association for Computational Linguistics. 2016: 112.
- Xiao H, Huang M, Zhu X. TransG: A Generative Model for Knowledge Graph Embedding[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016, 1: 2316-2325.

Reference . KB Embeddings

- Lin Y, Liu Z, Sun M. Knowledge representation learning with entities, attributes and relations[J]. ethnicity, 2016, 1: 41.52.
- Wen J, Li J, Mao Y, et al. On the representation and embedding of knowledge bases beyond binary relations[J]. arXiv preprint arXiv:1604.08642, 2016.
- Xie R, Liu Z, Sun M. Representation Learning of Knowledge Graphs with Hierarchical Types[C]//IJCAI. 2016: 2965-2971.
- Wang Z, Li J Z. Text-Enhanced Representation Learning for Knowledge Graph[C]//IJCAI. 2016: 1293-1299.
- Guo S, Wang Q, Wang L, et al. Jointly Embedding Knowledge Graphs and Logical Rules[C]//EMNLP. 2016: 192-202.
- Demeester T, Rocktäschel T, Riedel S. Lifted rule injection for relation embeddings[J]. arXiv preprint arXiv:1606.08359, 2016.
- Feng J, Huang M, Yang Y. GAKE: Graph aware knowledge embedding[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016: 641-651.

Reference . KB Embeddings

- Li Y, Zheng R, Tian T, et al. Joint embedding of hierarchical categories and entities for Concept Categorization and Dataless Classification[J]. arXiv preprint arXiv:1607.07956, 2016.
- Tay Y, Luu A T, Hui S C. Non-Parametric Estimation of Multiple Embeddings for Link Prediction on Dynamic Knowledge Graphs[C]//AAAI. 2017: 1243-1249.
- Xiao H, Huang M, Meng L, et al. SSP: Semantic Space Projection for Knowledge Graph Embedding with Text Descriptions[C]//AAAI. 2017: 3104-3110.
- Cao Y, Huang L, Ji H, et al. Bridging text and knowledge by learning multi-prototype entity mention embedding[C]//Proceedings of ACL. 2017.
- Liu L, Li X, Cheung W K, et al. A Structural Representation Learning for Multi-relational Networks[J].
- Zhu H, Xie R, Liu Z, et al. Iterative Entity Alignment via Joint Knowledge Embeddings[J].
- Xu J, Chen K, Qiu X, et al. Knowledge Graph Representation with Jointly Structural and Textual Encoding[J]. arXiv preprint arXiv:1611.08661, 2016.

Reference . KB Embeddings

- Chen M, Tian Y, Yang M, et al. Multi-lingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment[J]. arXiv preprint arXiv:1611.03954, 2016.
- Xiao H, Huang M, Zhu X. From one point to a manifold: Knowledge graph embedding for precise link prediction[J]. arXiv preprint arXiv:1512.04792, 2015.
- Ye H, Chao W, Luo Z. Jointly Extracting Relations with Class Ties via Effective Deep Ranking[J]. arXiv preprint arXiv:1612.07602, 2016.

Reference . KB Completion

- Ji G, Liu K, He S, et al. Knowledge Graph Completion with Adaptive Sparse Transfer Matrix[C]//AAAI. 2016: 985-991.
- Wang Q, Liu J, Luo Y, et al. Knowledge Base Completion via Coupled Path Ranking[C]//ACL (1). 2016.
- Wei Z, Zhao J, Liu K. Mining Inference Formulas by Goal-Directed Random Walks[C]//EMNLP. 2016: 1379-1388.
- Jiang T, Liu T, Ge T, et al. Towards Time-Aware Knowledge Graph Completion[C]//COLING. 2016: 1715-1724.
- Shi B, Weninger T. ProjE: Embedding Projection for Knowledge Graph Completion[C]//AAAI. 2017: 1236-1242.
- Xie Q, Ma X, Dai Z, et al. An Interpretable Knowledge Transfer Model for Knowledge Base Completion[J]. arXiv preprint arXiv:1704.05908, 2017.
- Komninos A, Manandhar S. Feature-Rich Networks for Knowledge Base Completion[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2017, 2: 324-329.

Reference . KB Completion

- Hayashi K, Shimbo M. On the Equivalence of Holographic and Complex Embeddings for Link Prediction[J]. arXiv preprint arXiv:1702.05563, 2017.
- Mazumder S, Liu B. Context-aware Path Ranking for Knowledge Base Completion[J].
- Hamaguchi T, Oiwa H, Shimbo M, et al. Knowledge Transfer for Out-of-Knowledge-Base Entities: A Graph Neural Network Approach[J]. arXiv preprint arXiv:1706.05674, 2017.
- Xiong W, Hoang T, Wang W Y. DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning[J]. arXiv preprint arXiv:1707.06690, 2017.