中国中文信息学会
第十三届暑期学校 & 前沿技术讲习班

语义表示学习

刘知远
清华大学自然语言处理实验室
liuzy@tsinghua.edu.cn

深度学习 = 数据表示 + 网络框架 +学习优化

Yoshua Bengio. Deep Learning of Representations. AAAI 2013 Tutorial.

词汇表示代表方案

- 1-hot representation: basis of Bag-of-Word Model

star [0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, ⋯]
sun [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, ⋯]

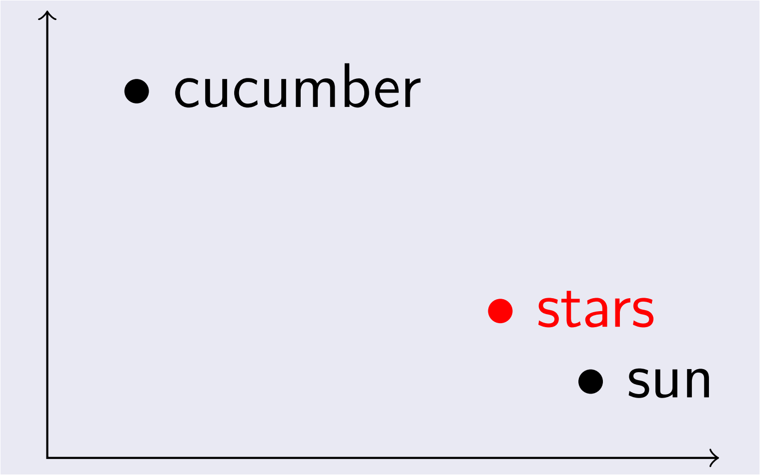
sim(star, sun) = 0 ☹

词汇表示代表方案

- Count-based distributional representation

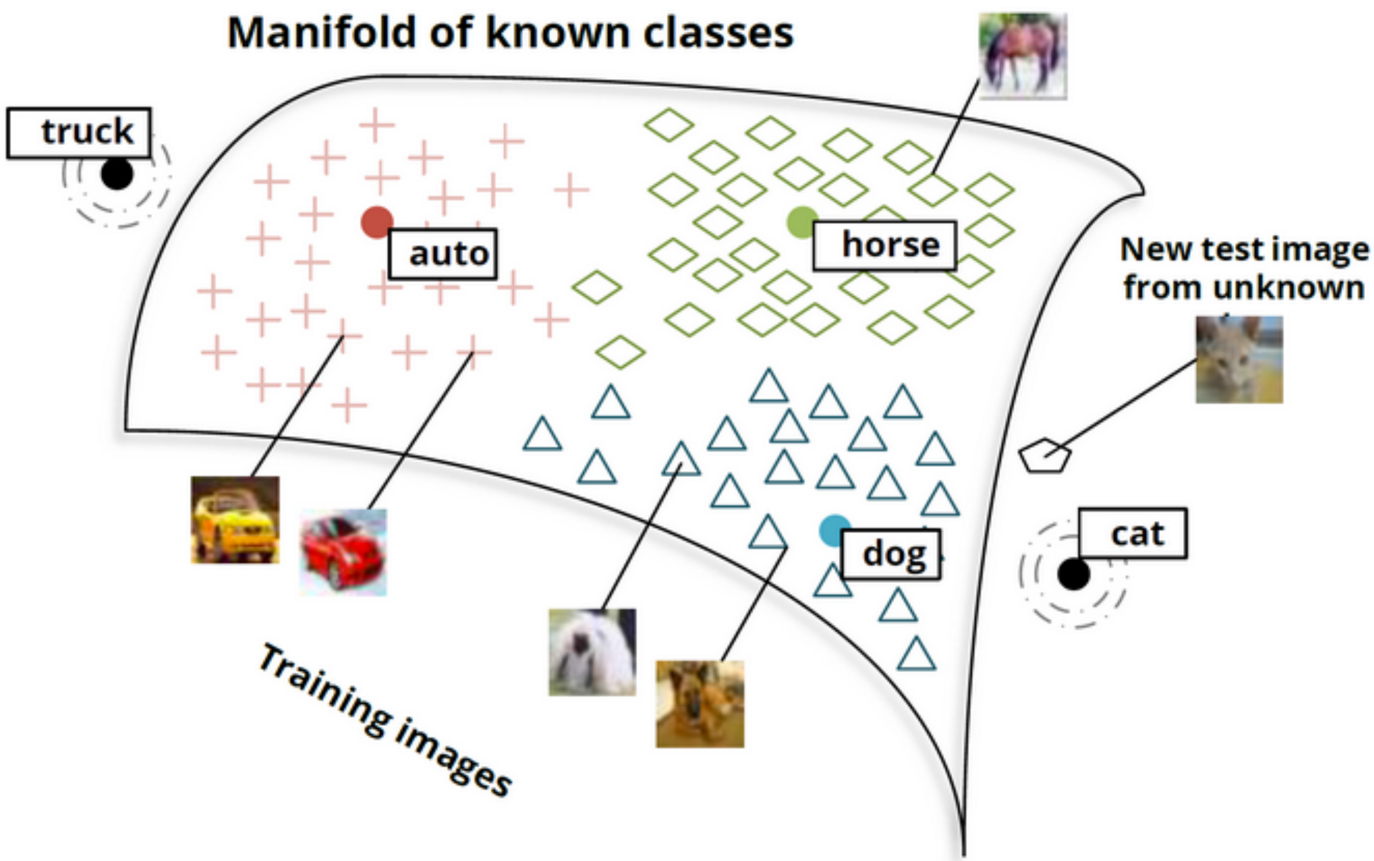
```
he curtains open and the stars shining in on the barely
ars and the cold , close stars " . And neither of the w
rough the night with the stars shining so brightly , it
made in the light of the stars . It all boils down , wr
 surely under the bright stars , thrilled by ice-white
sun , the seasons of the stars ? Home , alone , Jay pla
m is dazzling snow , the stars have risen full and cold
```

	shining	bright	trees	dark	look
stars	38	45	2	27	12



词汇表示最新方案

- Distributed Representation (Word Embeddings)
- 每个词被表示成稠密、实值、低维向量



表示学习的理论基础

- 类脑学习机制

- 人脑的学习能力
  - 信号传输慢 vs 计算速度快

  - 处理能力强 vs 能量消耗低

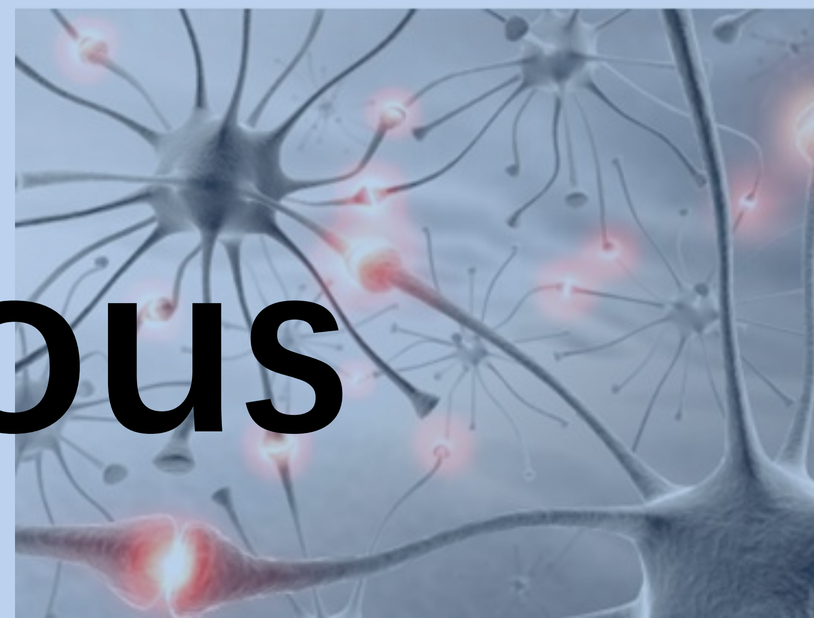
表示学习的理论基础

真实世界

Discrete

认知世界

Continuous



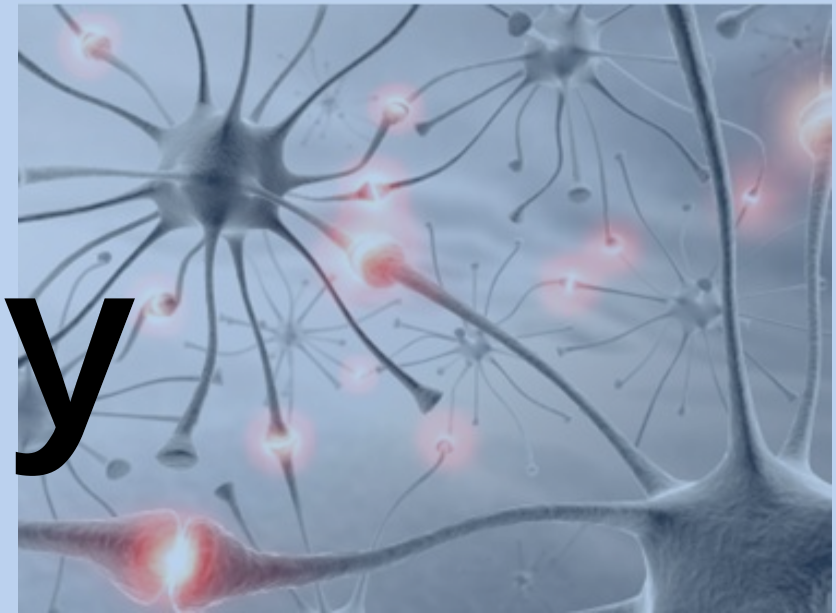
表示学习的理论基础

真实世界

Hierarchy

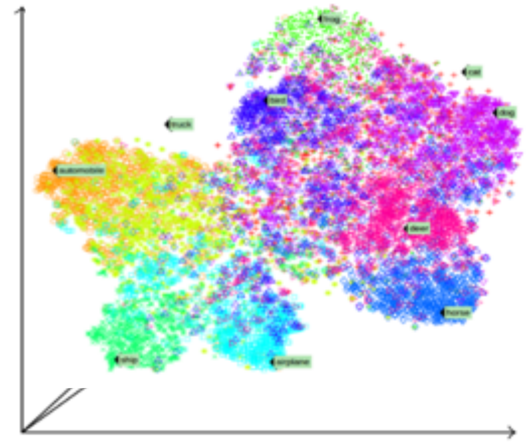
认知世界

Hierarchy

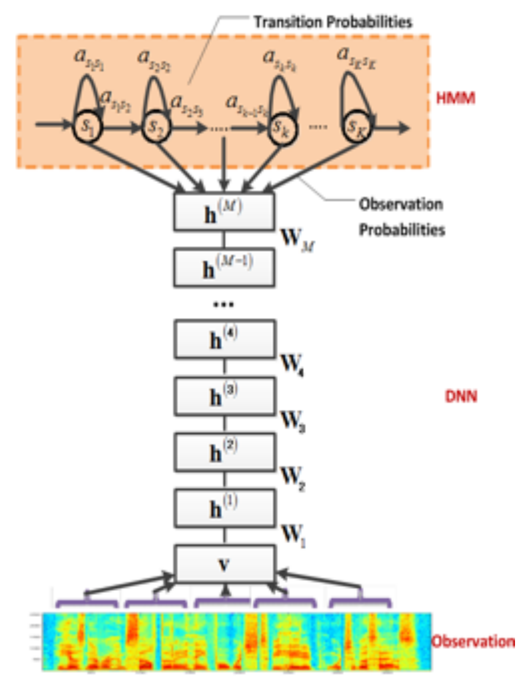


表示学习的主要思想

- 分布式（Distributed）表示
  - 嵌入（Embeddings）
  - 稠密、实值、低维向量

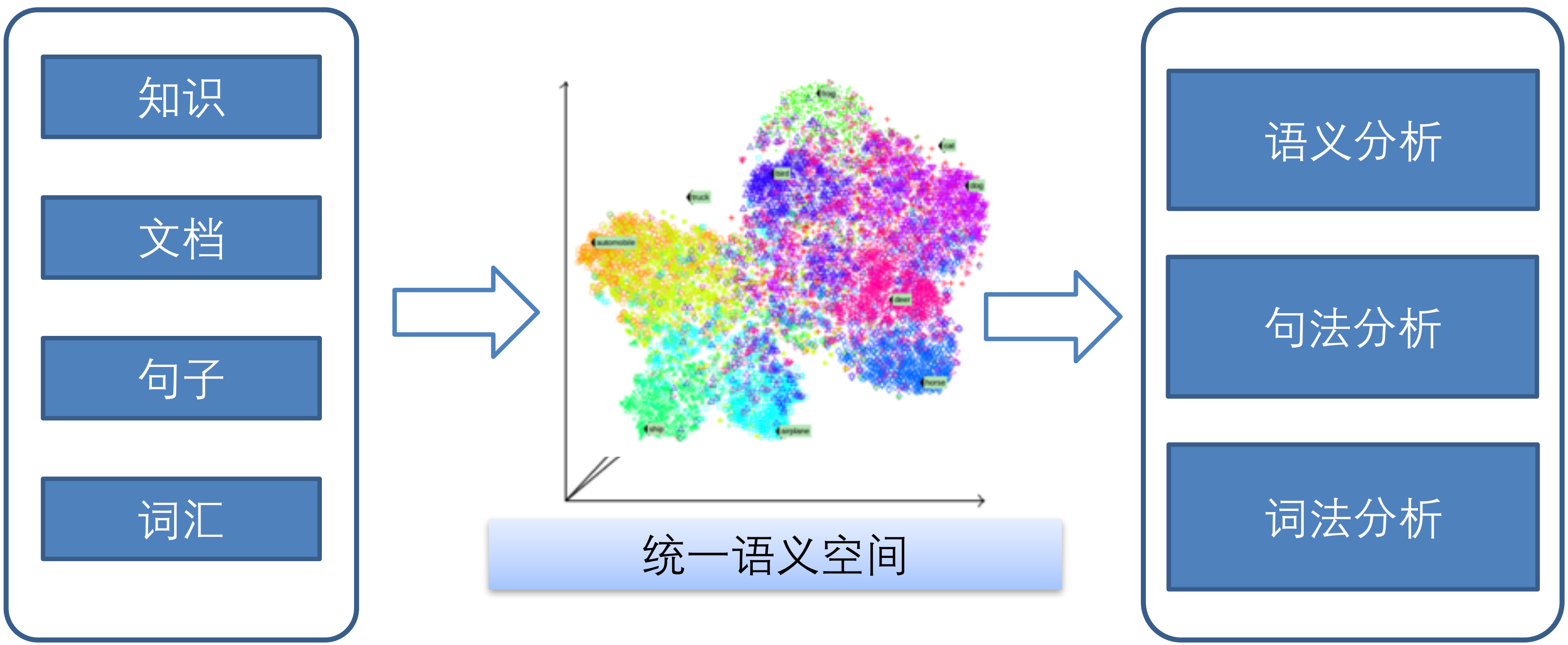


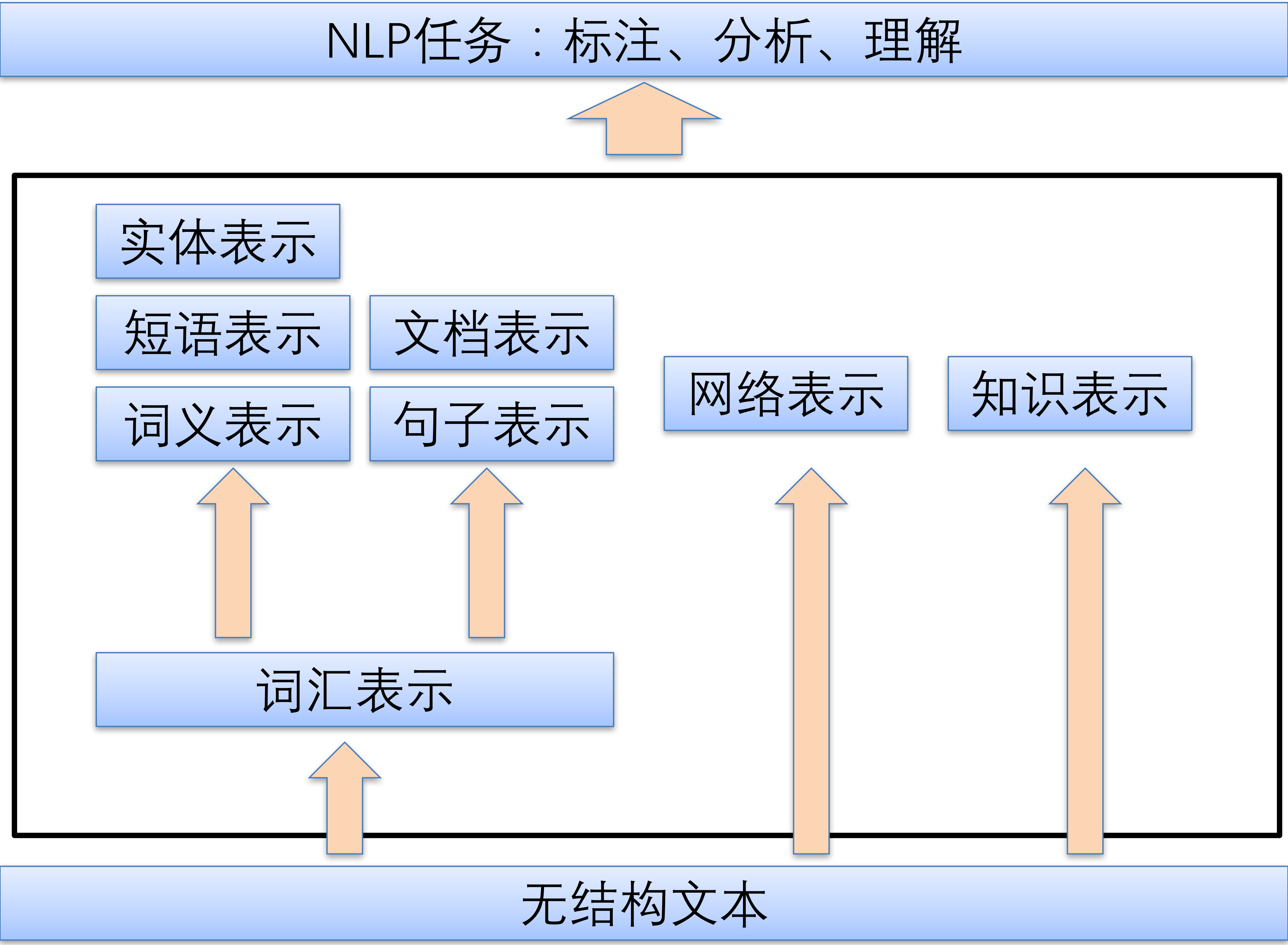
- 层次（ Hierarchical ）结构
  - 对应层次的真实世界
  - 具有抽象和泛化能力

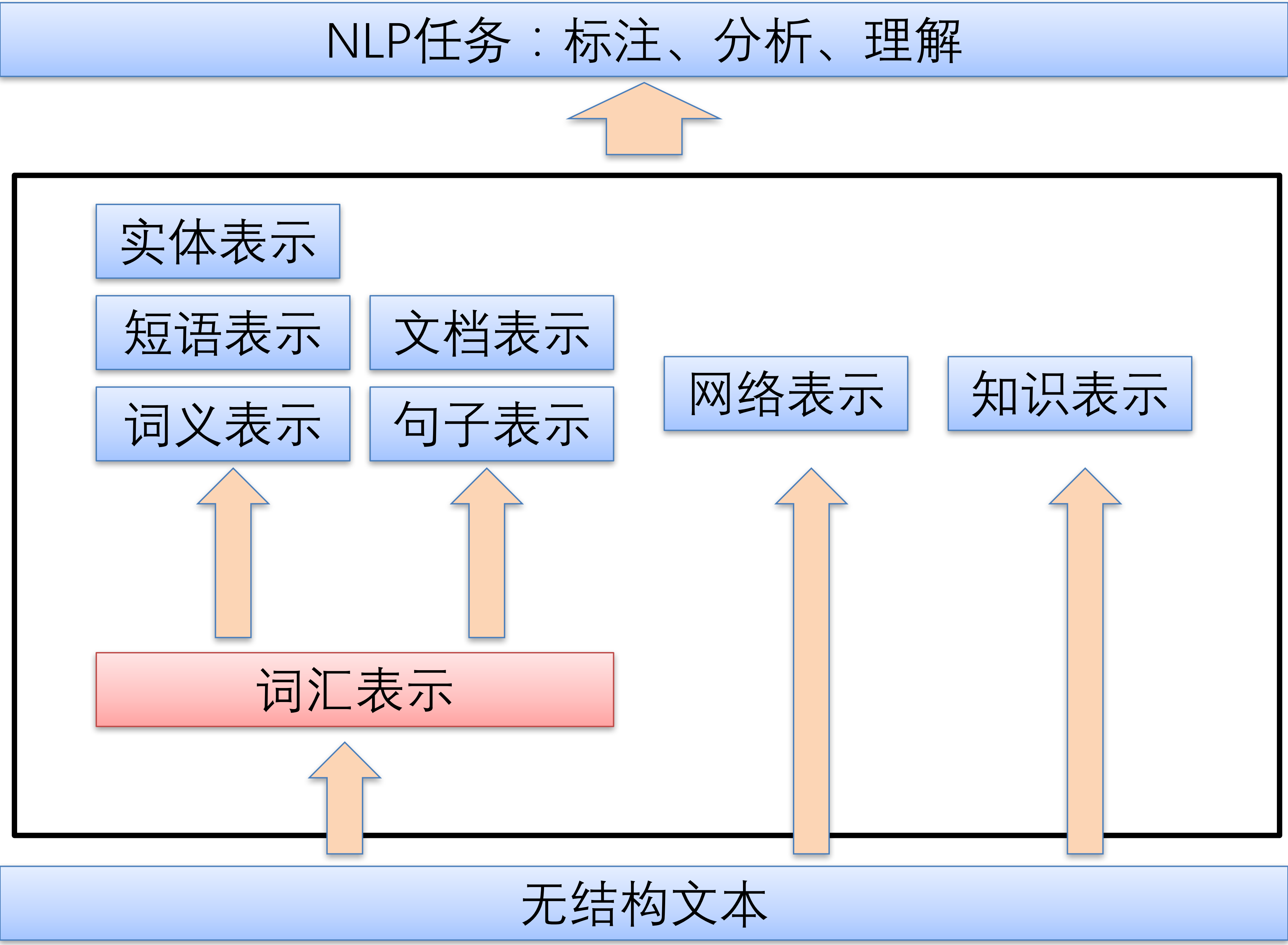


分布式表示的优势

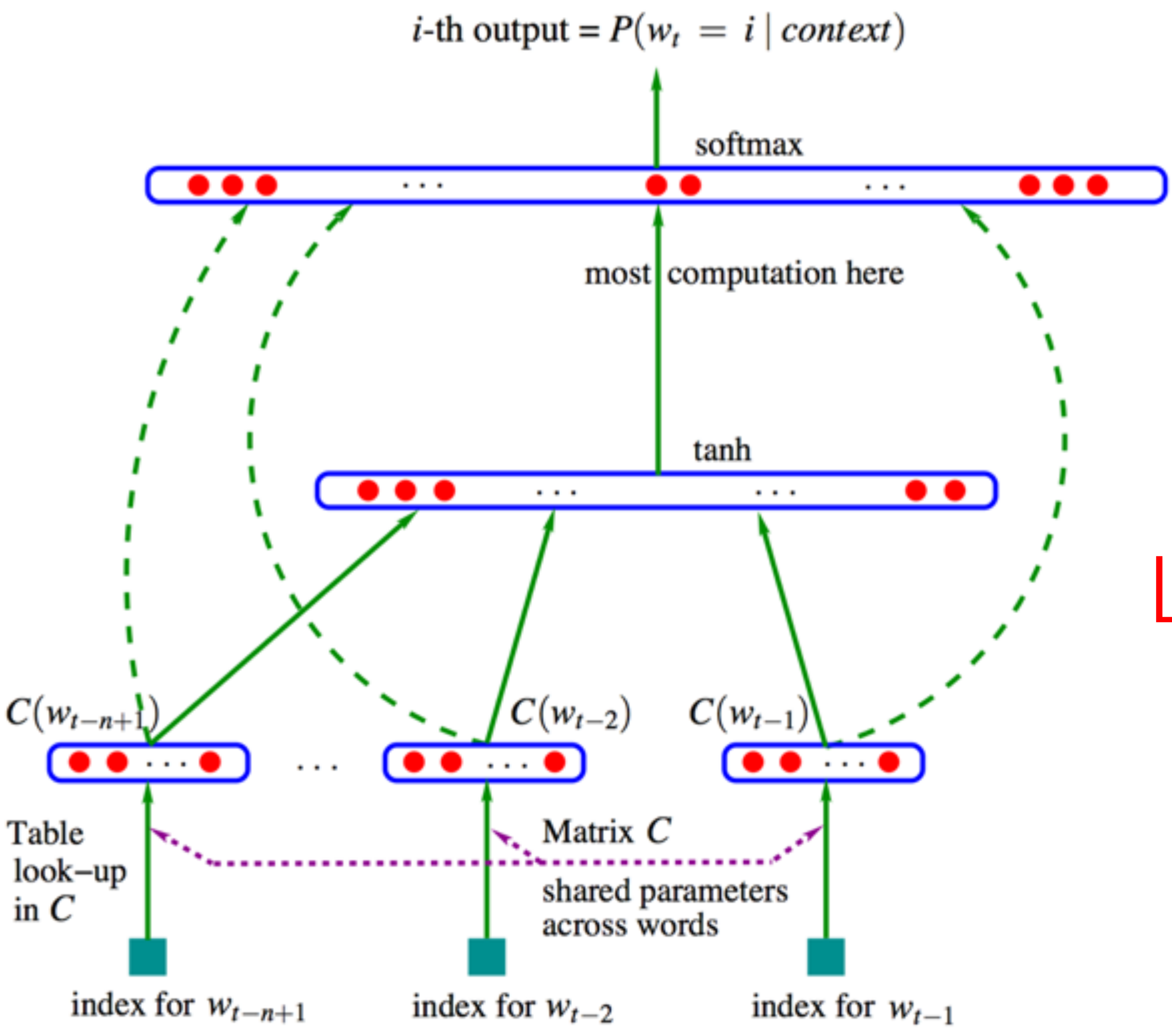
- 解决大数据NLP的数据稀疏问题
- 实现跨领域、跨对象的知识迁移
- 提供多任务学习的统一底层表示




NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
词义表示
句子表示
网络表示
知识表示
词汇表示
无结构文本


NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
网络表示
知识表示
词义表示
句子表示
词汇表示
无结构文本

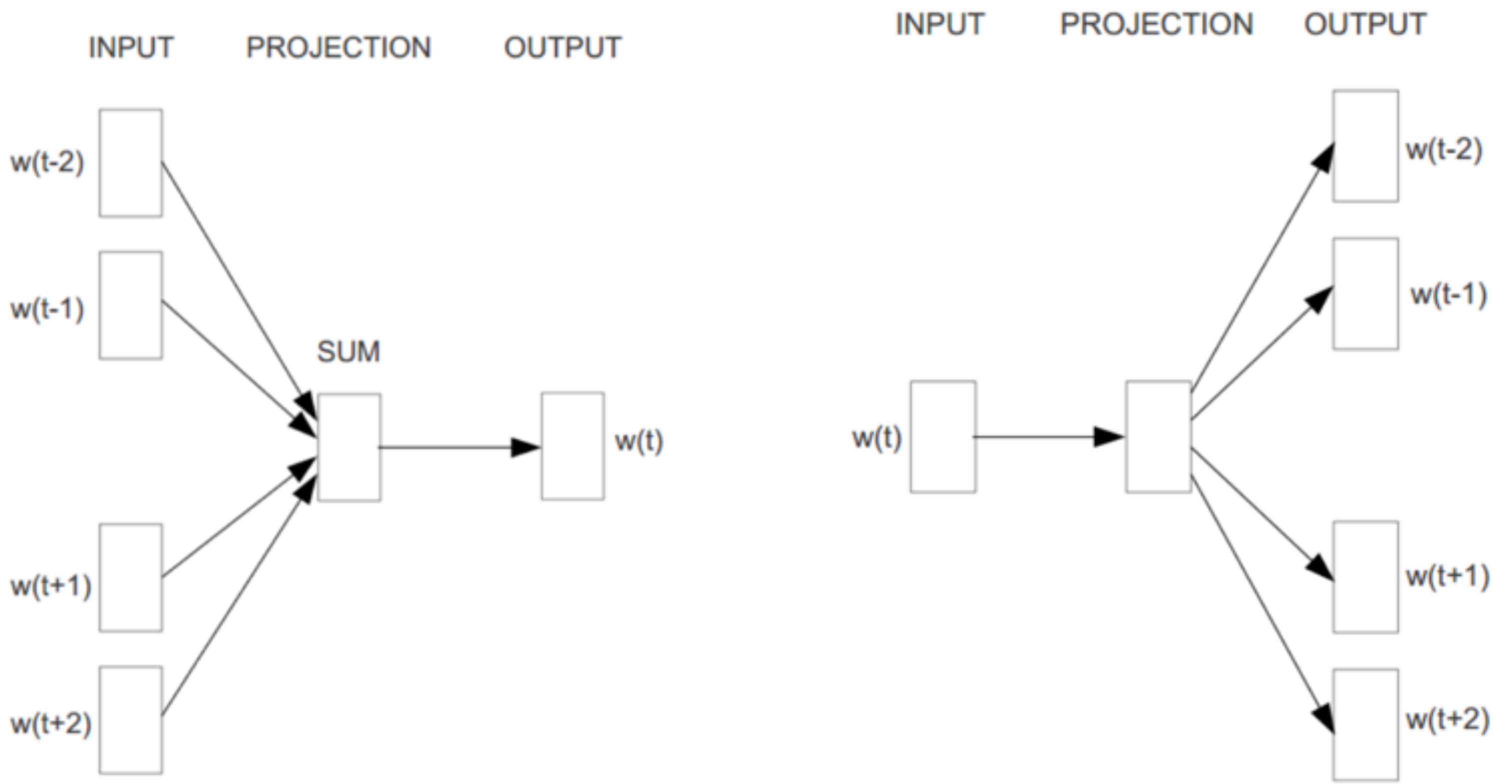
分布式表示代表模型



Neural Language Model

Yoshua Bengio. A neural probabilistic language model. JMLR 2003.

分布式表示代表模型



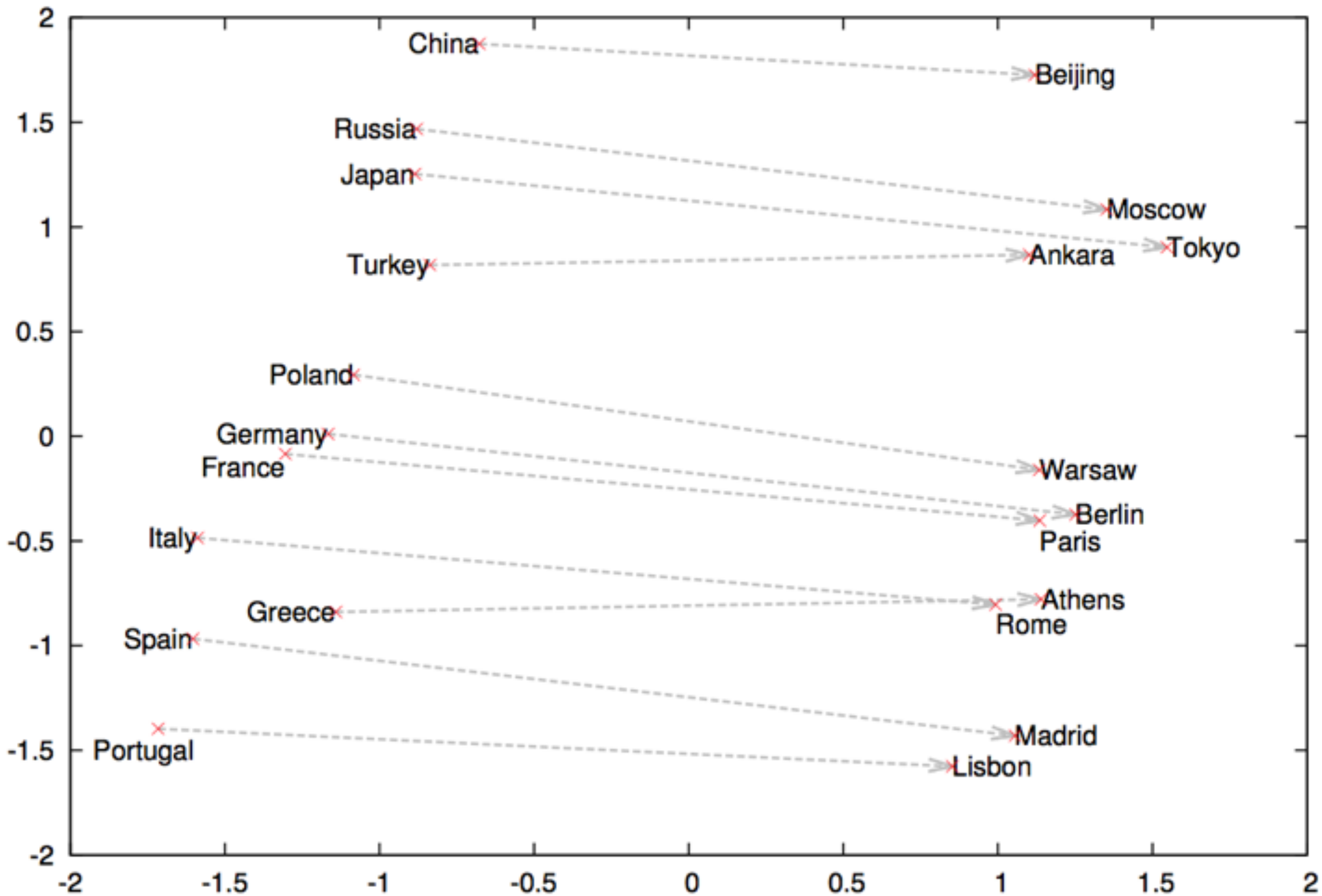
word2vec

Tomas Mikolov et al. Distributed representations of words and phrases and their compositionality. NIPS 2013.

词汇表示用于词汇相似度计算

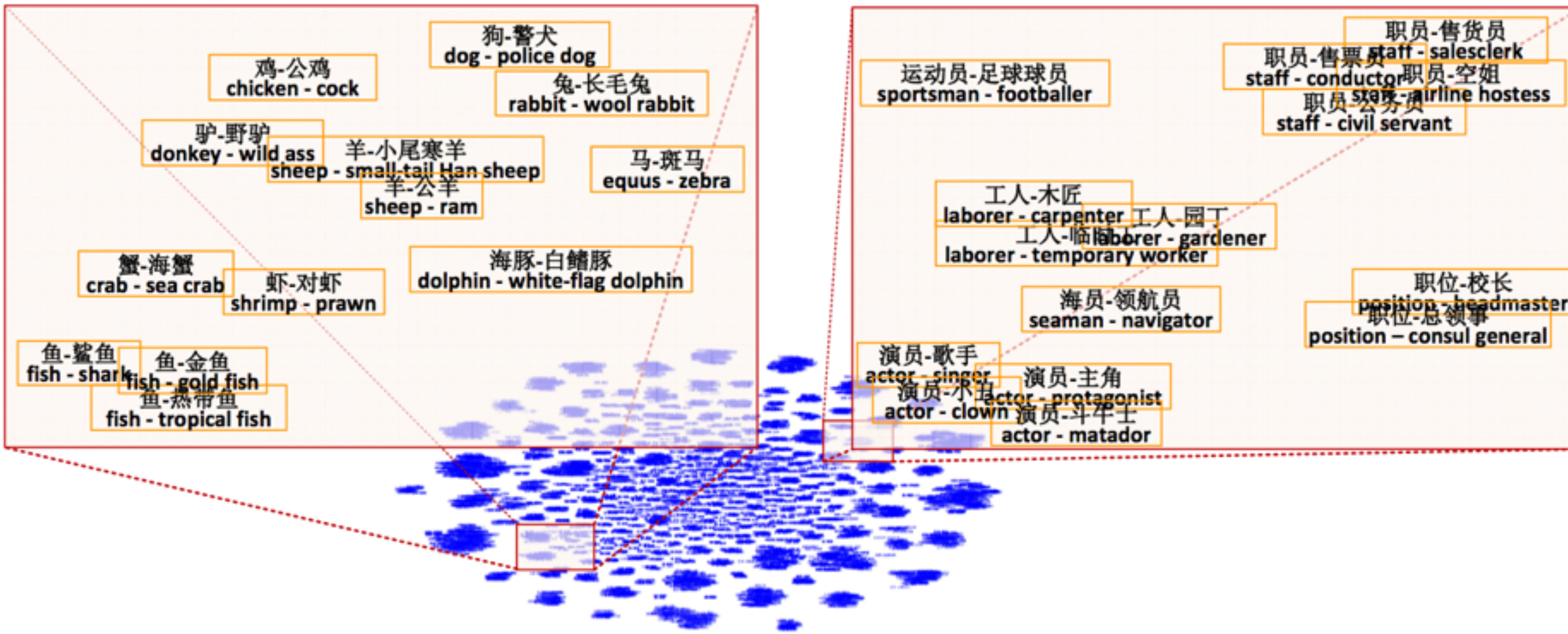
```
 (EXIT to break): china

.n vocabulary: 486

                          Word       Cosine distance
----------------------------------------------------
                        taiwan              0.768188
                         japan              0.652825
                         macau              0.614888
                         korea              0.614887
                           prc              0.613579
                       beijing              0.605946
                        taipei              0.592367
                      thailand              0.577905
                      cambodia              0.575681
                     singapore              0.569950
                      republic              0.567597
                      mongolia              0.554642
                       chinese              0.551576
```

词汇表示发现词汇间的隐含关系



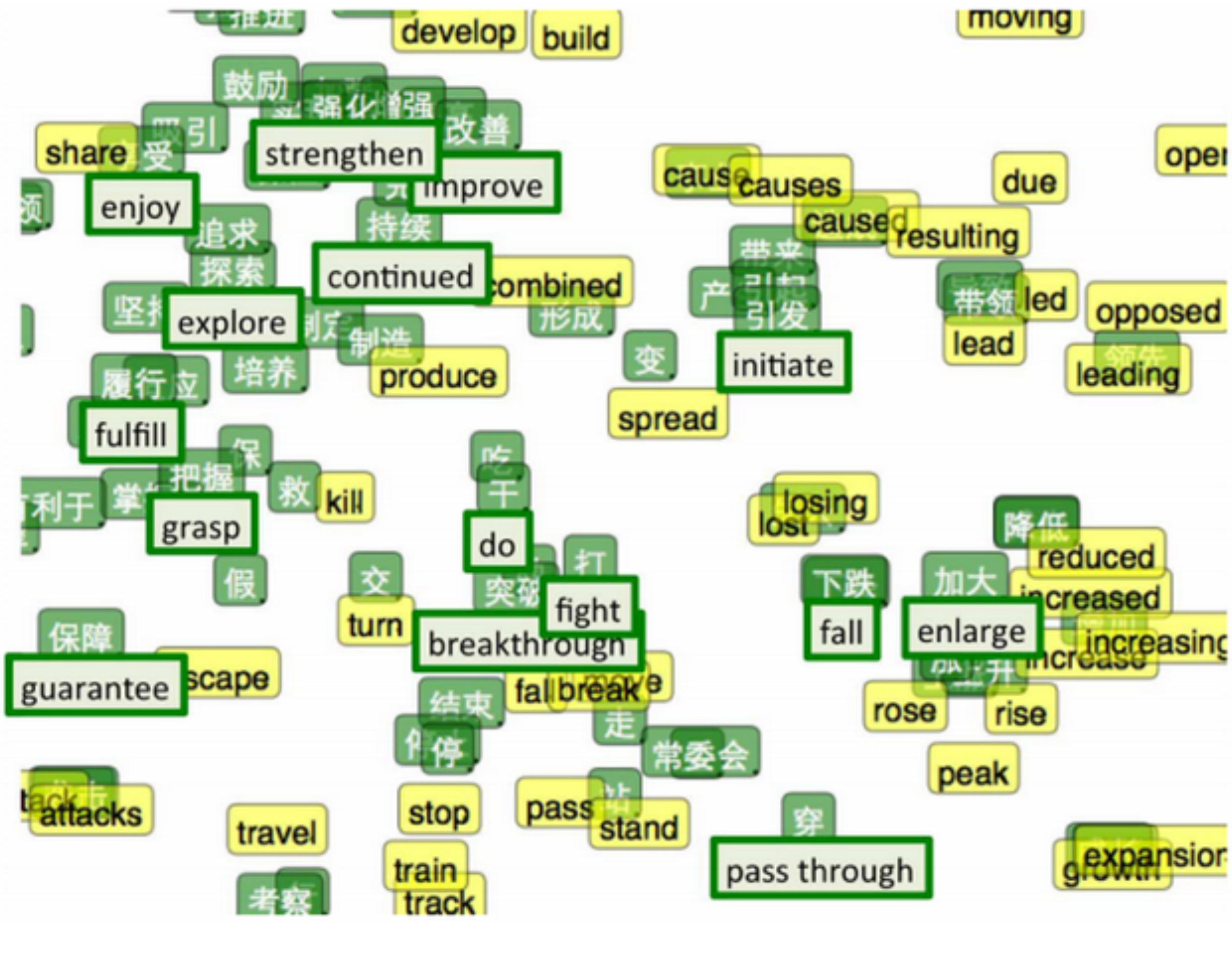
W(“China“) – W(“Beijing”) ≃ W(“Japan“) – W(“Tokyo")

词汇表示发现词汇语义层级



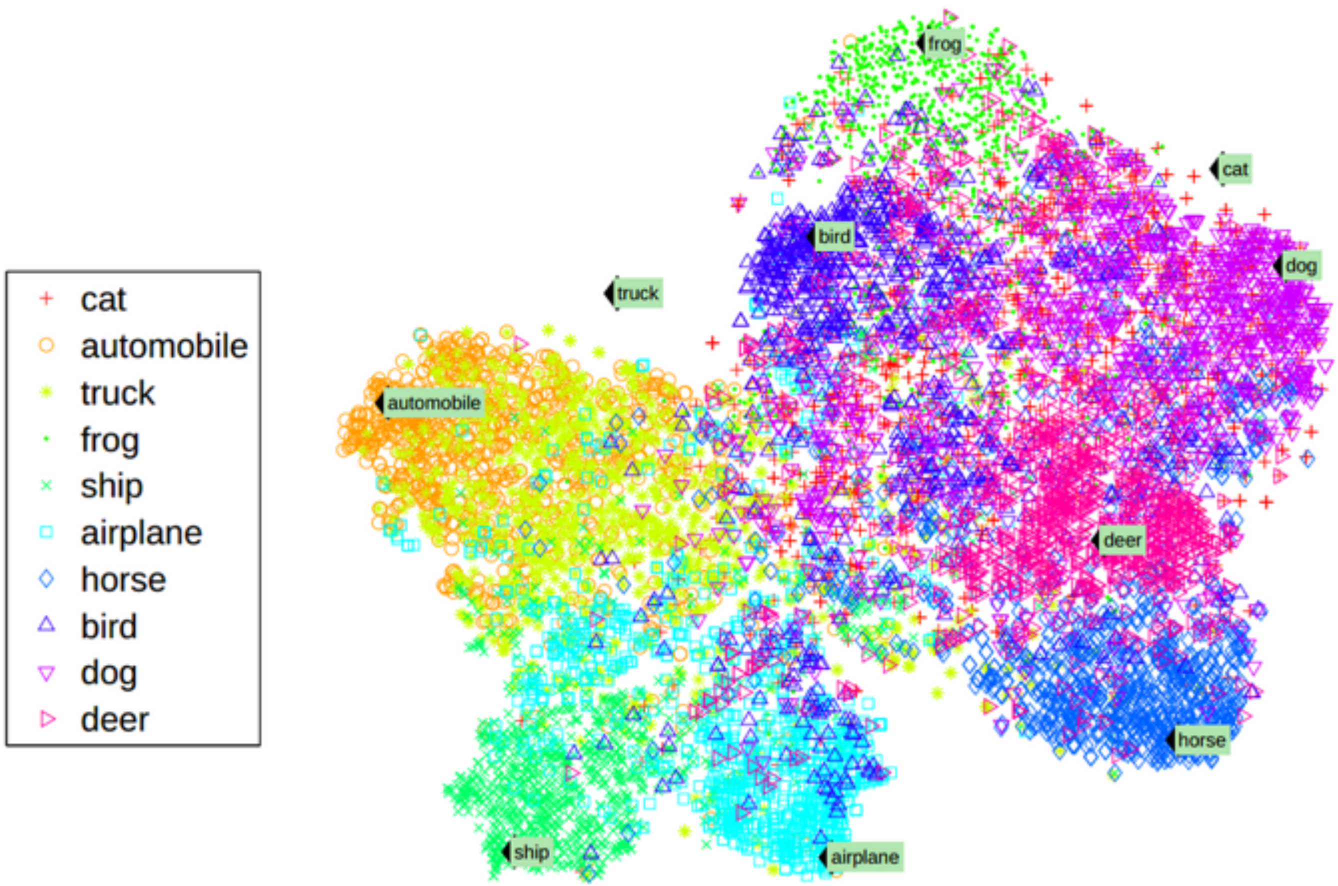
Fu, Ruiji, et al. Learning semantic hierarchies via word embeddings. ACL 2014.

分布式表示建立跨语言词汇联合表示



Zou, Will Y., et al. Bilingual word embeddings for phrase-based machine translation. EMNLP 2013.

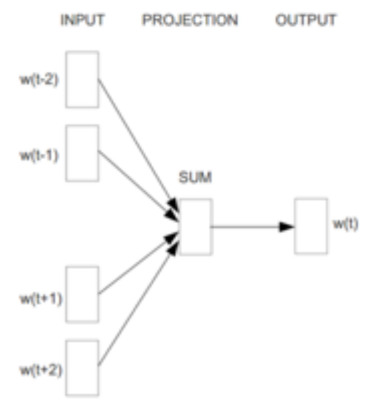
分布式表示建立视觉-文本联合表示

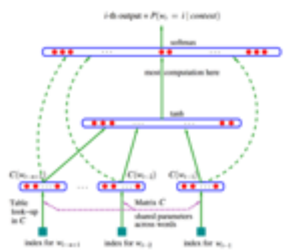


Richard Socher, et al. Zero-Shot Learning Through Cross-Modal Transfer. ICLR 2013.

Re-search, Re-invent

word2vec ≃ MF

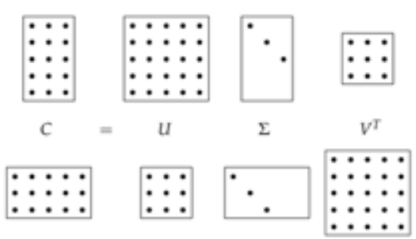




Neural Language Models

Distributional Representation

he curtains open and the stars shining in on the barely
ars and the cold , close stars " . And neither of the w
rough the night with the stars shining so brightly , it
made in the light of the stars . It all boils down , wr
surely under the bright stars , thrilled by ice-white
sun , the seasons of the stars ? Home , alone , Jay pla
m is dazzling snow , the stars have risen full and cold



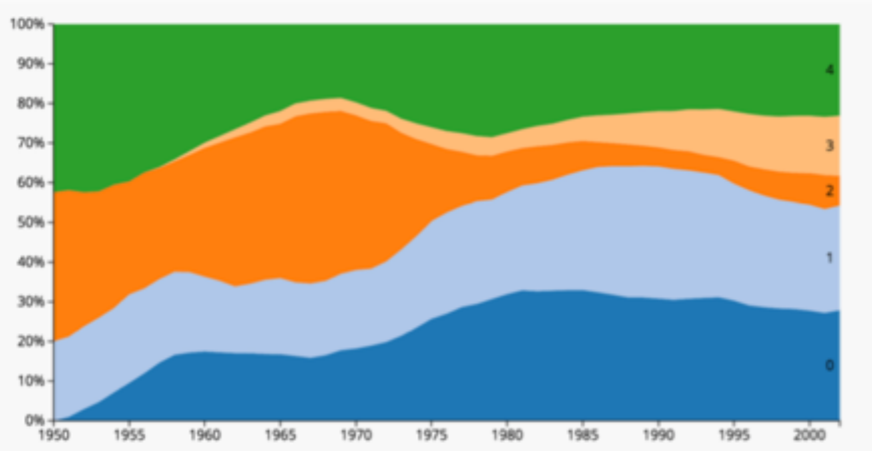
SVD

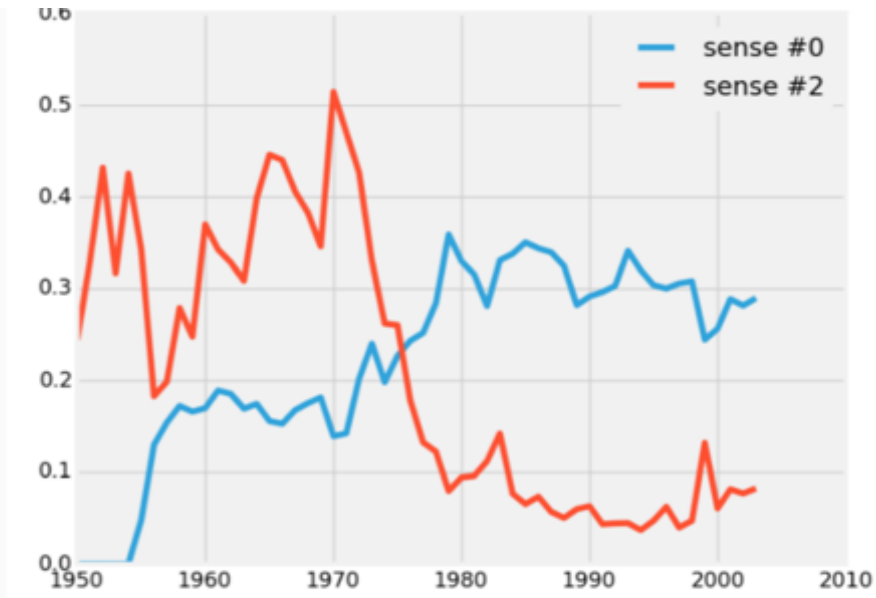
Levy and Goldberg. Neural word embedding as implicit matrix factorization. NIPS 2014.

词汇语义变迁研究

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. ACL 2016. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change.

词汇语义变迁研究





美国 0　英国，布什，法国，白宫，美国政府，国务卿，克林顿

美国 2　战争，美英，美军，发动，法国，美国政府，朝鲜

刘知远等. 词汇语义变化与社会变迁定量观测与分析.语言战略研究, 2016年.

语言分布式表示在大脑中的体现

- 来自认知科学的研究成果，利用分布式表示绘制词汇大脑地图

- 作者：Jack L. Gallant等
- 单位：美国加州伯克利大学
- 将985个常见英语词汇的对应大脑区域画了出来。7名志愿者躺在功能性核磁共振(fMRI)中两个多小时，过程中给他们播放The Moth Radio Hour有一万多字的故事
- 注：Thomas L. Griffiths是LDA Gibbs Sampling算法发明者

Huth, Alexander G., Wendy A. de Heer, Thomas L. Griffiths, et al. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, no. 7600 (2016): 453.

语言分布式表示在大脑中的体现

- 来自认知科学的研究成果，利用分布式表示绘制词汇大脑地图

- 词汇们分布在大脑四周，没有绝对的语言区域
- 意义相关的词语所激活的大脑区域相似
- 与词义对应的大脑区域呈双脑对称，这与过去一直以为的「左脑负责语义」的认识相悖
- 这份大脑词汇地图在人与人之间一致性很高

Huth, Alexander G., Wendy A. de Heer, Thomas L. Griffiths, et al. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, no. 7600 (2016): 453.

基于词汇表示的人类偏见研究

- 2017年Science发表论文指出，文本语料库包含可重现且准确的偏见印记，并能够被机器习得

RESEARCH

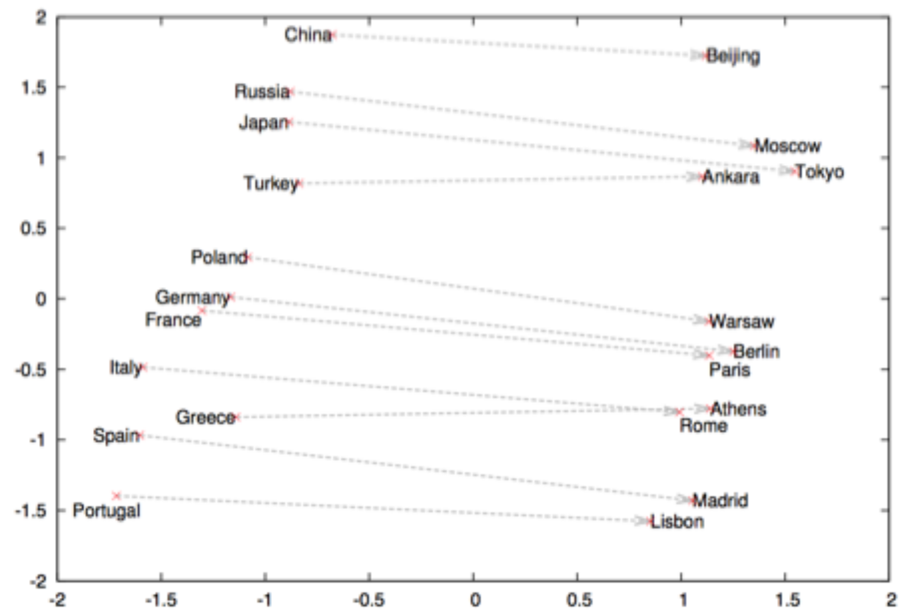
REPORT

COGNITIVE SCIENCE

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan,[1*] Joanna J. Bryson,[1,2*] Arvind Narayanan[1*]

Machine learning is a means to derive artificial intelligence by discovering patterns in existing data. Here, we show that applying machine learning to ordinary human language results in human-like semantic biases. We replicated a spectrum of known biases, as measured by the Implicit Association Test, using a widely used, purely statistical machine-learning model trained on a standard corpus of text from the World Wide Web. Our results indicate that text corpora contain recoverable and accurate imprints of our historic biases, whether morally neutral as toward insects or flowers, problematic as toward race or gender, or even simply veridical, reflecting the status quo distribution of gender with respect to careers or first names. Our methods hold promise for identifying and addressing sources of bias in culture, including technology.



基于词汇表示的人类偏见研究

- 论文指出，文本语料库包含可重现且准确的偏见印记，并能够被机器习得

- 心理实验：内隐测试，反应延迟时间之差

Target words	Attribute words	Original finding				Our finding			
		Ref.	N	d	P	N_T	N_A	d	P
Flowers vs. insects	Pleasant vs. unpleasant	(5)	32	1.35	10^{-8}	25 × 2	25 × 2	1.50	10^{-7}
Instruments vs. weapons	Pleasant vs. unpleasant	(5)	32	1.66	10^{-10}	25 × 2	25 × 2	1.53	10^{-7}
European-American vs. African-American names	Pleasant vs. unpleasant	(5)	26	1.17	10^{-5}	32 × 2	25 × 2	1.41	10^{-8}
European-American vs. African-American names	Pleasant vs. unpleasant from (5)	(7)	Not applicable			16 × 2	25 × 2	1.50	10^{-4}
European-American vs. African-American names	Pleasant vs. unpleasant from (9)	(7)	Not applicable			16 × 2	8 × 2	1.28	10^{-3}
Male vs. female names	Career vs. family	(9)	39k	0.72	$<10^{-2}$	8 × 2	8 × 2	1.81	10^{-3}
Math vs. arts	Male vs. female terms	(9)	28k	0.82	$<10^{-2}$	8 × 2	8 × 2	1.06	.018
Science vs. arts	Male vs. female terms	(10)	91	1.47	10^{-24}	8 × 2	8 × 2	1.24	10^{-2}
Mental vs. physical disease	Temporary vs. permanent	(23)	135	1.01	10^{-3}	6 × 2	7 × 2	1.38	10^{-2}
Young vs. old people's names	Pleasant vs. unpleasant	(9)	43k	1.42	$<10^{-2}$	8 × 2	8 × 2	1.21	10^{-2}

基于词汇表示的人类偏见研究

- 论文指出，文本语料库包含可重现且准确的偏见印记，并能够被机器习得

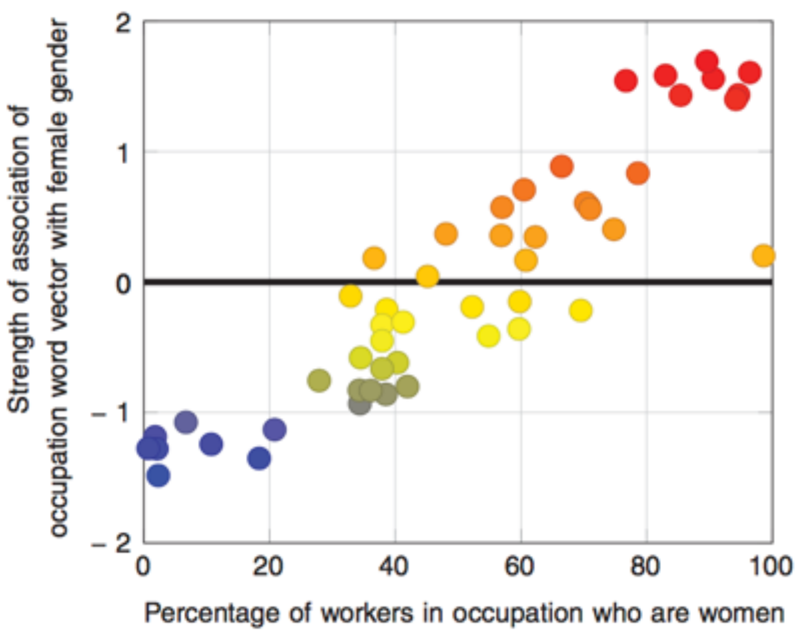


Fig. 1. Occupation-gender association. Pearson's correlation coefficient $\rho = 0.90$ with $P < 10^{-18}$.

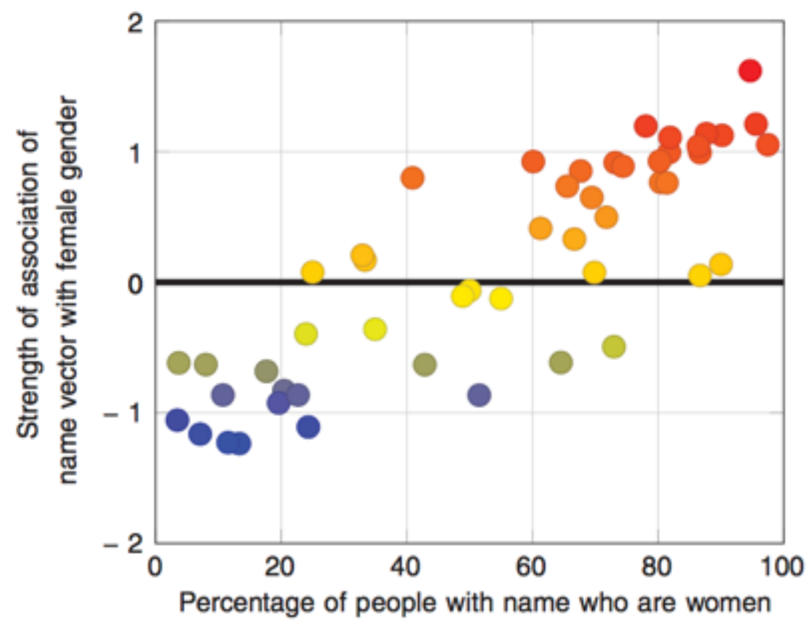


Fig. 2. Name-gender association. Pearson's correlation coefficient $\rho = 0.84$ with $P < 10^{-13}$.

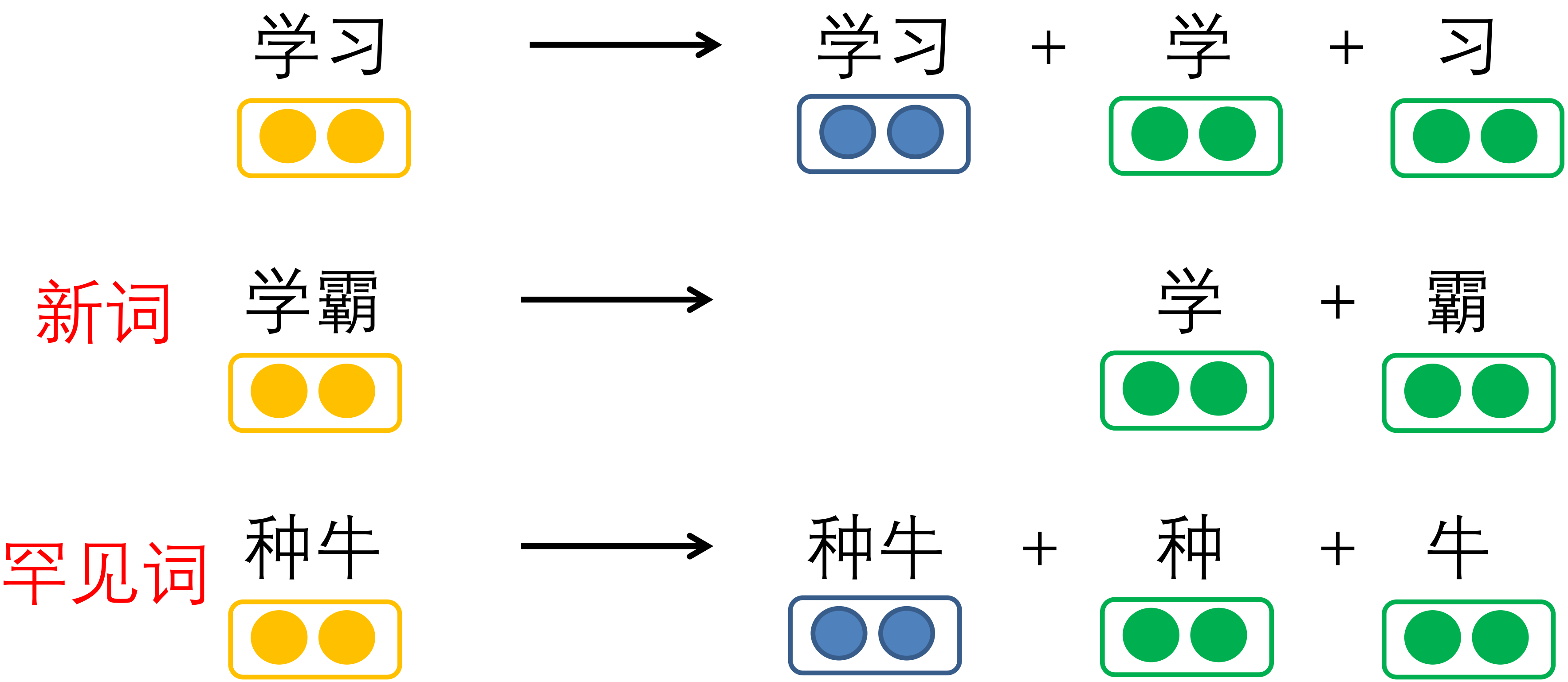
字符与词汇表示的联合学习模型

Xinxiong Chen, Lei Xu, Zhiyuan Liu, Maosong Sun, Huanbo Luan. Joint Learning of Character and Word Embeddings. *International Joint Conference on Aritificail Intelligence(IJCAI'15)*.

问题介绍

一般词向量对于新词和罕见词无法得到好的表示

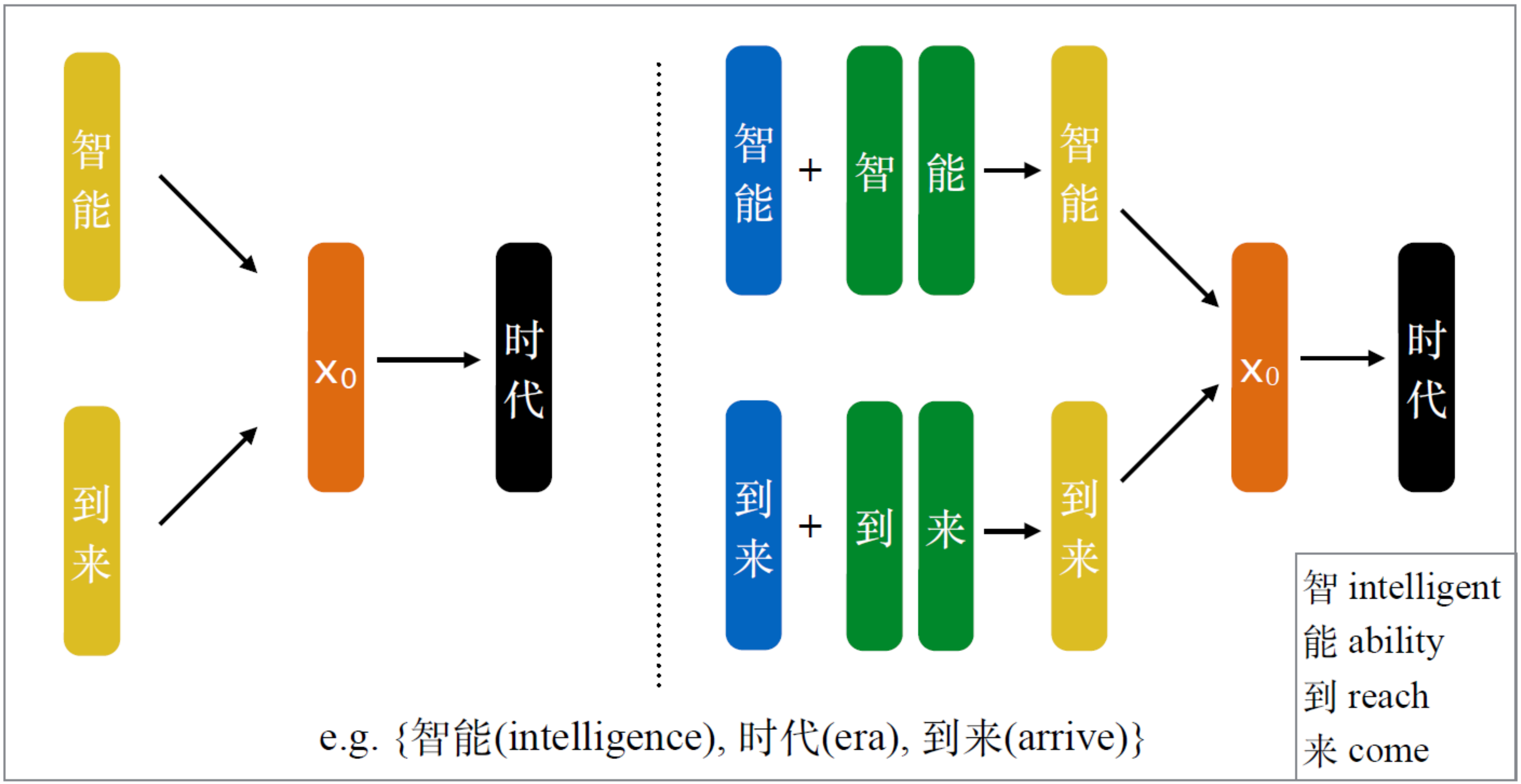
解决思路



使用词汇内部的字符信息（向量）来加强表示

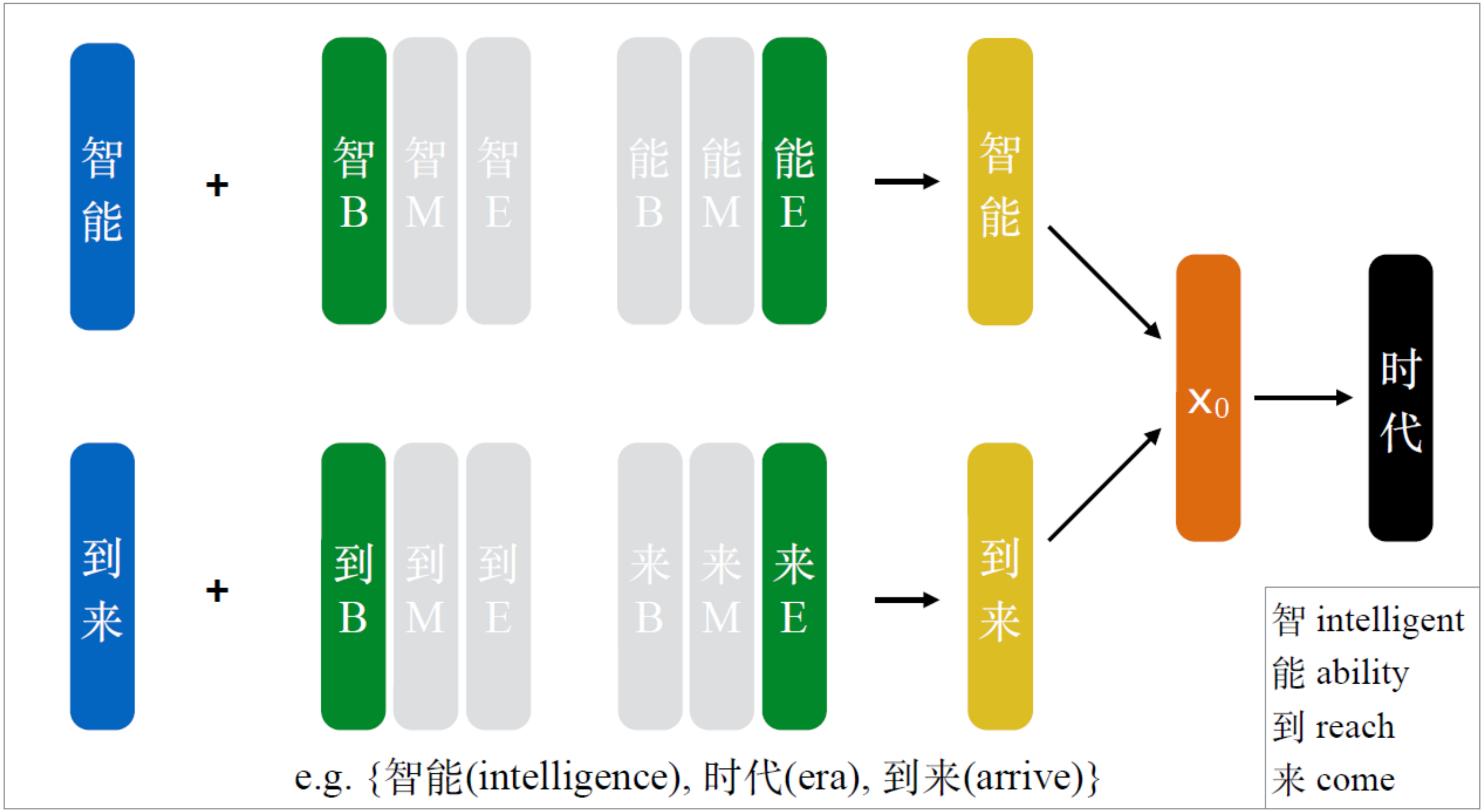
模型介绍

- 提出Character-enhanced word embedding(CWE)
- CWE可与已有各种框架(下为CBOW)进行融合



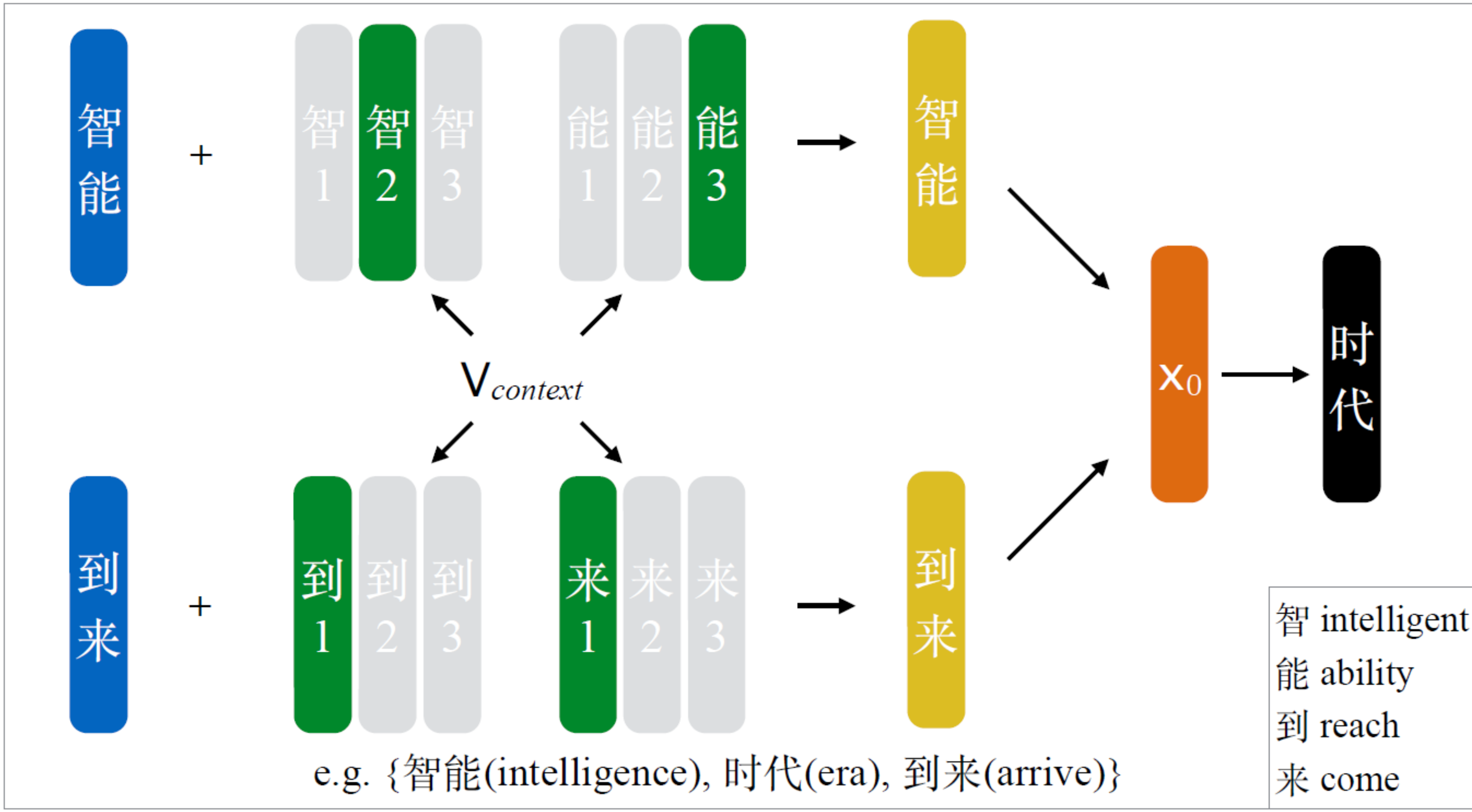
字的歧义

- 在中文中一个字可能包含有多种意思，针对字的歧义问题提出三种模型
- 第一种模型考虑位置信息



字的歧义

- 第二种将字的意思分为K类



- 第三种是动态决定每个字有多少意思

示例

- 字向量的k近邻
  - 基于位置：B和E分别代表词首(Begin)和词尾(End)
  - 基于聚类：I和II分别代表“道”的两种不同意思

道-B	道行 (attainments of a Taoist priest), 道经 (Taoist scriptures), 道法 (an oracular rule), 道人 (Taoist)
道-E	直道 (straight way), 近道 (shortcut), 便道 (sidewalk), 半道 (halfway), 大道 (revenue), 车道 (traffic lane)
道-I	直道 (straight way), 就道 (get on the way), 便道 (sidewalk), 巡道 (inspect the road), 大道 (revenue)
道-II	道行 (attainments of a Taoist priest), 邪道 (evil ways), 道法 (an oracular rule), 论道 (talk about methods)

实验：词相似度

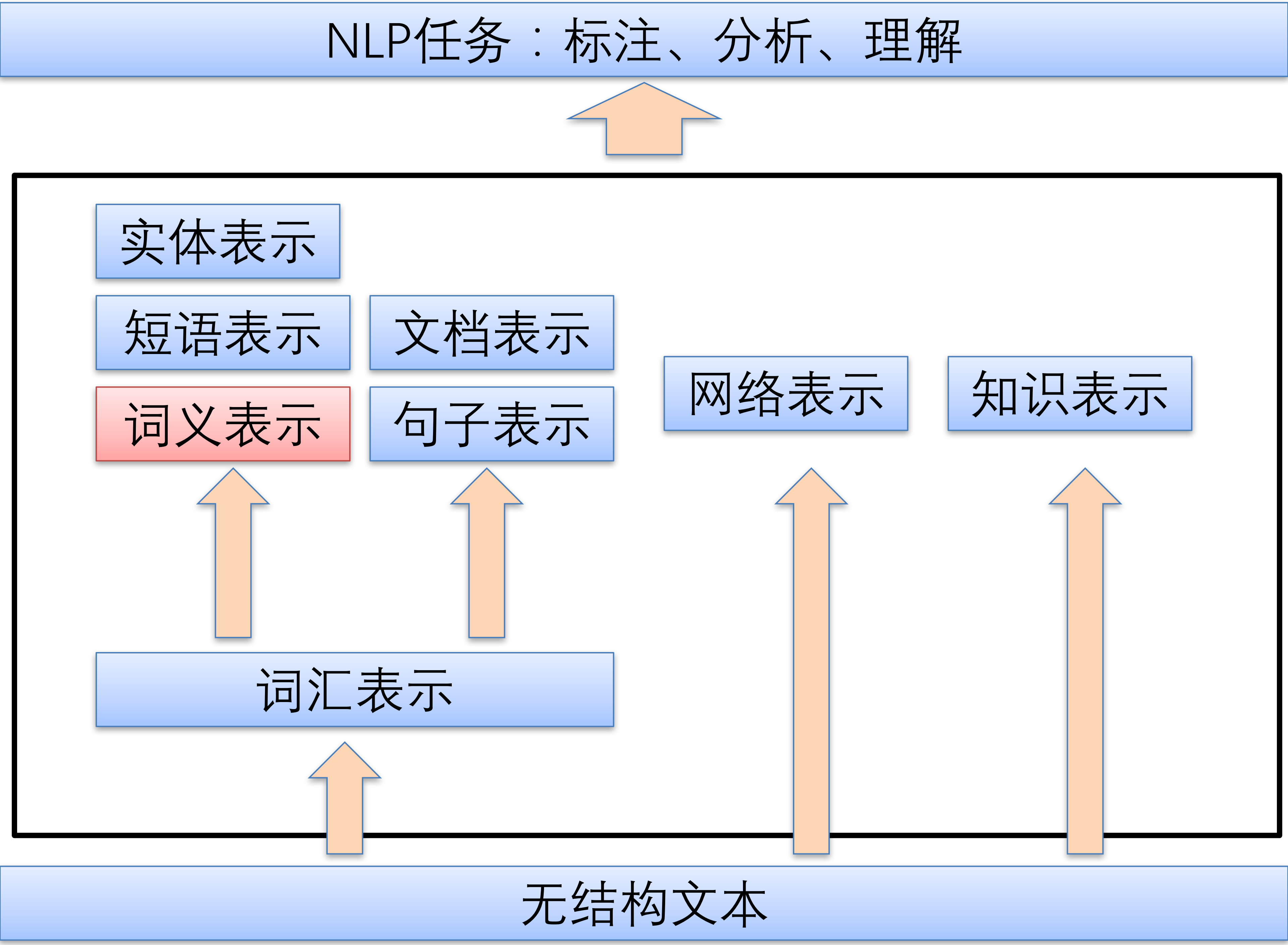
- Wordsim240和Wordsim296
  - 分别包含240和296个词对及对应人工相似度打分
  - 用Spearman系数计算与人工打分的相关程度

Dataset	wordsim-240		wordsim-296	
Method	233 Pairs	240 Pairs	280 Pairs	296 Pairs
CBOW	55.69	55.85	61.81	55.75
Skip-Gram	56.27	56.12	58.79	51.71
GloVe	47.72	48.22	48.22	43.06
CWE	56.90	57.56	64.02	63.57
CWE+P	56.34	57.30	62.39	62.41
CWE+L	**59.00**	59.53	**64.53**	**63.58**
CWE+LP	57.98	58.84	63.63	63.01
CWE+N	58.81	**59.64**	62.89	61.08

实验：类比

- 人工建立1125个四元类比对
  - 男人：女人::父亲：母亲
  - 给定三个词，猜第四个词，计算回答的准确率

Method	Total	Capital	State	Family
CBOW	54.85	51.40	66.29	62.92
+CWE	58.24	53.32	66.29	70.00
+CWE+P	60.07	54.36	66.29	73.75
Skip-Gram	69.14	62.78	82.29	80.83
+CWE	68.04	63.66	81.14	78.75
+CWE+P	72.07	65.44	**84.00**	**84.58**
GloVe	67.44	69.22	58.05	69.25
+CWE	70.42	70.01	64.00	76.25
+CWE+P	**72.99**	**73.26**	65.71	81.25

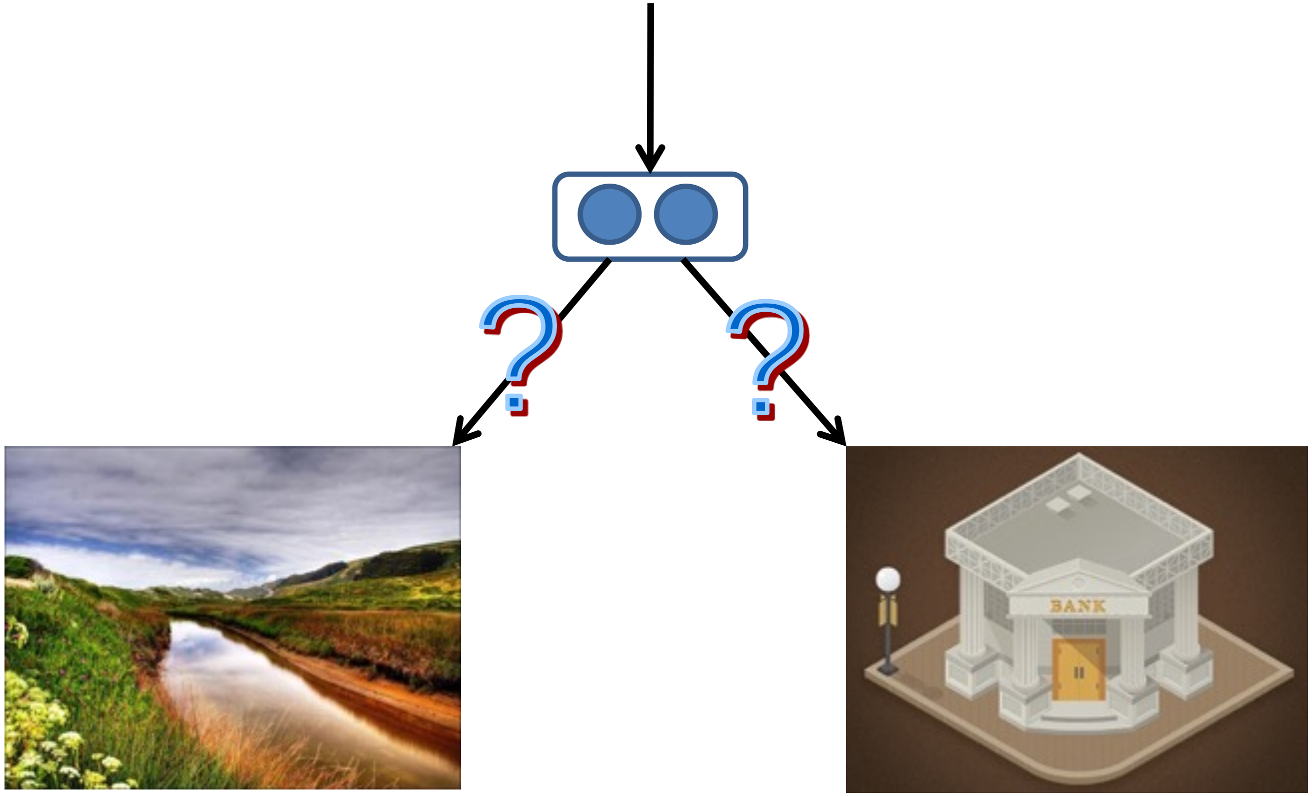

NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
词义表示
句子表示
网络表示
知识表示
词汇表示
无结构文本

面向一词多义的词汇表示模型

Xinxiong Chen, Zhiyuan Liu, Maosong Sun. A Unified Model for Word Sense Representation and Disambiguation. *The Conference on Empirical Methods in Natural Language Processing (EMNLP'14)*.

面向一词多义的词汇表示模型

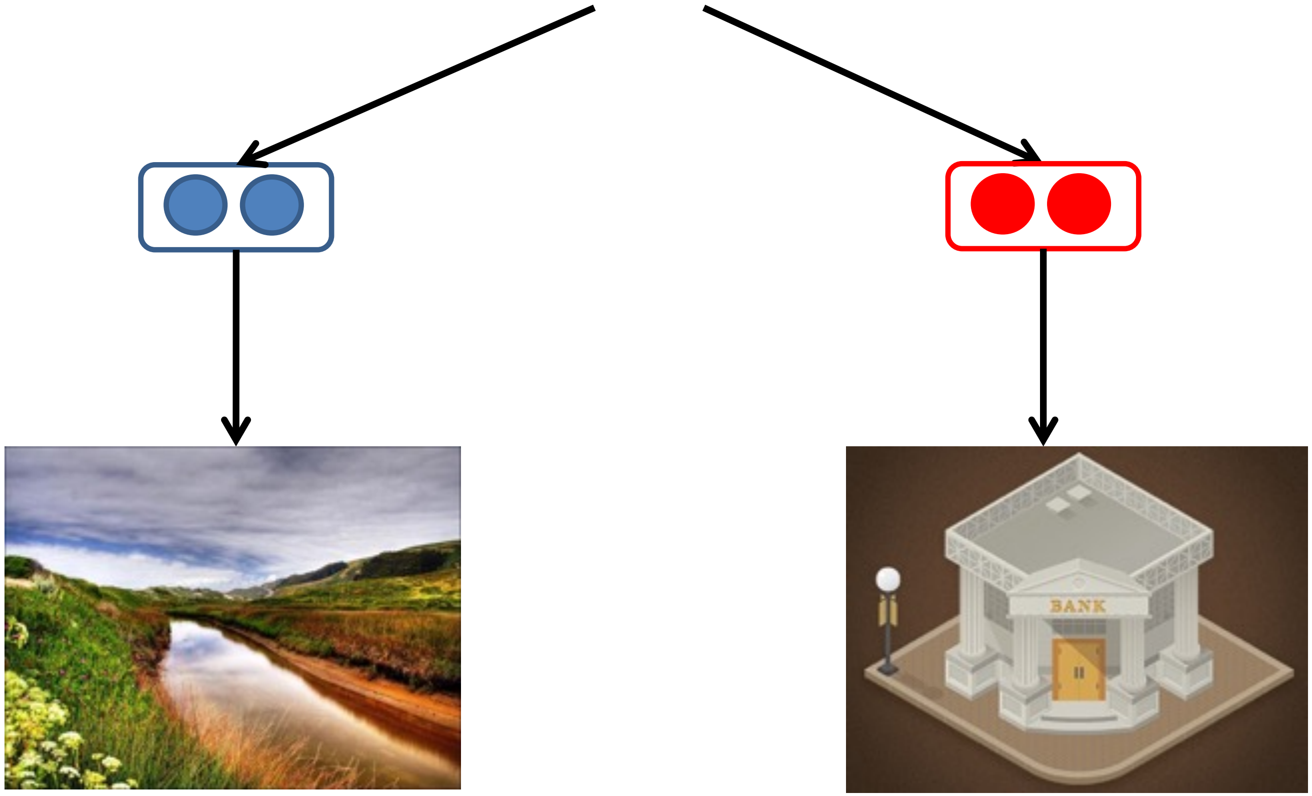
He sat on the bank of the lake



传统单一词向量无法解决一词多义问题

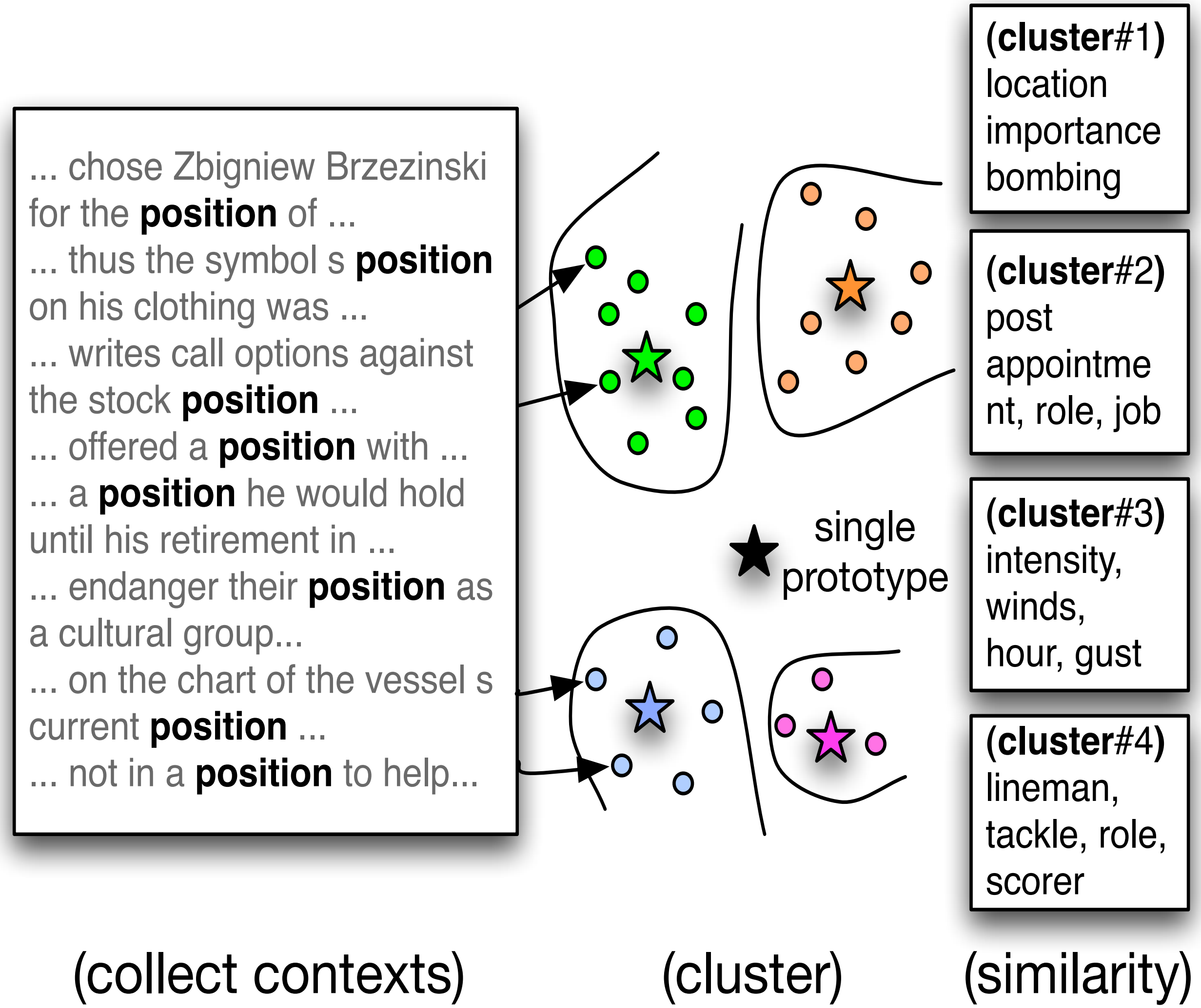
面向一词多义的词汇表示模型

He sat on the bank of the lake



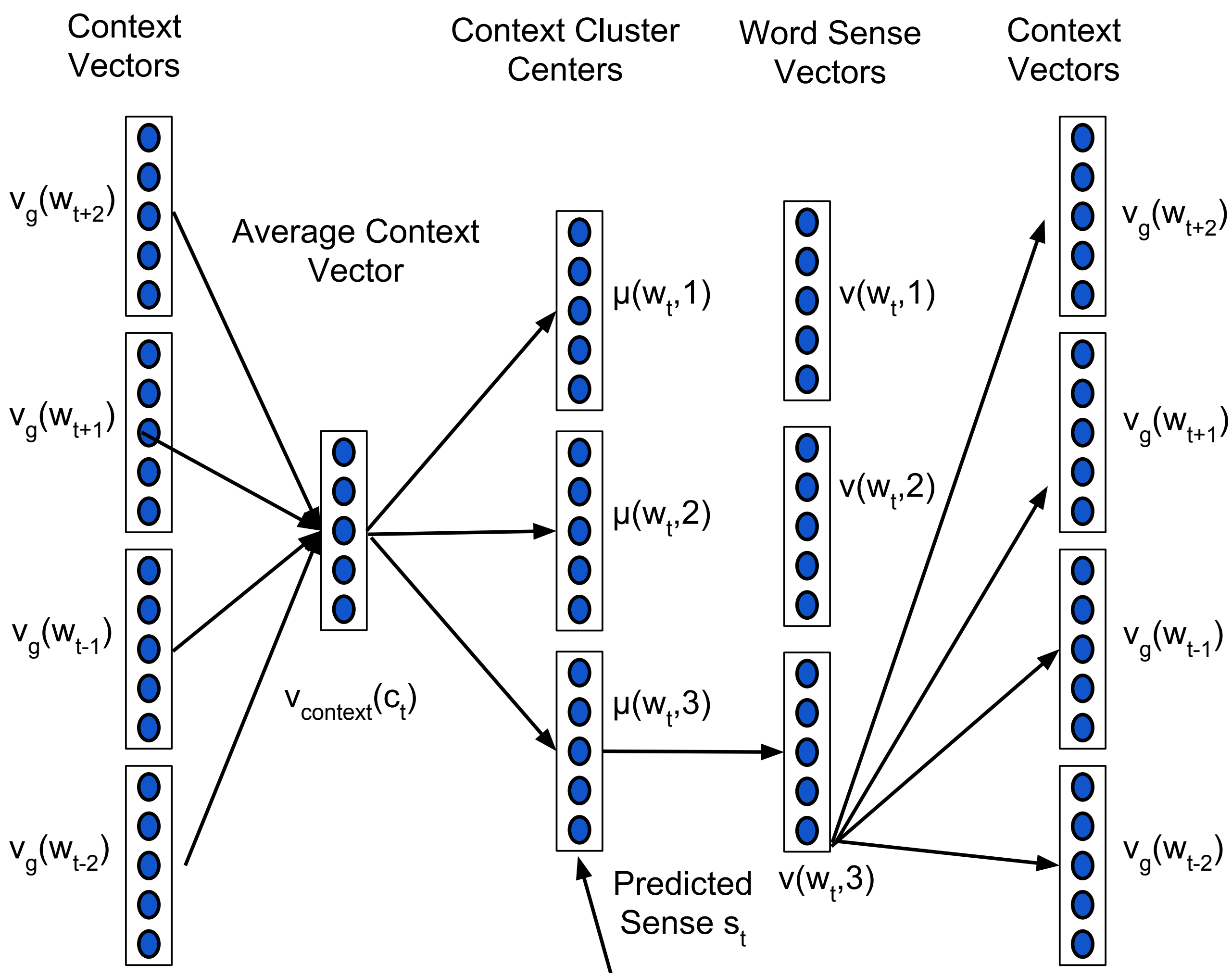
解决方案：为不同词义学习不同向量

Multiple Prototype Methods



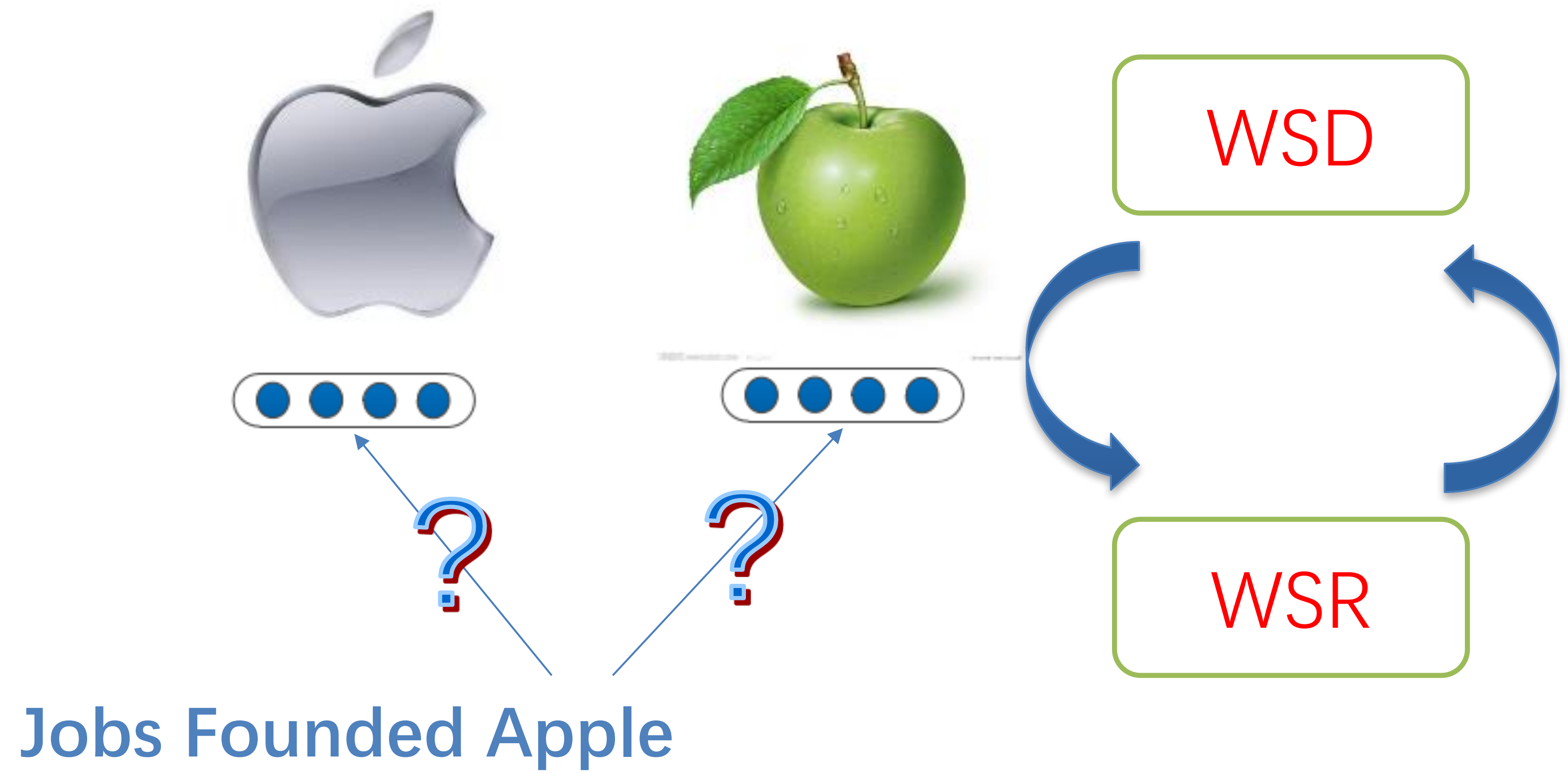
J. Reisinger and R. Mooney. Multi-prototype vector-space models of word meaning. HLT-NAACL 2010.
E Huang, et al. Improving word representations via global context and multiple word prototypes. ACL 2012.

Nonparametric Methods



Neelakantan et al. Efficient Non-parametric Estimation of Multiple Embeddings per Word in Vector Space. EMNLP 2014.

词义消歧与词义表示联合建模



Chen Xinxiong, et al. A Unified Model for Word Sense Representation and Disambiguation. EMNLP 2014.

模型介绍

- 首先，用word2vec技术学习单一词汇表示
- 然后，将WordNet作为词义来源，用gloss words中词汇向量的平均值作为词义的初始向量

bank$_1$

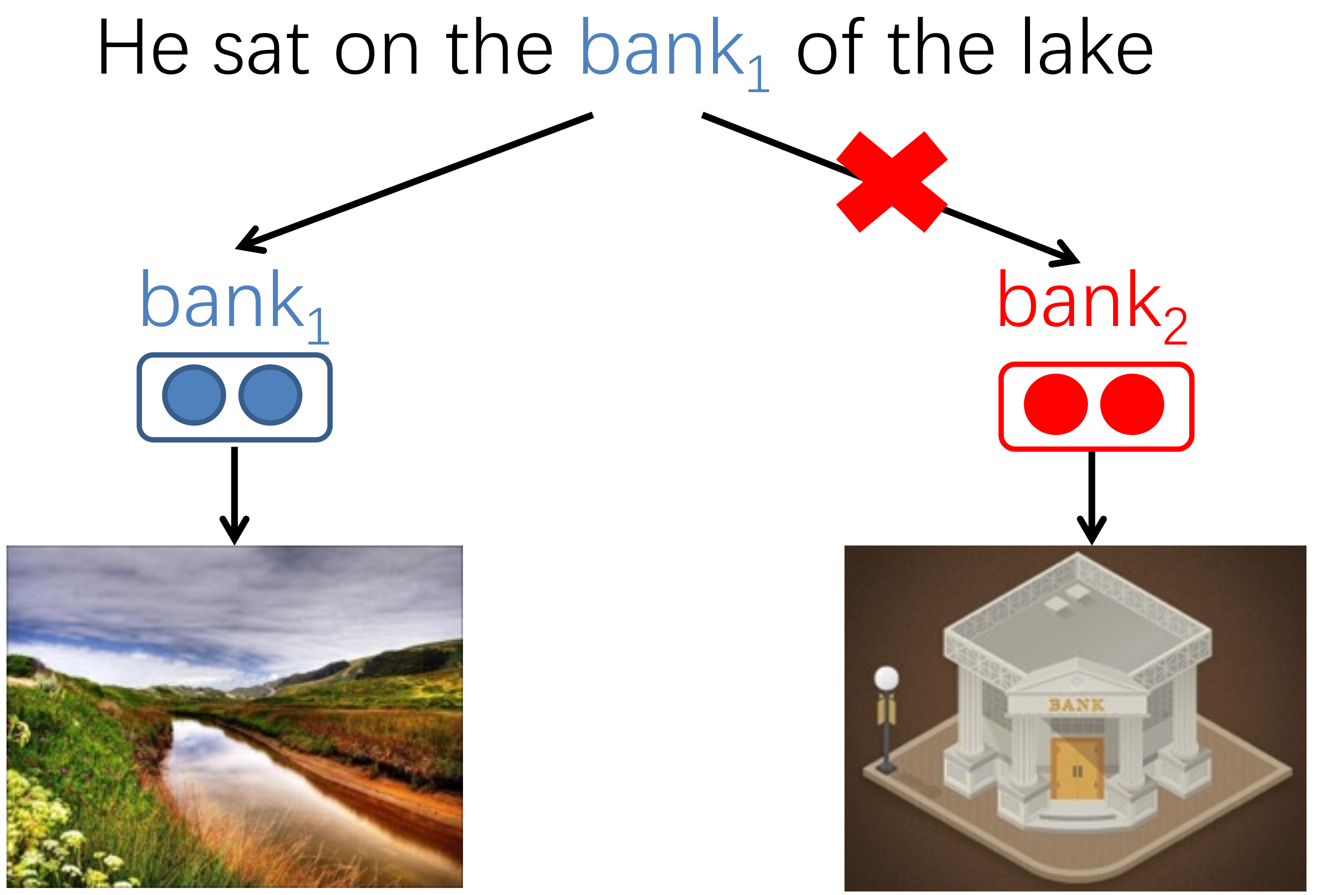
sloping land (especially the slope beside a body of water)

bank$_2$

a financial institution that accepts deposits and channels the money into lending activities

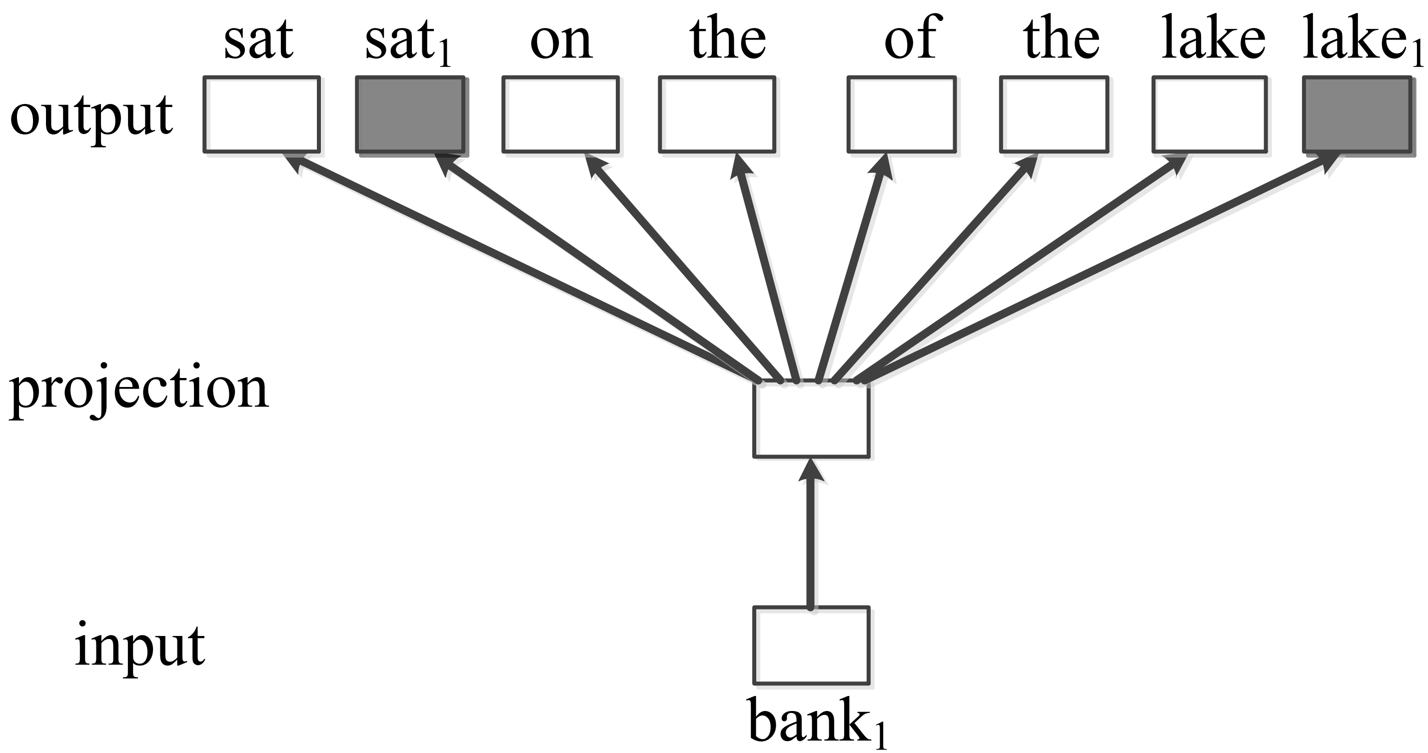
模型介绍

- 利用词义向量对大规模数据进行词义消歧



词义消歧与词义表示联合建模

- 利用消歧结果学习更新词义向量



示例

Word or Sense	Nearest neighbors
bank	banks, IDBI, CitiBank
$bank_1$	river, slope, Sooes
$bank_2$	mortgage, lending, loans
star	stars, stellar, trek
$star_1$	photosphere, radiation, gamma-rays
$star_2$	someone, skilled, genuinely
plant	plants, glavaticevo, herbaceous
$plant_1$	factories, machinery, manufacturing
$plant_2$	locomotion, organism, organisms

实验：特定领域词义消歧

- 体育与金融领域
  - 将句子中指定的词消歧到WordNet义项上
  - 共41个词，每个词对应约100个待消歧句子

Algorithm	Sports Recall	Finance Recall
Random BL	19.5	19.6
MFS BL	19.6	37.1
k-NN	30.3	43.4
Static PR	20.1	39.6
Personalized PR	35.6	46.9
Degree	42.0	47.8
Our Model	**57.3**	**60.6**

实验：全文词义消歧

- SemEval2007
  - 为5篇文章进行全文消歧，共包括2269个待消歧词，其中名词1108个

Algorithm	Type	Nouns F1	All F1
Random BL	U	63.5	62.7
MFS BL	Semi	77.4	78.9
SUSSX-FR	Semi	81.1	77.0
NUS-PT	S	82.3	**82.5**
SSI	Semi	**84.1**	**83.2**
Degree	Semi	**85.5**	**81.7**
Our Model	U	81.6	75.8
Our Model	Semi	**85.3**	**82.6**

实验：上下文敏感的词汇相似度计算

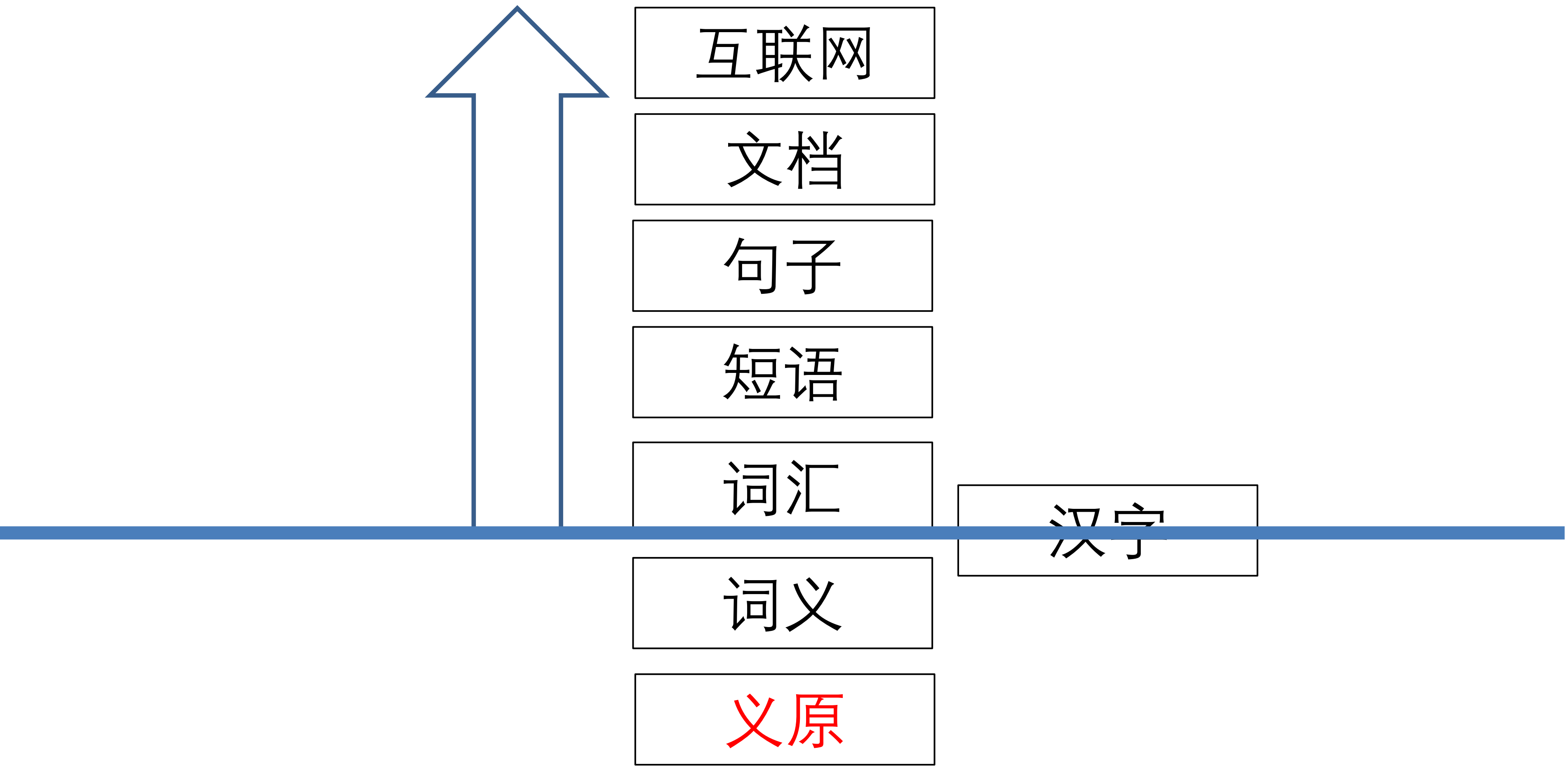
- SCWS数据集
  - 共有2003个词对及其所在句子，计算词对的相似度
  - 用Spearman系数计算与人工标注答案的相关程度

Model	$\rho \times 100$
C&W-S	57.0
Huang-S	58.6
Huang-M AvgSim	62.8
Huang-M AvgSimC	65.7
Our Model-S	64.2
Our Model-M AvgSim	66.2
Our Model-M AvgSimC	**68.9**

融合HOWNET的词义表示学习

自然语言特点

- 词汇或汉字是最小使用单位，但不是最小语义单位



义原知识与HowNet

- HowNet是**董振东、董强**父子毕三十年之功标注的大型语言知识库，主要面向中文的词汇与概念标注义原知识
- 秉承**还原论**思想， 用义原（Sememe）标注词汇语义，义原顾名思义就是**原子语义**，即最基本的、不宜再分割的最小语义单位
- HowNet逐渐构建出一套精细的义原体系（包含约2000个义原）， 累计标注了数十万词汇/词义的语义信息

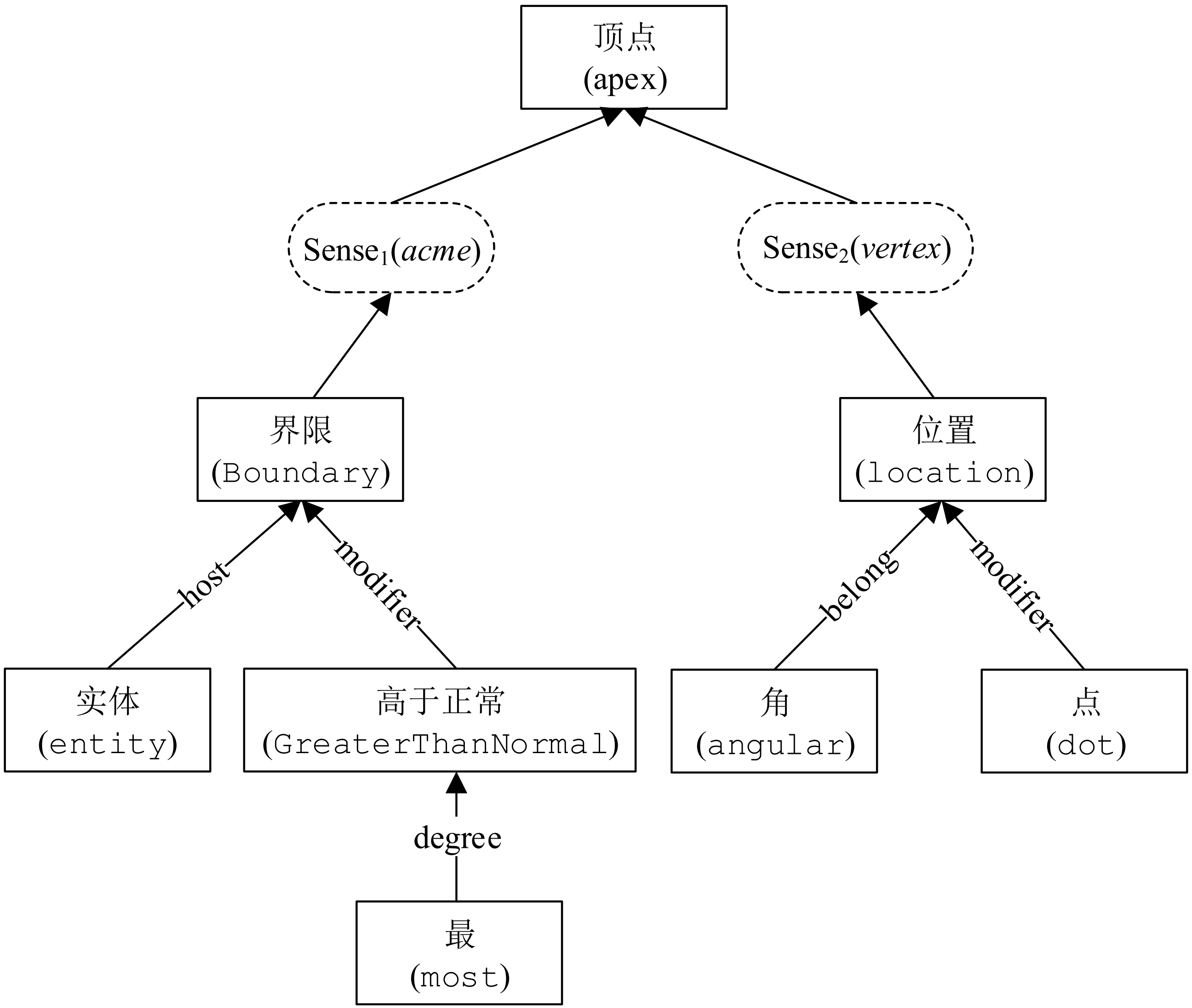
HowNet—瞥

- 每个词义信息用义原标注，每个义原用 英文|中文 标明
- 义原之间还标记语义关系，如modifier，host，belong等

> 顶点#1
> DEF={Boundary|界限:host={entity|实体},modifier={GreaterThanNormal|高于正常:degree={most|最}}}
> 顶点#2
> DEF={location|位置:belong={angular|角},modifier={dot|点}}

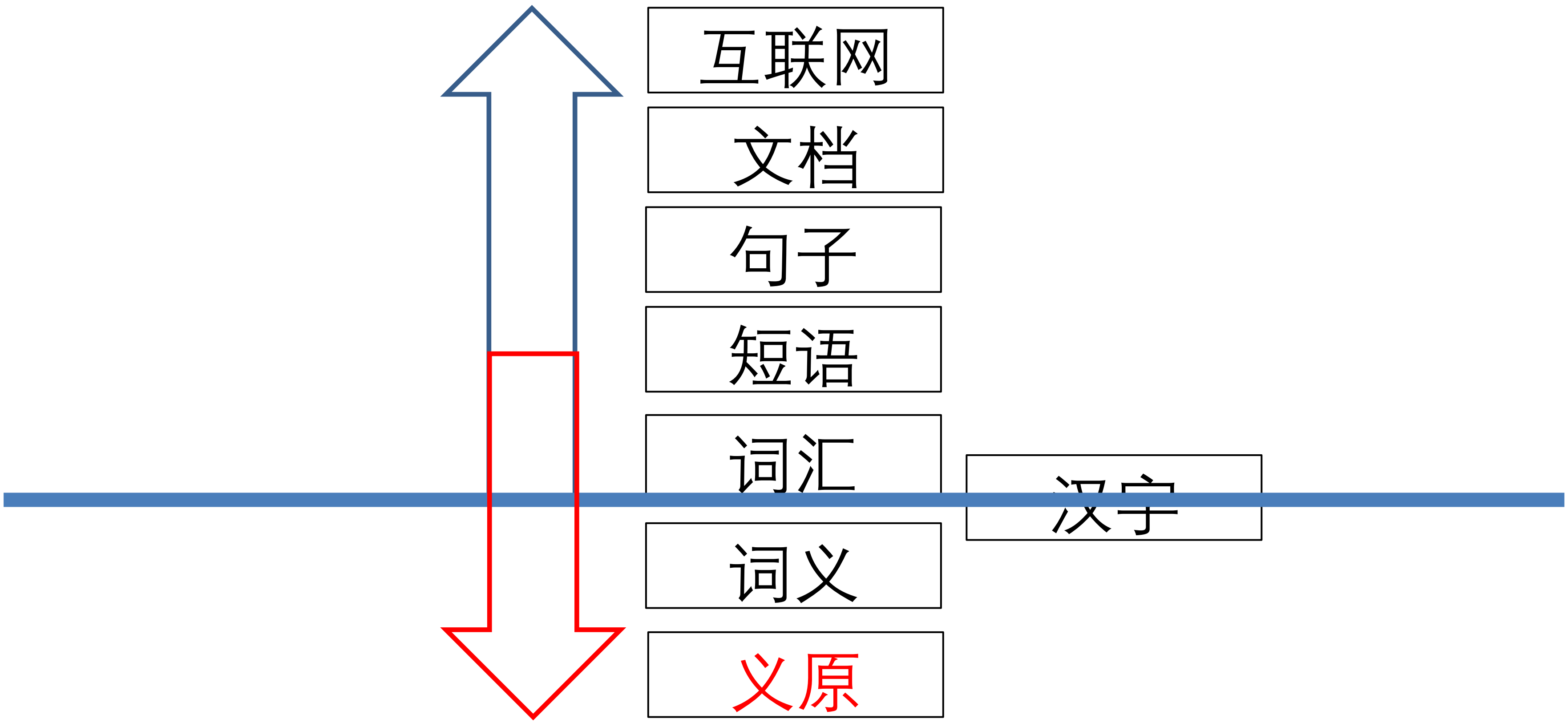
HowNet—瞥

- 义原知识带有层次结构



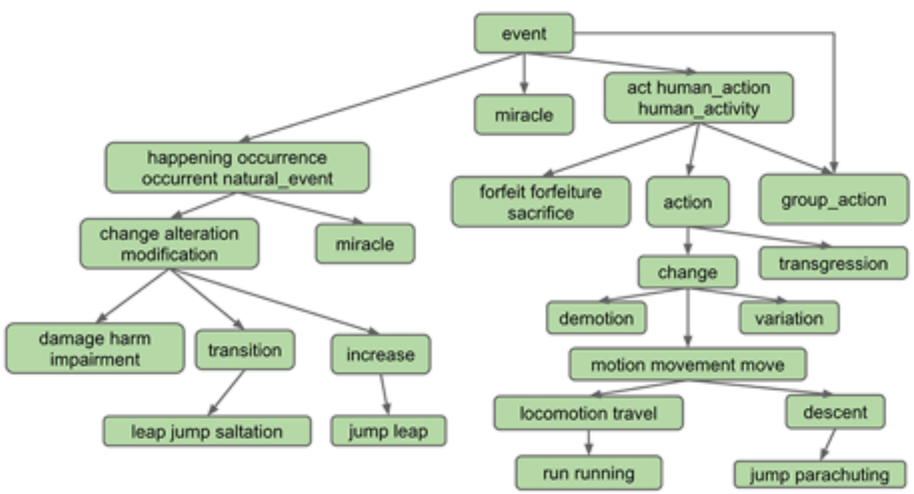
HowNet特点

- **在自然语言理解方面，更贴近语言本质特点**
  - 义原标注体系是突破词汇屏障，深入了解词汇背后丰富语义信息的重要通道

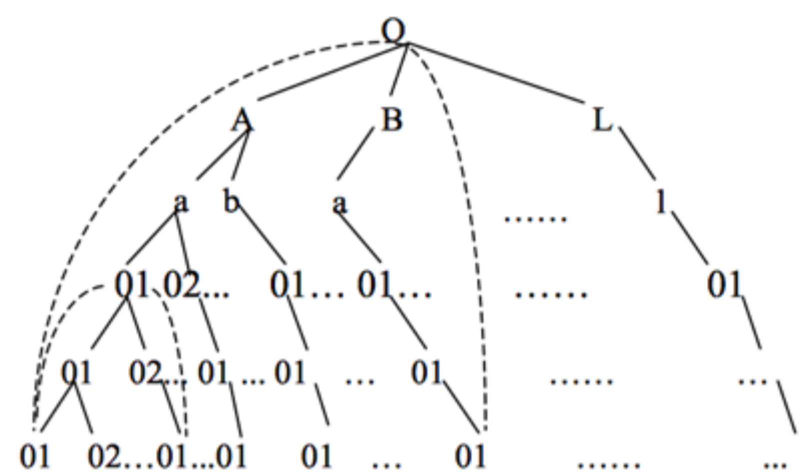


HowNet特点

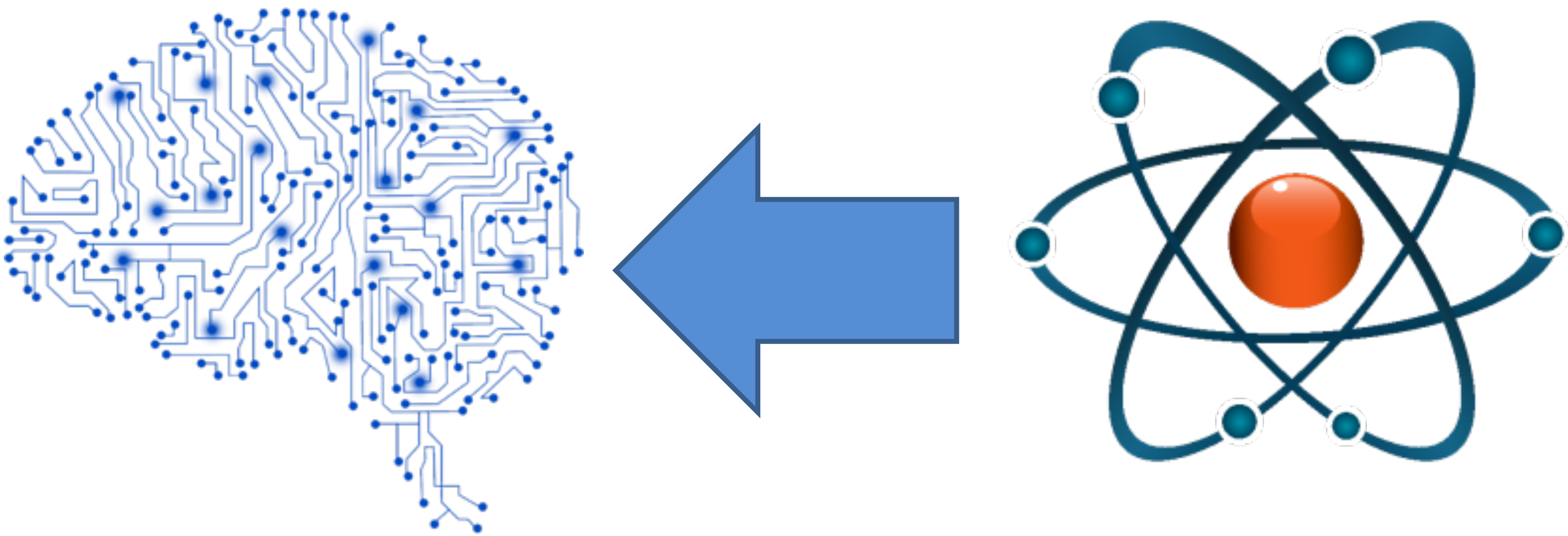
- **在融入深度学习方面，具有无可比拟优势**
  - 与WordNet、同义词词林等知识库组织模式不同
  - HowNet通过统一义原标注体系直接精准刻画语义信息。每个义原含义明确固定，可被直接作为语义标签融入机器学习模型



WordNet Synset体系



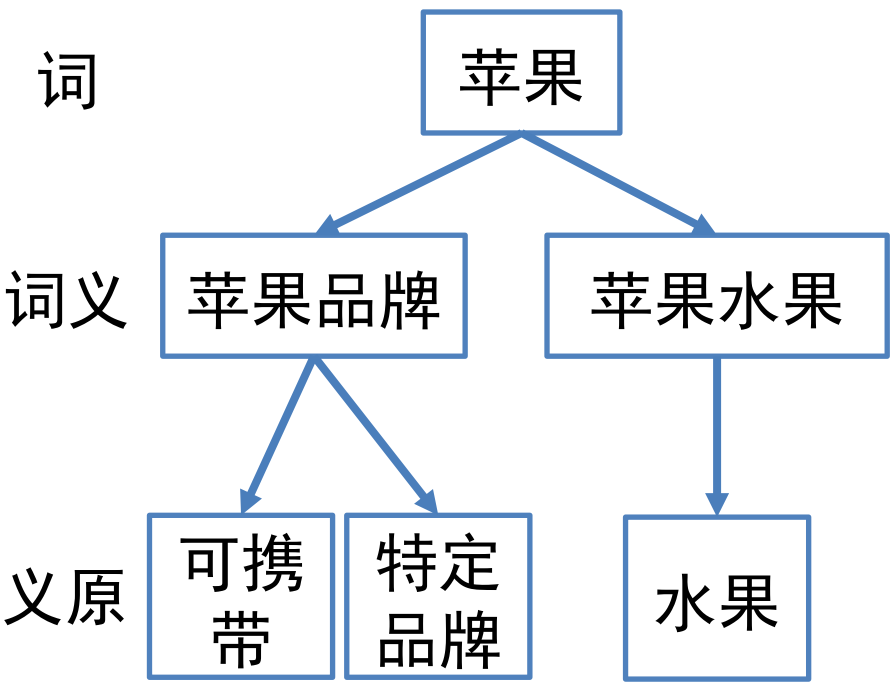
同义词词林层次类别体系



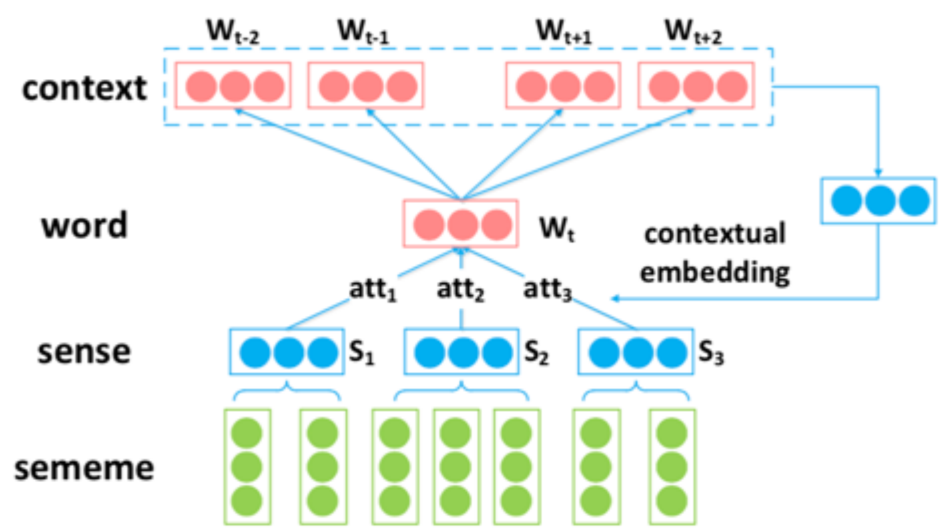
深度学习　　　　义原知识

融合义原知识的词义表示

- 考虑HowNet的词义-义原标注信息，提升词义表示性能



HowNet词义-义原标注示例



义原-词义-词汇的联合表示学习模型
（ACL 2017）

实验结果

- 在词相似度计算和类比推理任务上的性能得到显著提升

Model	Accuracy Capital	Accuracy City	Accuracy Relationship	Accuracy All	Mean Rank Capital	Mean Rank City	Mean Rank Relationship	Mean Rank All
CBOW	49.8	85.7	**86.0**	64.2	36.98	1.23	62.64	37.62
GloVe	57.3	74.3	81.6	65.8	19.09	1.71	3.58	12.63
Skip-gram	66.8	93.7	76.8	73.4	137.19	1.07	2.95	83.51
SSA	62.3	93.7	81.6	71.9	45.74	1.06	3.33	28.52
MST	65.7	95.4	82.7	74.5	50.29	1.05	2.48	31.05
SAC	79.2	97.7	75.0	81.0	28.88	1.02	2.23	18.09
SAT	**82.6**	**98.9**	80.1	**84.5**	**14.78**	**1.01**	**1.72**	**9.48**

类比推理任务评测结果，其中SAC、SAT代表两种本工作提出的模型

实验结果

- 能够有效根据上下文信息实现词义消歧

上下文词	义原“首都”	义原“古巴”
古巴	0.39	0.42
俄罗斯	0.39	-0.09
雪茄	0.00	0.36

上下文词对“哈瓦那”义原注意力值示例

例句	词义1：概率	词义2：概率
苹果素有果中王美称	苹果品牌：0.28	苹果水果：0.72
苹果电脑无法正常启动	苹果品牌：0.87	苹果水果：0.13
八支**队伍**进入第二阶段团体赛	团体：0.90	部队：0.10
公安基层**队伍**组织建设	团体：0.15	部队：0.85

根据上下文消歧结果示例

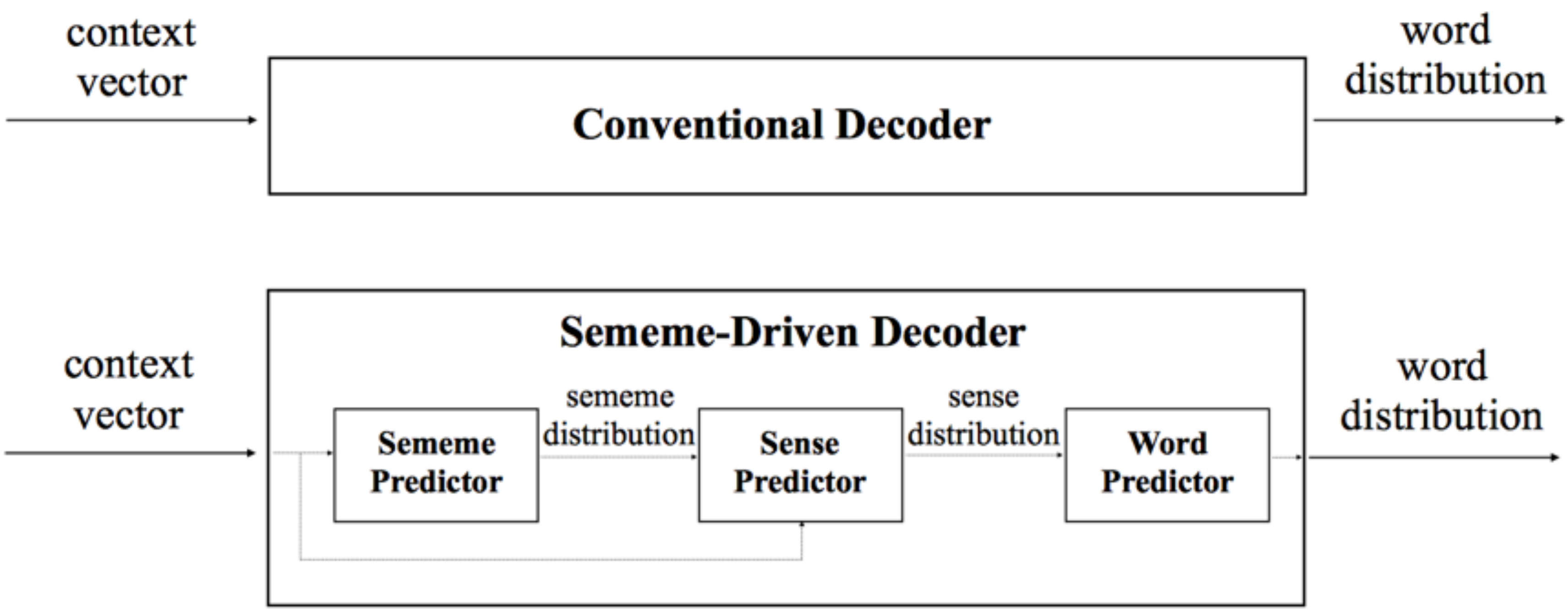
融合义原知识的神经语言模型

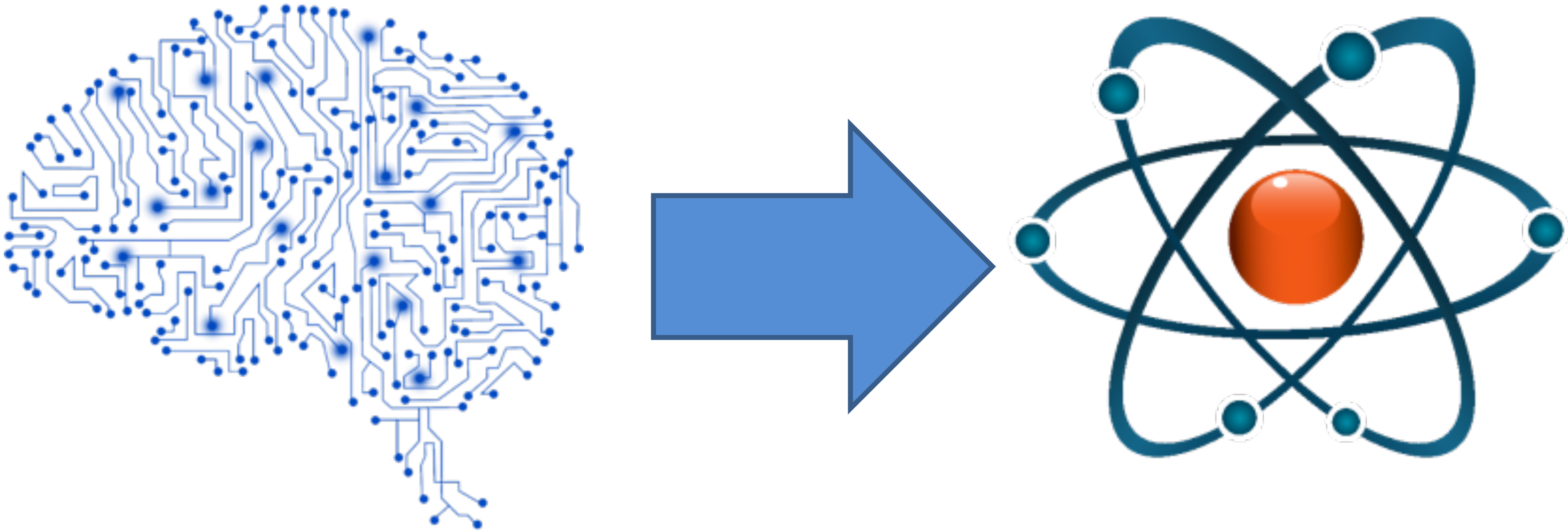
- 语言模型是自然语言处理的核心任务
- NGram是前深度学习时代的代表语言模型，深度学习框架CNN、RNN即用来学习语言模型
- 马尔科夫性：当前词出现的概率，依赖于上下文出现的词

The U.S. trade deficit last year is initially estimated to be 40 billion ______ .

融合义原知识的神经语言模型

- 传统深度学习语言模型是纯数据驱动模型
- 目标：建立义原知识驱动的语言模型





深度学习　　　　义原知识

基于语义表示学习的义原推荐

- HowNet等知识库主要依赖人工标注，费时费力
- **义原自动推荐**：实现义原知识库与时俱进，提升标注一致性

基于词向量的近邻协同过滤方法（SPWE）

$$P(s_j, w) = \sum_{w_i \in W} \cos(\mathbf{w}, \mathbf{w_i}) \cdot \mathbf{M}_{ij} \cdot c^{r_i}$$

基于词义-义原矩阵分解的推荐方法（SPSE）

$$\mathcal{L} = \sum_{w_i \in W, s_j \in S} \left(\mathbf{w}_i \cdot (\mathbf{s}_j + \bar{\mathbf{s}}_j) + \mathbf{b}_i + \mathbf{b}'_j - \mathbf{M}_{ij}\right)^2 + \lambda \sum_{s_j, s_k \in S} \left(\mathbf{s}_j \cdot \bar{\mathbf{s}}_k - \mathbf{C}_{jk}\right)^2,$$

（IJCAI 2017）

实验结果

- 将两种方法相融合，能够显著提升义原推荐效果。词性、词频有显著影响。

Method	MAP
SPSE	0.554
SPASE	0.506
GloVe+LR	0.662
SPWE	0.676
SPWE+SPASE	0.683
SPWE+SPSE	**0.713**

义原推荐效果

POS	number of words	MAP
adverb	136	0.568
adjective	808	0.544
verb	1,867	0.583
noun	3,556	0.747

不同词性的词汇义原推荐效果

word frequency	number of words	MAP
<800	1,659	0.817
800 - 3,000	1,494	0.736
3,001 - 15,000	1,672	0.690
>15,000	1,311	0.596

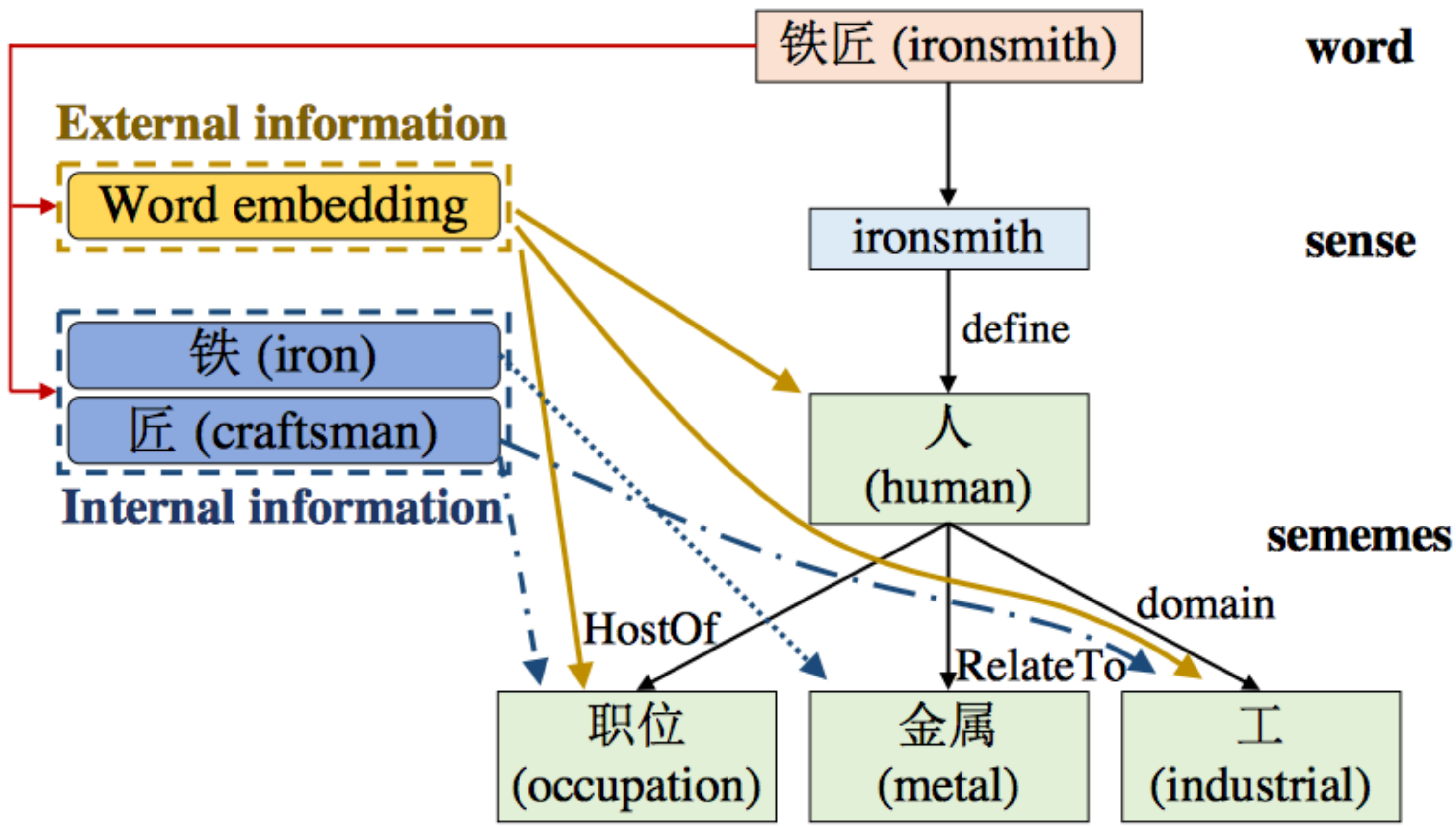
不同词频的词汇义原推荐效果

实验结果

words	Top 5 sememes prediction
网迷(webaholic)	**人(human)**, **因特网(internet)**, **经常(frequency)**, **利用(use)**, 喜欢(fond of)
专递(express mail)	**邮寄(post)**, **信件(letter)**, **快(fast)**, 事情(fact), 车(landvehicle)
电影业(film industry)	**事务(affairs)**, **艺(entertainment)**, **表演物(shows)**, 拍摄(take picture), **制造(produce)**
漂流(rafting)	船(ship), 旅游(tour), 游(swim), 水域(waters), 消闲(whileaway)
公羊(ram)	**牲畜(livestock)**, **男(male)**, 女(female), 走兽(beast), 饲养(foster)

考虑内部汉字信息的义原推荐

- 词汇内部的汉字信息对语义理解具有重要意义，提出同时考虑内部汉字信息进行义原推荐



(ACL 2018)

实验结果

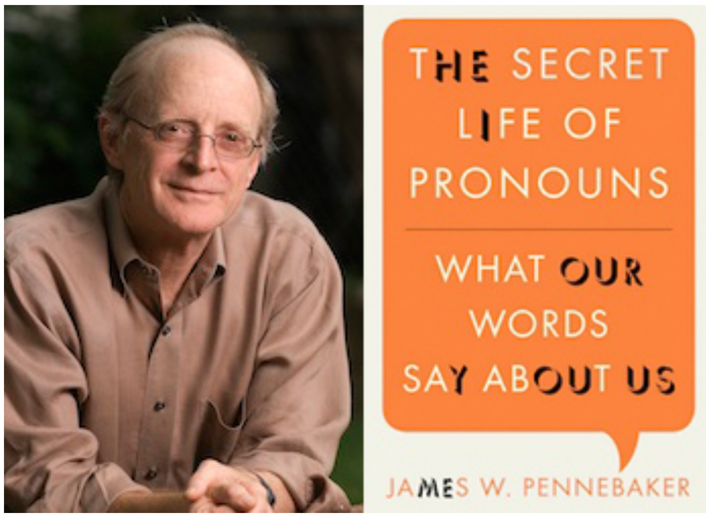
- 实验证明，考虑内部汉字信息，对于低频词的义原推荐提升尤为明显

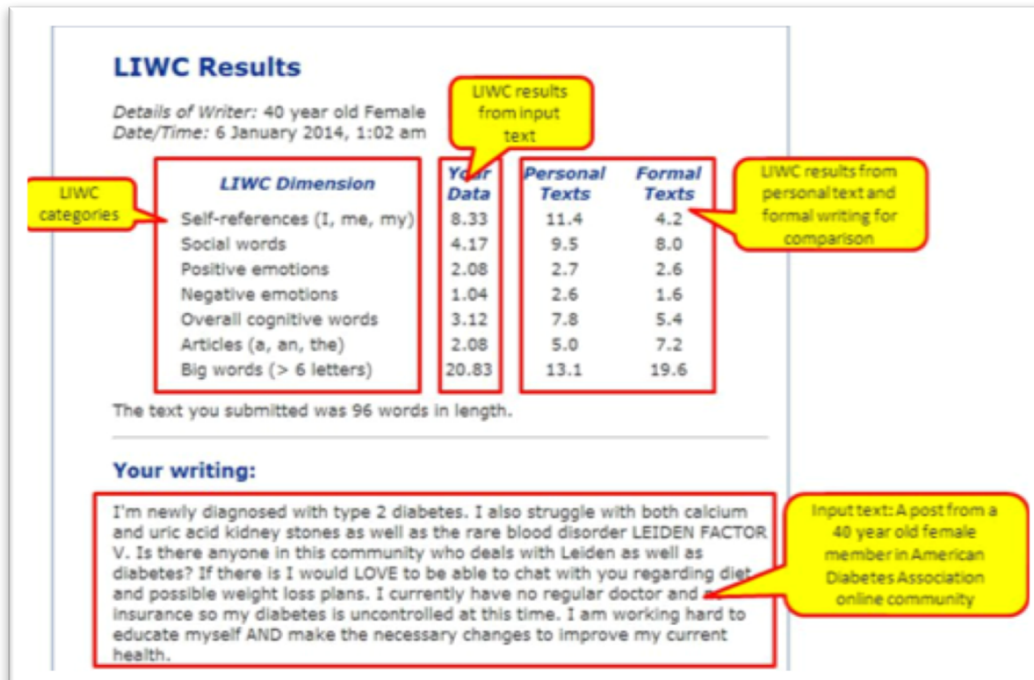
word frequency occurrences	⩽ 50 8537	51– 100 4868	101 – 1,000 3236	1,001 – 5,000 2036	5,001 – 10,000 663	10,001 – 30,000 753	>30,000 686
SPWE	0.312	0.437	0.481	0.558	0.549	0.556	0.509
SPSE	0.187	0.273	0.339	0.409	0.407	0.424	0.386
SPWE + SPSE	0.284	0.414	0.478	0.556	0.548	0.554	0.511
SPWCF	0.456	0.414	0.400	0.443	0.462	0.463	0.479
SPCSE	0.309	0.291	0.286	0.312	0.339	0.353	0.342
SPWCF + SPCSE	0.467	0.437	0.418	0.456	0.477	0.477	0.494
SPWE + fastText	0.495	0.472	0.462	0.520	0.508	0.499	0.490
CSP	**0.527**	**0.555**	**0.555**	**0.626**	**0.632**	**0.641**	**0.624**

words	models	Top 5 sememes
钟表匠 (clockmaker)	internal	**人(human)**, **职位(occupation)**, 部件(part), **时间(time)**, **告诉(tell)**
	external	**人(human)**, 专(ProperName), 地方(place), 欧洲(Europe), 政(politics)
	ensemble	**人(human)**, **职位(occupation)**, **告诉(tell)**, **时间(time)**, **用具(tool)**
奥斯卡 (Oscar)	internal	**专(ProperName)**, 地方(place), 市(city), 人(human), 国都(capital)
	external	**奖励(reward)**, **艺(entertainment)**, **专(ProperName)**, 用具(tool), **事情(fact)**
	ensemble	**专(ProperName)**, **奖励(reward)**, **艺(entertainment)**, 著名(famous), 地方(place)

融合HowNet义原标注的词典扩展

- 以中文LIWC (Linguistic Inquiry and Word Count) 为例，是计算社会科学中的著名词典





LIWC Results

Details of Writer: 40 year old Female
Date/Time: 6 January 2014, 1:02 am

LIWC Dimension	Your Data	Personal Texts	Formal Texts
Self-references (I, me, my)	8.33	11.4	4.2
Social words	4.17	9.5	8.0
Positive emotions	2.08	2.7	2.6
Negative emotions	1.04	2.6	1.6
Overall cognitive words	3.12	7.8	5.4
Articles (a, an, the)	2.08	5.0	7.2
Big words (> 6 letters)	20.83	13.1	19.6

The text you submitted was 96 words in length.

Your writing:

I'm newly diagnosed with type 2 diabetes. I also struggle with both calcium and uric acid kidney stones as well as the rare blood disorder LEIDEN FACTOR V. Is there anyone in this community who deals with Leiden as well as diabetes? If there is I would LOVE to be able to chat with you regarding diet and possible weight loss plans. I currently have no regular doctor and insurance so my diabetes is uncontrolled at this time. I am working hard to educate myself AND make the necessary changes to improve my current health.

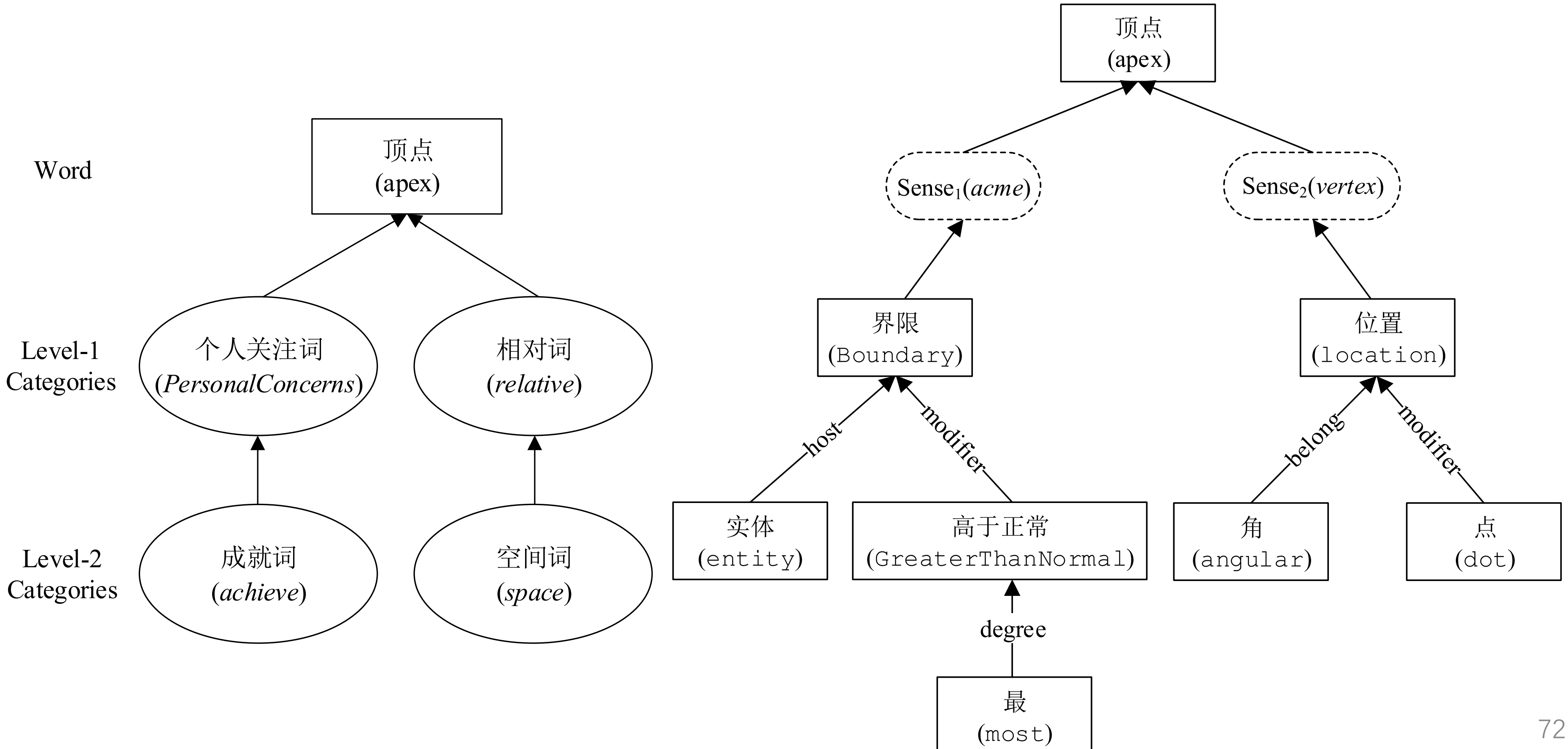
融合HowNet义原标注的词典扩展

- 以中文LIWC (Linguistic Inquiry and Word Count)为例，是计算社会科学中的著名词典
- LIWC中包含不到7000词，但中文中至少包括5万常用词

类别名称	英文简写	总词数	范例
认知历程词	cogmech	1255	理解、选择、质疑
洞察词	insight	328	了解、恍然大悟、体会
因果词	cause	128	引起、使得、变成
差距词	discrep	84	不足、纳闷、期待
暂订词	tentat	167	大约、未定、差不多
确切词	certain	145	不容置疑、必然、保证
限制词	inhib	292	废止、不准、规则
包含词	incl	82	包括、附近、添加
排除词	excl	39	取消、但是、除外

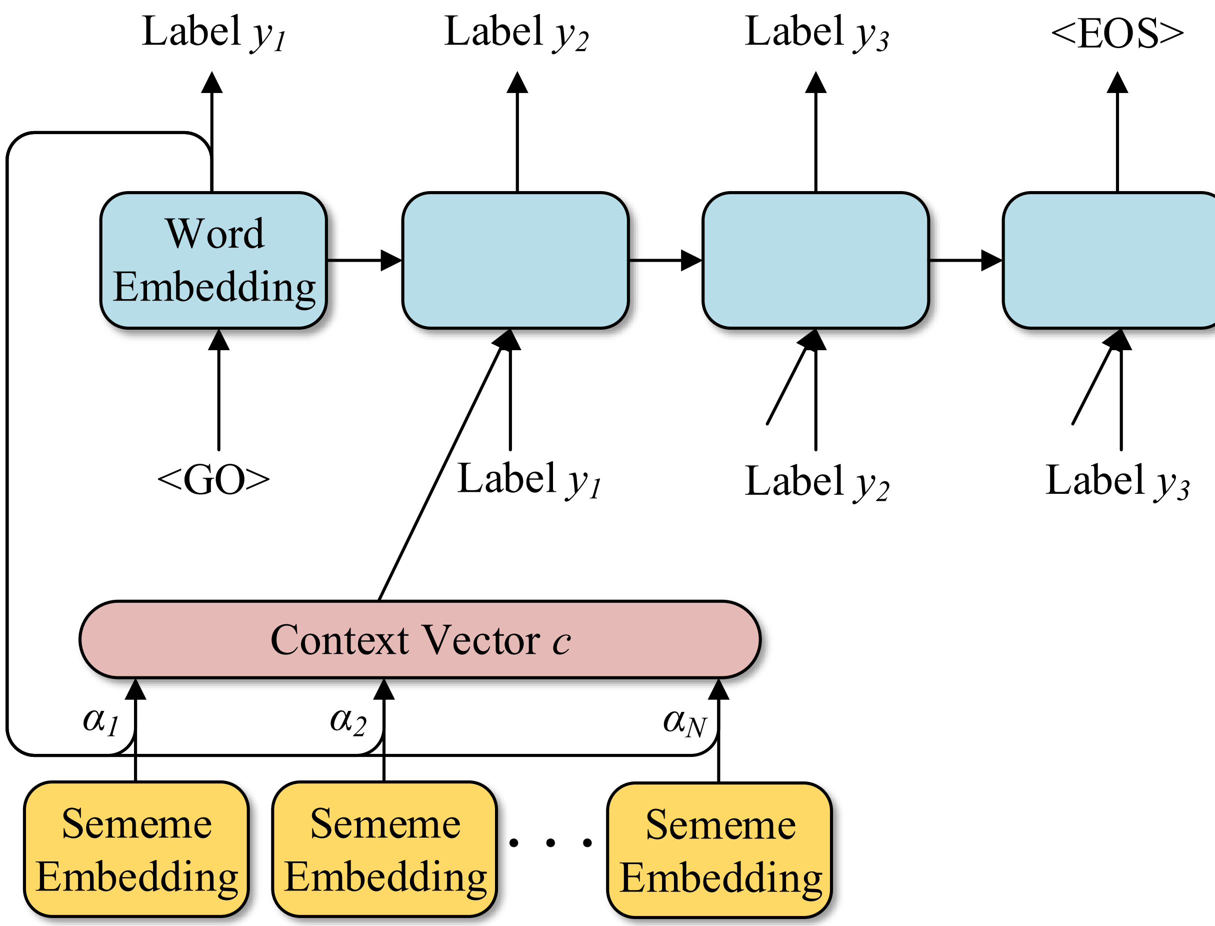
融合HowNet义原标注的词典扩展

- 以中文LIWC (Linguistic Inquiry and Word Count) 为例，是计算社会科学中的著名词典
- 可以看做对词汇的层次分类



融合HowNet义原标注的词典扩展

- Hierarchical Decoder with Sememe Attention (AAAI 2018)



实验结果

- 在CLIWC词汇层次分类任务上，我们提出的HDSA显著优于其他方法

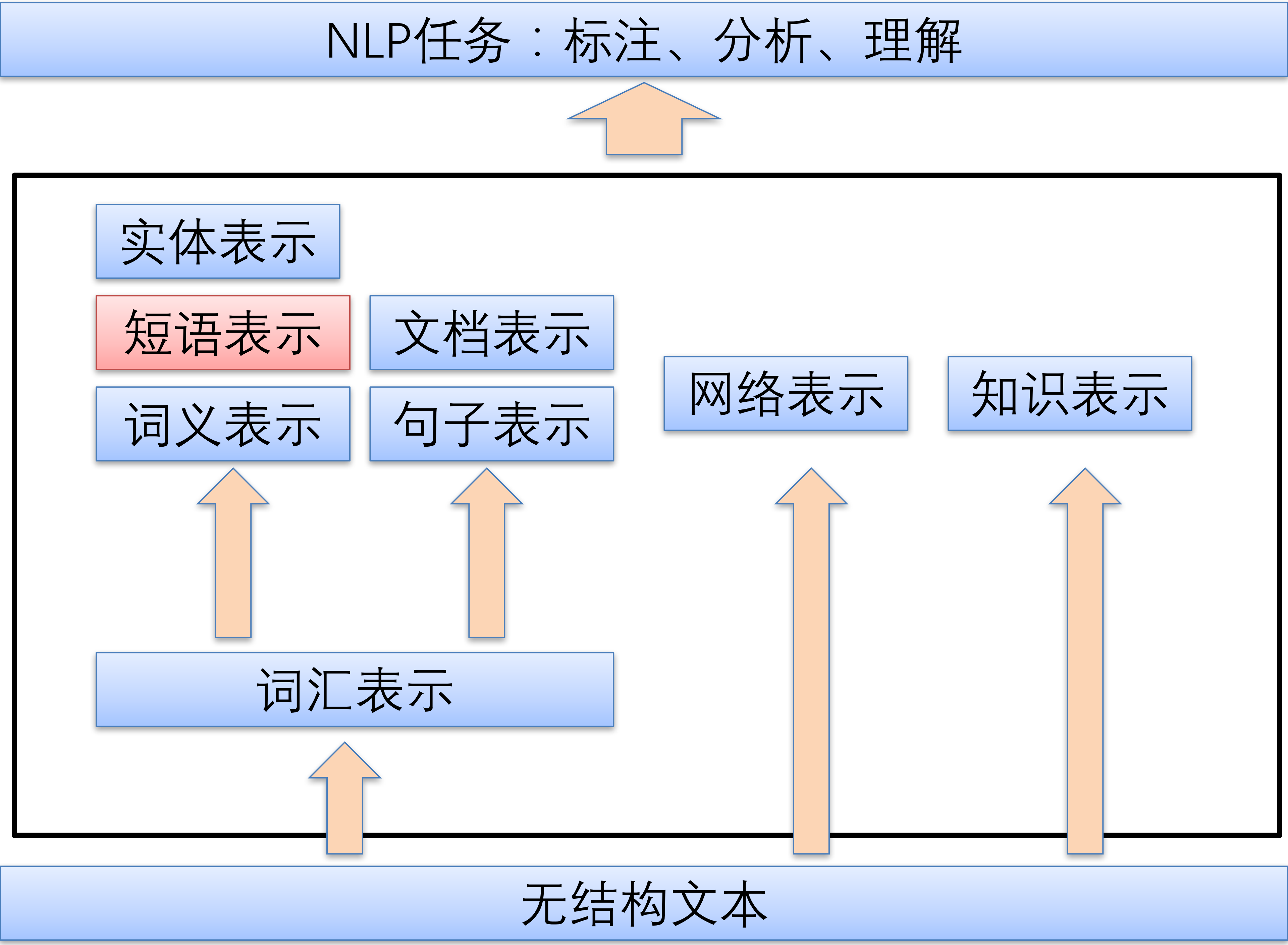
	Overall		Level 1		Level 2		Level 3	
Model	Micro-F_1	W-M-F_1	Micro-F_1	W-M-F_1	Micro-F_1	W-M-F_1	Micro-F_1	W-M-F_1
TD k-NN	0.6198	0.6169	0.6756	0.6772	0.5716	0.5646	0.4884	0.4858
TD SVM	0.6283	0.6106	0.6858	0.6785	0.5766	0.5557	0.4503	0.4142
Structural SVM	0.6444	0.6448	0.7011	0.7010	0.5919	0.5919	0.5725	0.5718
CSSA	0.6511	0.6319	0.6880	0.6864	0.6172	0.5914	0.4729	0.4322
HD	0.7023	0.7000	0.7495	0.7476	0.6658	0.6614	0.6113	0.6064
HDSA	**0.7224**	**0.7204**	**0.7636**	**0.7616**	**0.6927**	**0.6874**	**0.6270**	**0.6234**

实验结果

Word	Sememes	HD Prediction	HDSA Prediction	True Labels
恋人 (sweetheart)	交往 (associate), 人 (human), 爱恋 (love)	social←friend	social←friend, affect←posemo	social←friend, affect←posemo
今天 (today)	时间 (time), 现在 (present), 特定 (specific), 日 (day)	relativ←time	funct←TenseM←PresentM, relativ←time	funct←TenseM←PresentM, relativ←time
市镇 (town)	乡 (village), 市 (city), 地方 (place)	PersonalConcerns ←work	relativ←space	relativ←space
无望 (hopeless)	悲惨 (miserable)	cogmech←discrep	affect←negemo←sad	affect←negemo←sad
种种 (all kinds of)	多种 (various)	funct←negate	funct←quant	funct←quant
天空 (sky)	空域 (airspace)	relativ←time	relativ←space	relativ←space
联盟 (alliance)	结盟(ally), 团体(community)	PersonalConcerns ←work	social, PersonalConcerns←work	PersonalConcerns ←work
泪珠 (teardrop)	部件(part), 体液 (BodyFluid), 动物(AnimalHuman)	affect←negemo←sad	affect←negemo, bio←health	affect←negemo←sad

相关文献

- 下载地址：http://nlp.csai.tsinghua.edu.cn/~lzy/publication.html
- Huiming Jin, Hao Zhu, Zhiyuan Liu, Ruobing Xie, Maosong Sun, Fen Lin, Leyu Lin. **Incorporating Chinese Characters of Words for Lexical Sememe Prediction**. The 56th Annual Meeting of the Association for Computational Linguistics (ACL 2018).
- Xiangkai Zeng, Cheng Yang, Cunchao Tu, Zhiyuan Liu, Maosong Sun. **Chinese LIWC Lexicon Expansion via Hierarchical Classification of Word Embeddings with Sememe Attention**. The 32nd AAAI Conference on Artificial Intelligence (AAAI 2018).
- Ruobing Xie, Xingchi Yuan, Zhiyuan Liu, Maosong Sun. **Lexical Sememe Prediction via Word Embeddings and Matrix Factorization**. International Joint Conference on Artificial Intelligence (IJCAI 2017).
- Yilin Niu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. **Improved Word Representation Learning with Sememes**. The 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017).

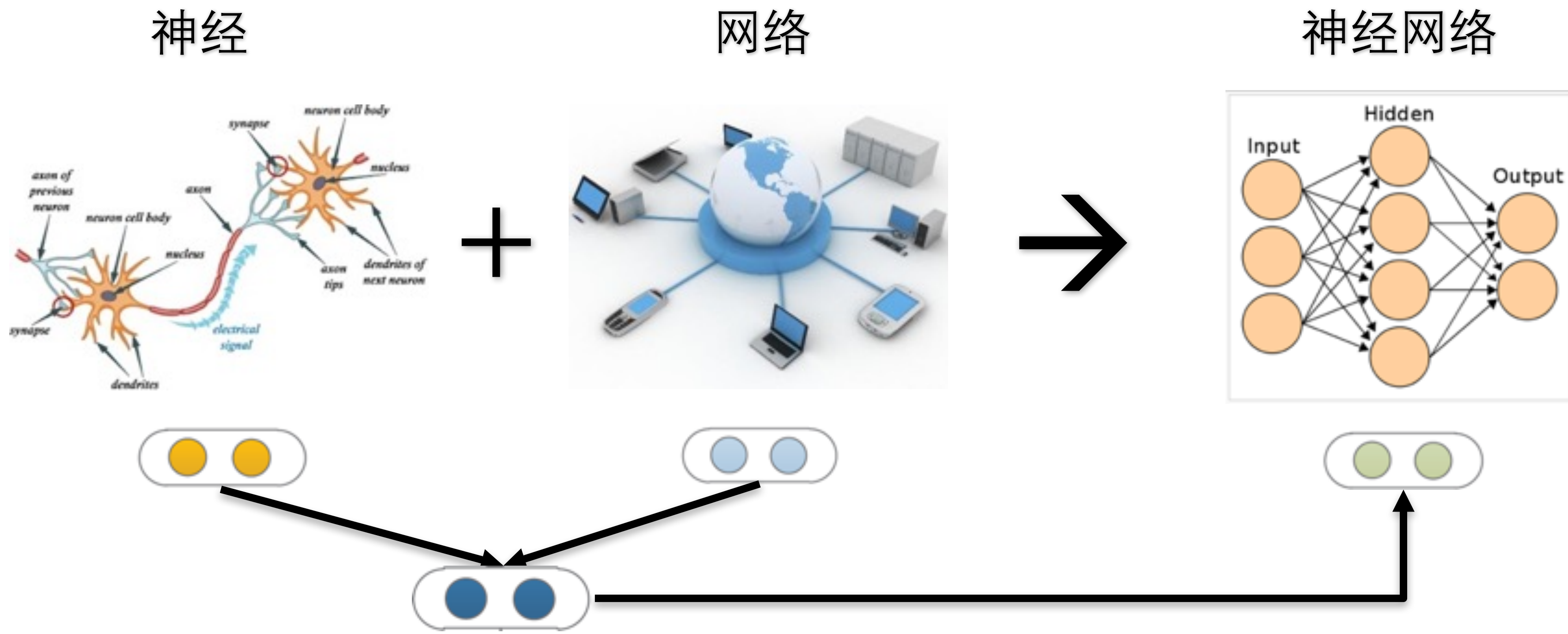

NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
网络表示
知识表示
词义表示
句子表示
词汇表示
无结构文本

基于词向量的组合语义建模

Yu Zhao, Zhiyuan Liu, Maosong Sun. Phrase Type Sensitive Tensor Indexing Model for Semantic Composition. AAAI 2015.

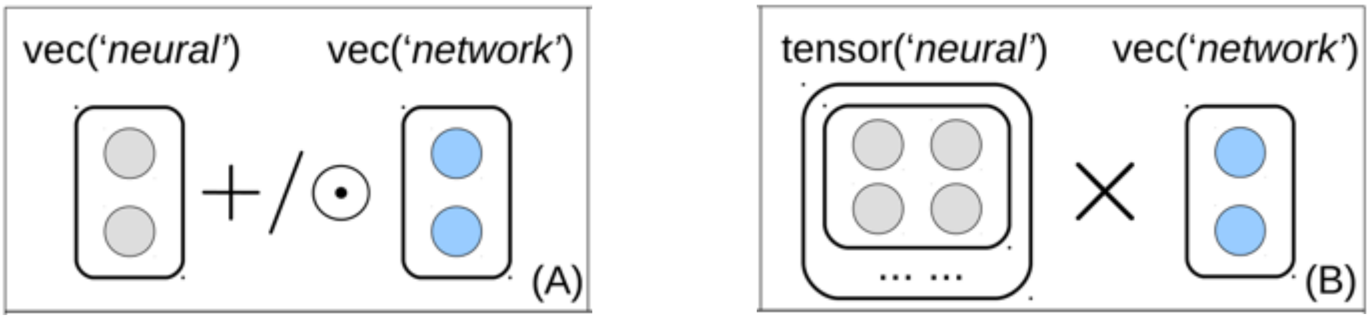
问题介绍

- 短语的语义可以由单词的词义组合而成
- 将单词表示进行组合得到高质量的短语表示



传统模型问题与解决思路

- 传统组合模型分别面临准确性低和数据稀疏问题

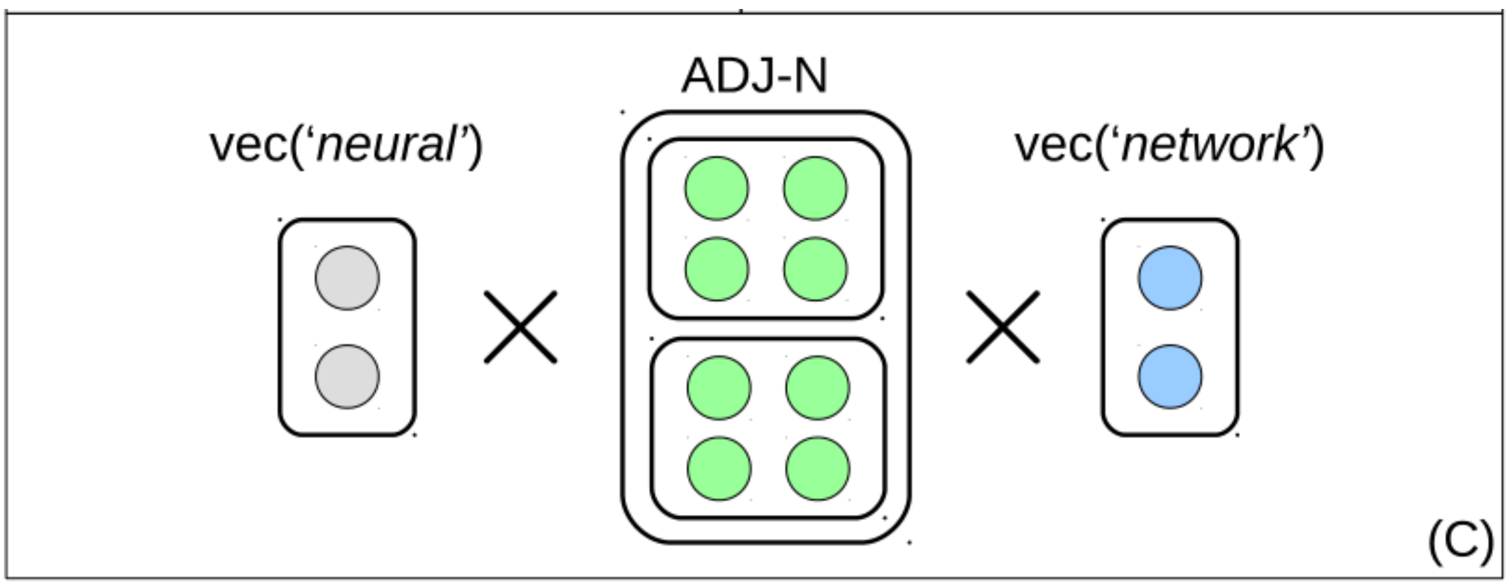


- 观察：直接将短语合并为长单词，通过词表示学习方法得到的向量较为合理，但无法对任意可能的短语都进行学习

v(神经)	**v**(网络)	**v**(神经) + **v**(网络)	**v**(神经网络)
皮层	广播网	基因调控	反向传播
突触	有线电视网	细胞	联结
传感系统	IPTV	线路交换	脉冲耦合神经网络

提出方法

- 三步策略
  - 提取高频短语
  - 同时学习单词和现有短语的表示
  - 训练组合模型
    - 使用张量表示组合函数



短语抽取

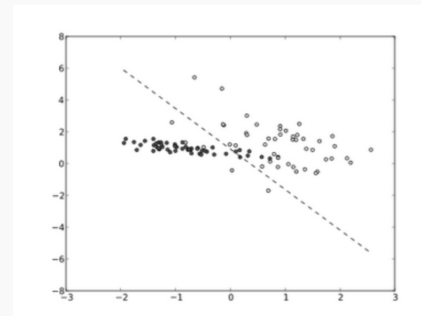
- 通过大规模自然标注，获取确定的短语边界

数据挖掘 [编辑]

数据挖掘（Data mining），又译为**资料探勘、数据挖掘、数据采矿**。它是**数据库知识发现**（英文：Knowledge-Discovery in Databases，缩写：KDD）中的一个步骤。数据挖掘一般是指从大量的数据中自动搜索隐藏于其中的有着特殊关系性（属于Association rule learning）的信息的过程。数据挖掘通常与计算机科学有关，并通过统计、在线分析处理、情报检索、机器学习、专家系统（依靠过去的经验法则）和模式识别等诸多方法来实现上述目标。

目录 [隐藏]

机器学习与数据挖掘



问题

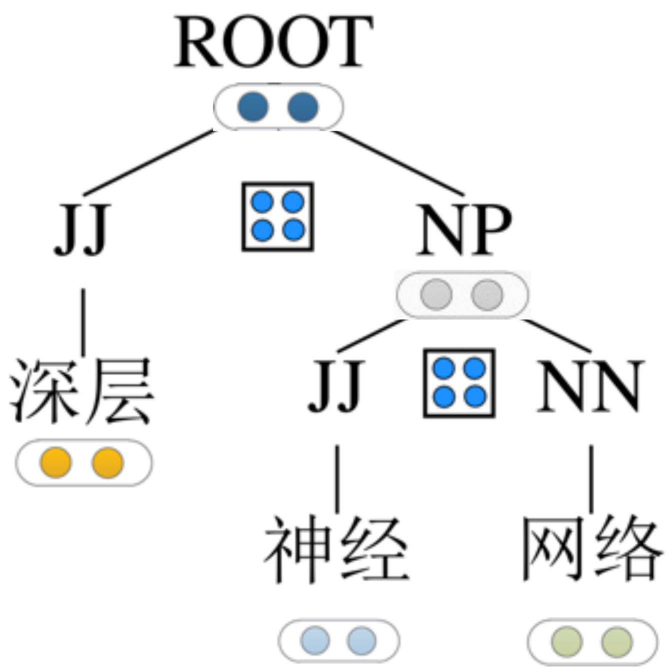
分类 • 聚类 • 回归 • 异常检测 • 关联规则 • 强化学习 • 结构预测 • 特征学习 • 在线学习 • 半监督学习 • 语法归纳

使用知识库提取出24万个高频短语
(针对三种短语类型: ADJ-N、N-N、V-N)

锚短语

模型扩展

- 该模型具有良好的可扩展性，可以延伸至句子、段落的语义表示



实验设置

- 短语间相似度的计算 [Lapata 2010 数据集]
  - 324 个短语对 (Adj-N, N-N , V-Obj)

- 评价公式
  - 对短语p1、p2的向量表示 Vc(p1)、Vc(p2) 计算余弦相似度

$$Sim_C(p_1, p_2) = \frac{V_C(p_1) \cdot V_C(p_2)}{|V_C(p_1)| \cdot |V_C(p_2)|}$$

  - 以人工标注结果为标准答案

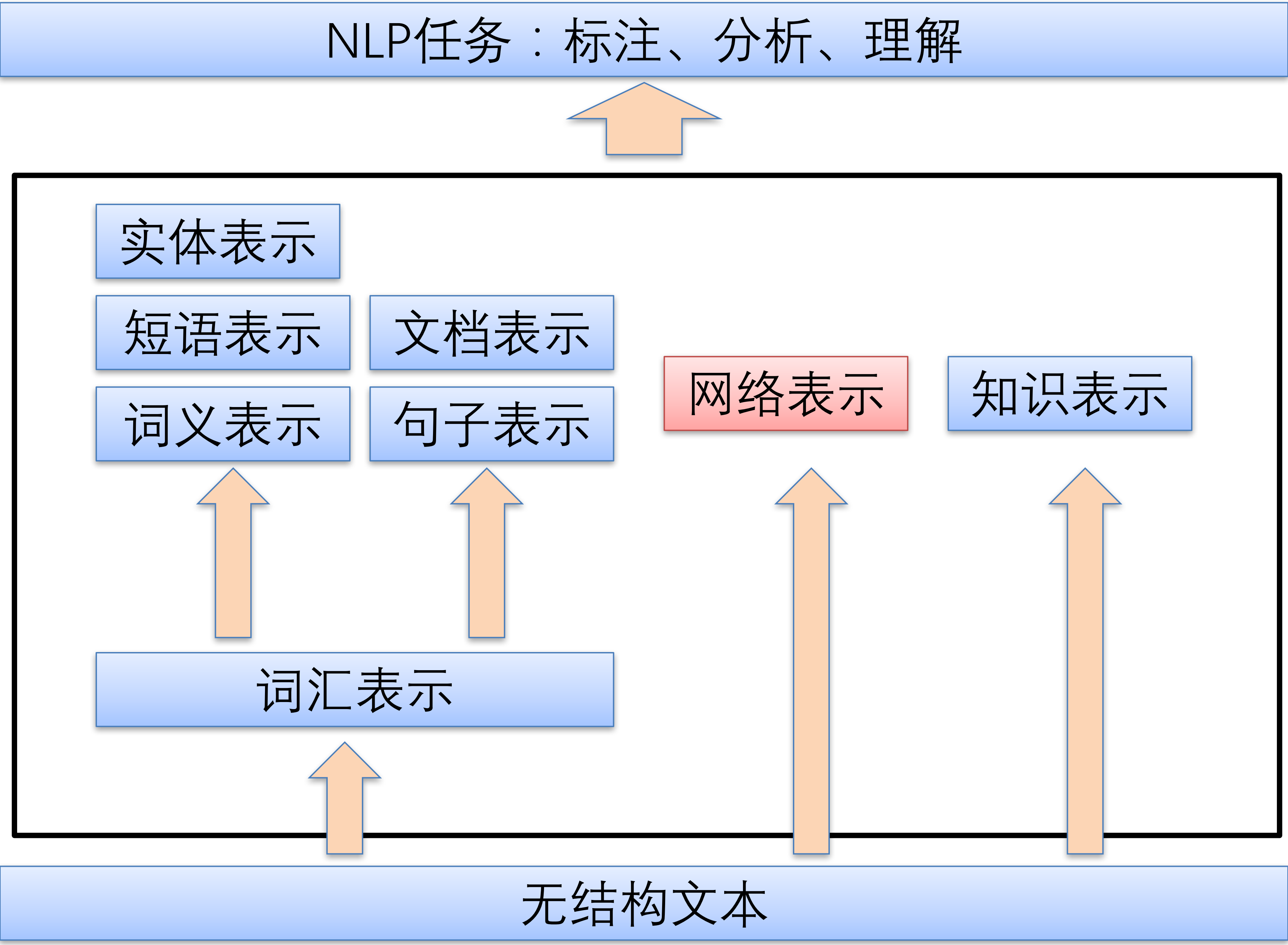
实验结果

- 我们的TIM模型与人工标注的相关系数最高

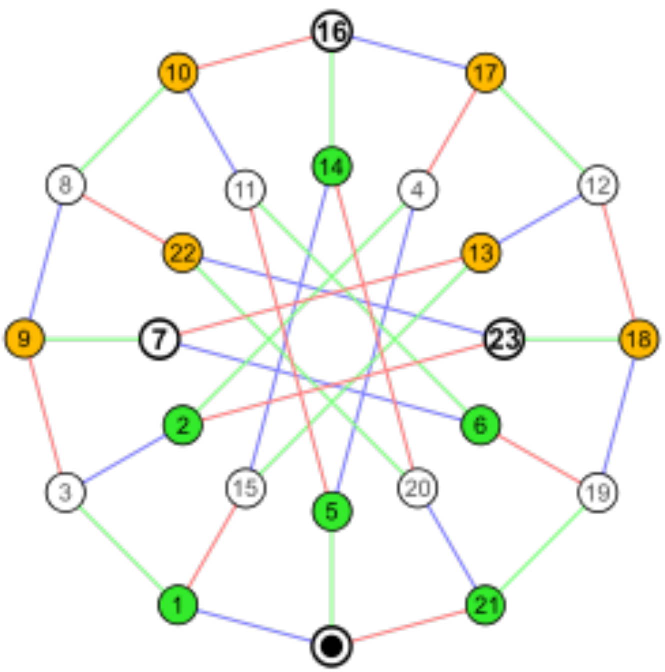
w.r	c.m	Adj-N	N-N	V-Obj
SDS (BNC)	ADD	0.37	0.38	0.28
	MUL	0.48	0.50	0.35
	RAE	0.31	0.30	0.28
SDS (Wiki)	ADD	0.34	0.44	0.13
	MUL	0.42	0.63	0.28
	AT	0.42	0.45	-
	FT	0.49	0.57	-
SGM (Wiki)	ADD	0.73	0.73	0.62
	MUL	0.39	0.34	0.41
	AT	0.36	0.51	-
	FT	0.57	0.66	-
	TIM	**0.77**	**0.75**	**0.66**

词汇级表示的研究趋势

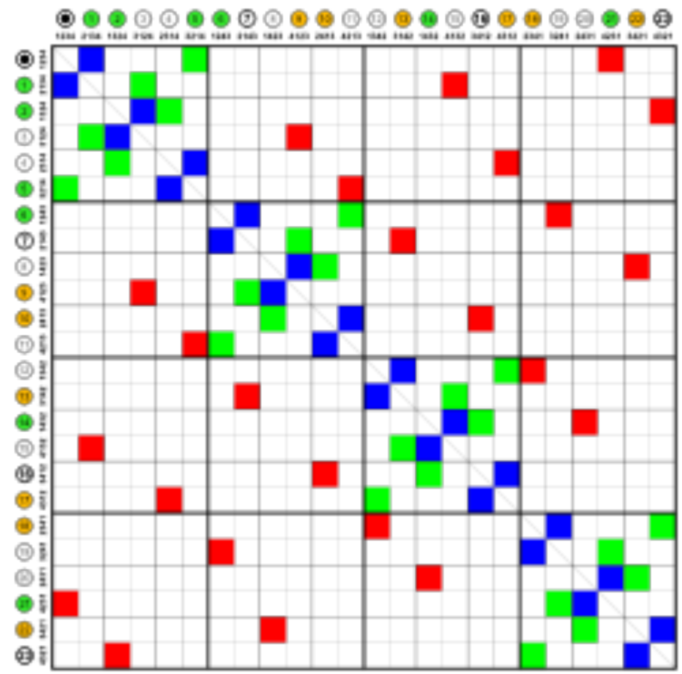
- 引入外部信息辅助表示学习
  - 词典自然语言定义
  - 相关概念图像信息
- 引入人类先验结构化语言知识
  - WordNet、HowNet、同义词词林等
- 辅助构建词汇级语言知识库
  - 知识指导的深度学习模型


NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
词义表示
句子表示
网络表示
知识表示
词汇表示
无结构文本

传统网络表示的问题



N个节点的网络



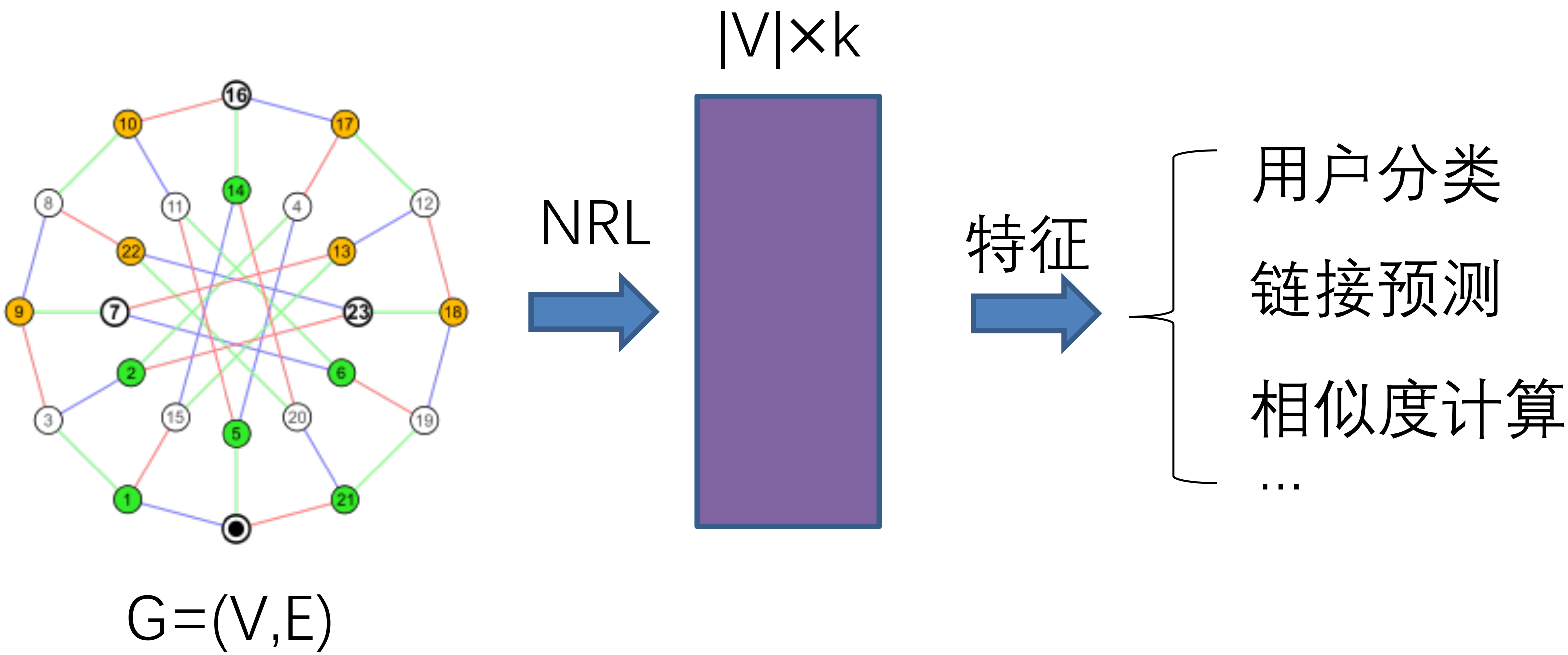
邻接矩阵

需要N×N个元素来表示

稀疏！不利于存储计算

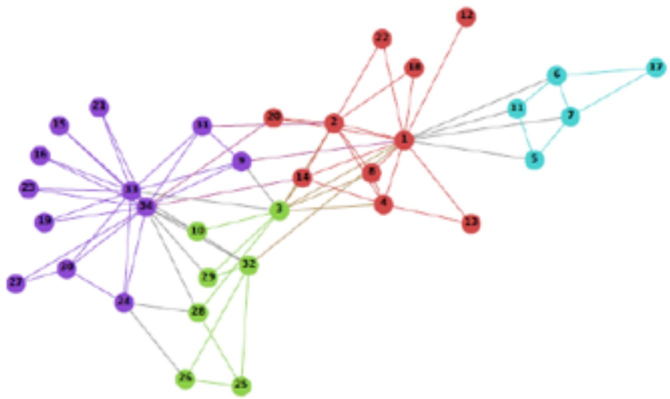
网络表示学习

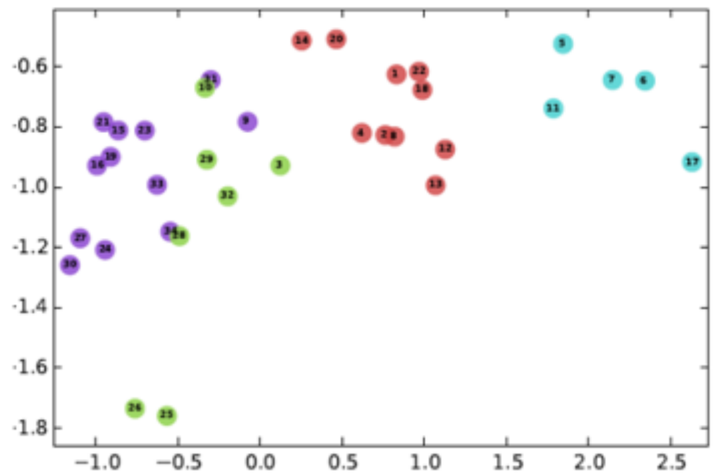
- 将网络中节点的语义信息表示为低维向量



网络表示学习

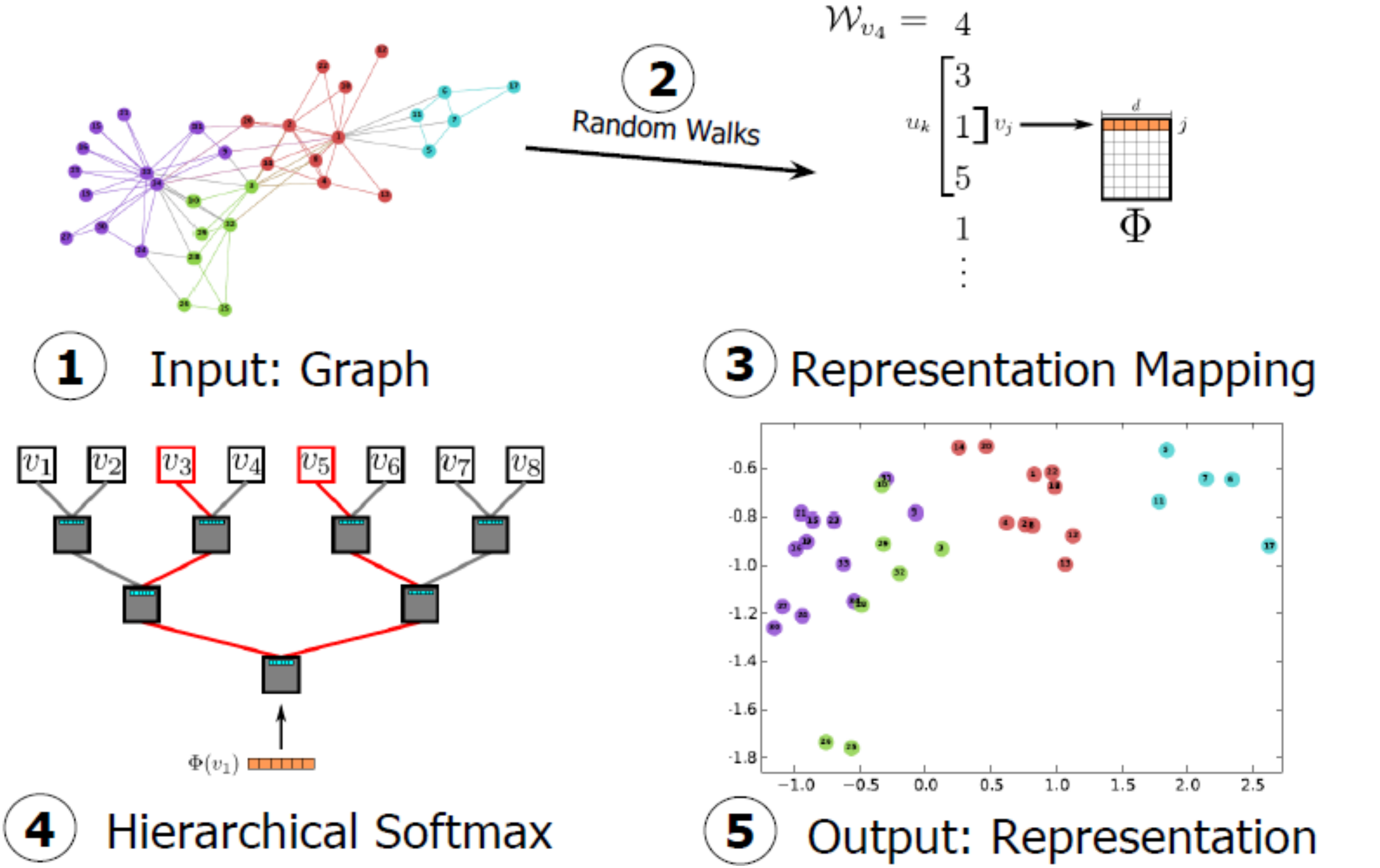
- 跆拳道俱乐部社会网络 (k=2)





Perozzi et al. DeepWalk: Online Learning of Social Representations. KDD 2014

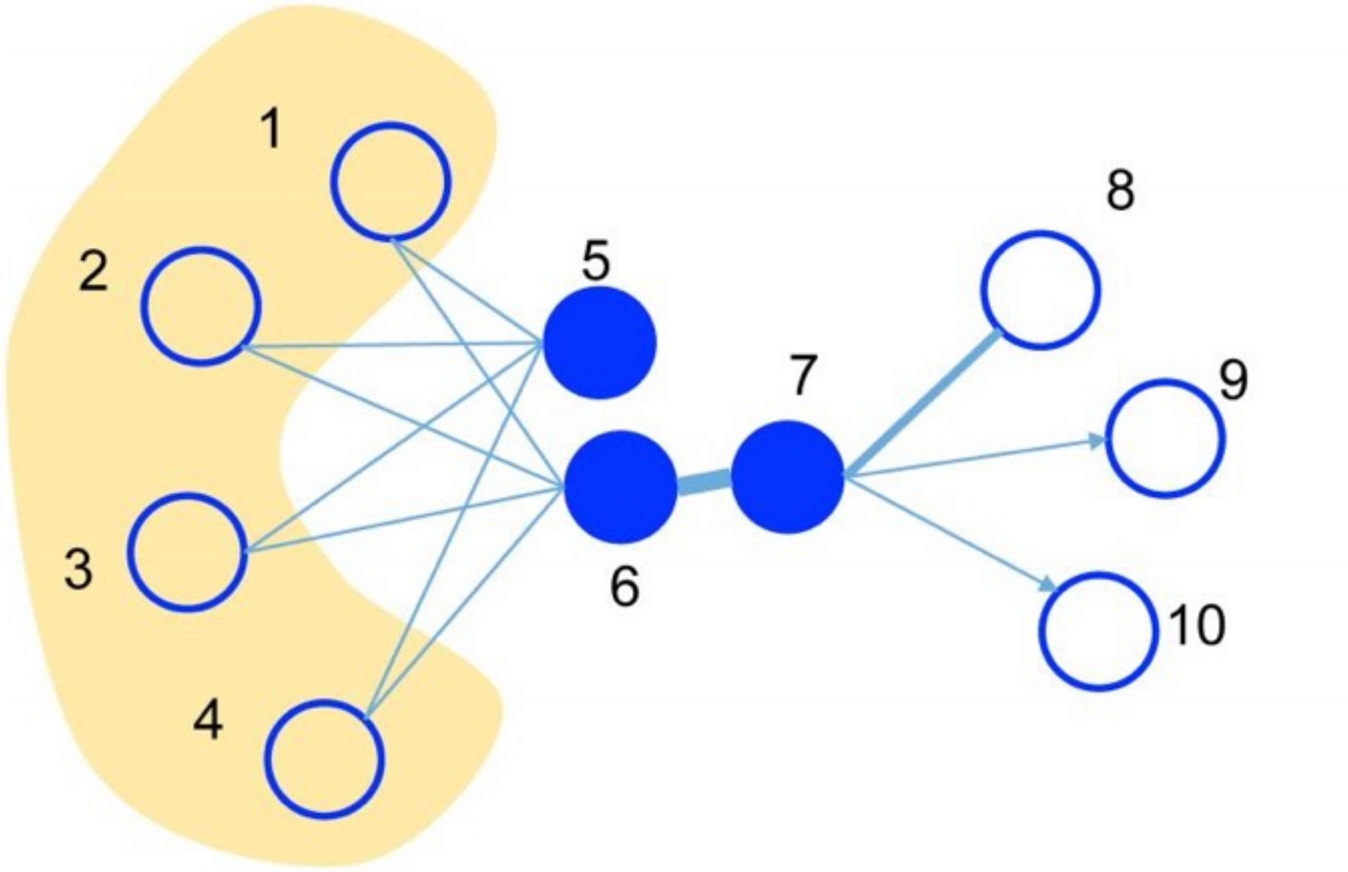
DeepWalk



$\mathcal{W}_{v_4} = 4$
3
u_k
1
v_j
5
1
2
Random Walks
d
j
Φ
1 Input: Graph
3 Representation Mapping
v_1 v_2 v_3 v_4 v_5 v_6 v_7 v_8
$\Phi(v_1)$
4 Hierarchical Softmax
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
−1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5
5 Output: Representation

LINE

- 一阶和二阶邻近度

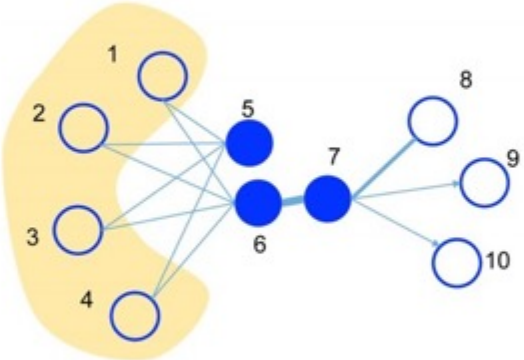


Tang et al. LINE: Large-scale Information Network Embedding. WWW 2015

LINE

- 一阶邻近度
  - 由强关系连接的6和7表示应该相近

$$p_1(v_i, v_j) = \frac{1}{1 + \exp(-\vec{u}_i^T \cdot \vec{u}_j)}$$

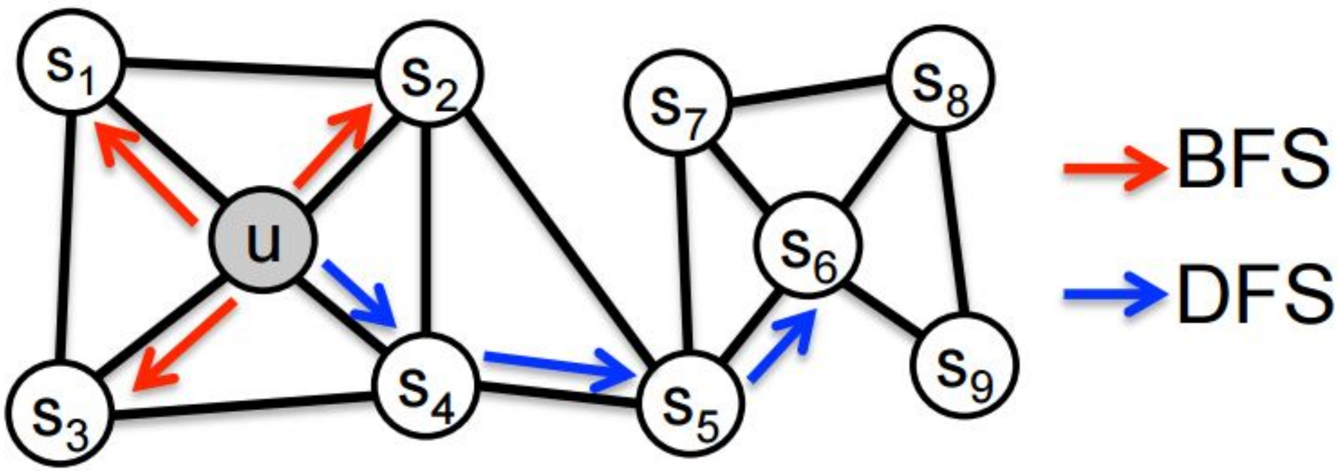


- 二阶邻近度
  - 共同邻居多的5和6的表示应该相近

$$p_2(v_j | v_i) = \frac{\exp(\vec{u}_j'^T \cdot \vec{u}_i)}{\sum_{k=1}^{|V|} \exp(\vec{u}_k'^T \cdot \vec{u}_i)}$$

node2vec

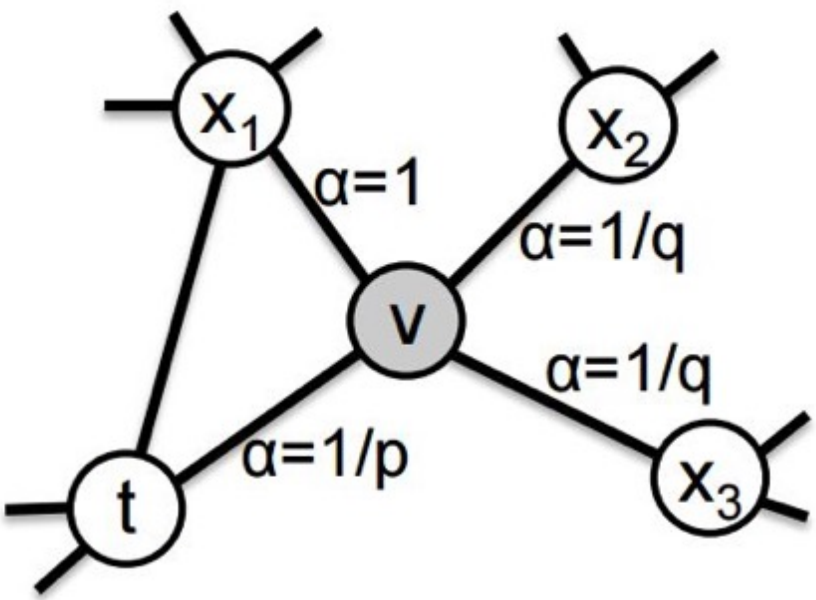
- 随机游走策略
  - 宽度优先搜索：微观局部信息
  - 深度优先搜索：宏观全局信息



Grover and Leskovec, node2vec: Scalable Feature Learning for Networks. KDD 2016

node2vec

- 参数控制的随机游走
  - 返回概率参数p，对应BFS
  - 离开概率参数q，对应DFS



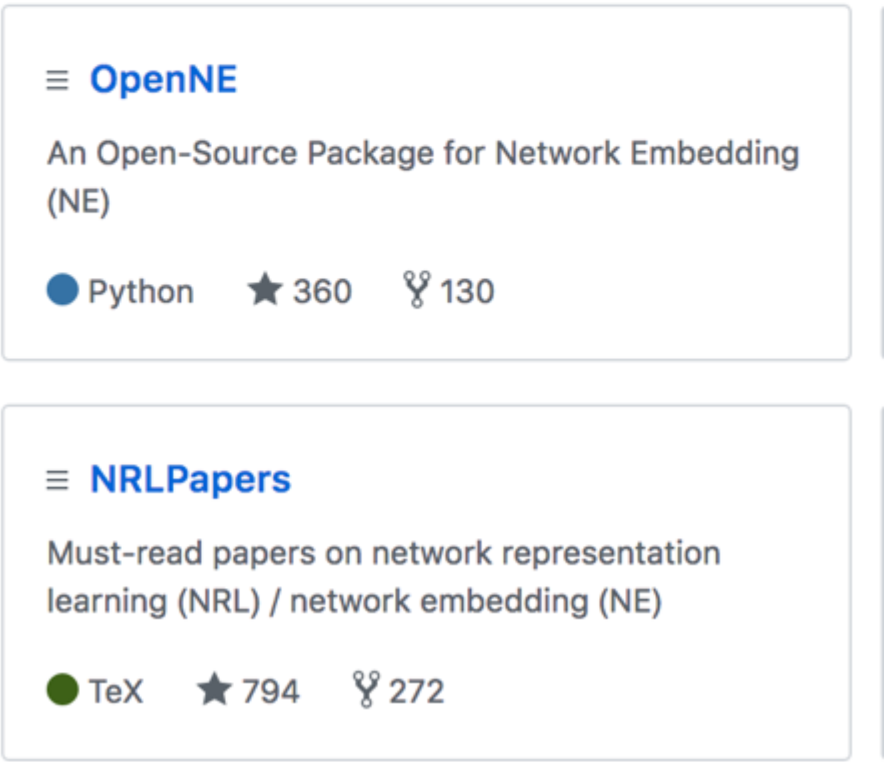
$$\alpha_{pq}(t,x) = \begin{cases} \frac{1}{p} & \text{if } d_{tx} = 0 \\ 1 & \text{if } d_{tx} = 1 \\ \frac{1}{q} & \text{if } d_{tx} = 2 \end{cases}$$

性能比较

Algorithm	**Dataset**		
	BlogCatalog	PPI	Wikipedia
Spectral Clustering	0.0405	0.0681	0.0395
DeepWalk	0.2110	0.1768	0.1274
LINE	0.0784	0.1447	0.1164
node2vec	**0.2581**	**0.1791**	**0.1552**

相关资源

- NRL Papers:
  – https://github.com/thunlp/NRLPapers
- OpenNE
  – https://github.com/thunlp/OpenNE

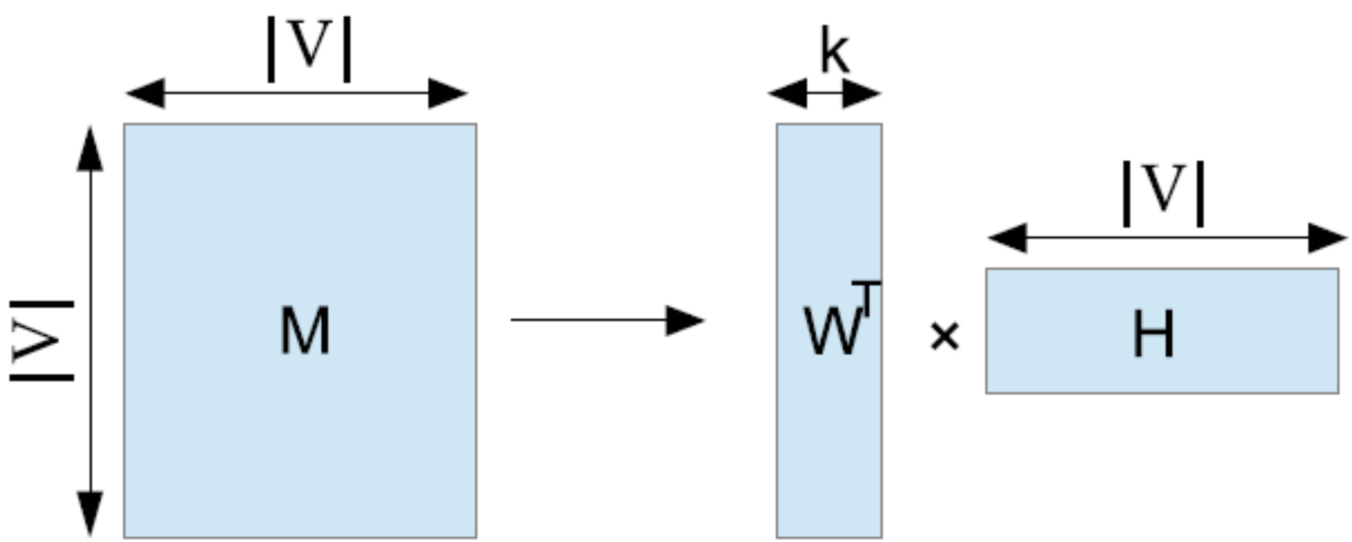


网络表示学习与矩阵分解

Cheng Yang, Maosong Sun, Zhiyuan Liu, Cunchao Tu. Fast Network Embedding Enhancement via High Order Proximity Approximation. International Joint Conference on Artificial Intelligence (IJCAI 2017).

DeepWalk算法和矩阵分解

- 我们从数学上证明了当前最好的网络表示学习算法DeepWalk等价于矩阵分解
- DeepWalk的矩阵分解形式理解



网络表示学习与矩阵分解的关系

- 构建矩阵M时
  - 引入的高阶邻近度越多效果越好
  - 高阶邻近度的计算越精确，节点表示效果越好
  - 但是计算精确的高阶邻近度复杂度高

Table 1: Comparisons among three NRL methods.

	SC	DeepWalk	GraRep
Proximity Matrix	L	$\sum_{k=1}^{K} \frac{A^k}{K}$	$A^k, k = 1 \dots K$
Computation	Accurate	Approximate	Accurate
Scalability	Yes	Yes	No
Performance	Low	Middle	High

问题：如何近似计算高阶邻近度矩阵M，同时不提高计算复杂度？

高阶邻近度近似的快速网络表示提升

- 解决思路
  - 不直接计算构建矩阵M
  - 而是更新已经训练好的网络表示R，其中A是行归一化的邻接矩阵：

$$R' = R + \lambda A \cdot R$$

  - 定理：如果已经训练好的网络表示R是通过分解最高阶为K的矩阵得到，那么更新后的表示R’实际上等价于分解最高阶为K+2的矩阵得到。

高阶邻近度近似的快速网络表示提升

- 分类实验结果

Table 2: Classification results on Cora dataset.

% Labeled Nodes	% Accuracy 10%	50%	90%	Time (s)
GF	50.8 (**68.0**)	61.8 (**77.0**)	64.8 (**77.2**)	4 (+0.1)
SC	55.9 (**68.7**)	70.8 (**79.2**)	72.7 (**80.0**)	1 (+0.1)
DeepWalk_{low}	71.3 (76.2)	76.9 (81.6)	78.7 (81.9)	31 (+0.1)
DeepWalk_{mid}	68.9 (**76.7**)	76.3 (82.0)	78.8 (84.3)	69 (+0.1)
DeepWalk_{high}	68.4 (**76.1**)	74.7 (80.5)	75.4 (81.6)	223 (+0.1)
LINE_{1st}	64.8 (70.1)	76.1 (80.9)	78.9 (82.2)	62 (+0.1)
LINE_{2nd}	63.3 (**73.3**)	73.4 (80.1)	75.6 (80.3)	67 (+0.1)
node2vec	76.9 (77.5)	81.0 (81.6)	81.4 (81.9)	56 (+0.1)
TADW	78.1 (84.4)	83.1 (86.6)	82.4 (87.7)	2 (+0.1)

高阶邻近度近似的快速网络表示提升

- 链接预测实验结果

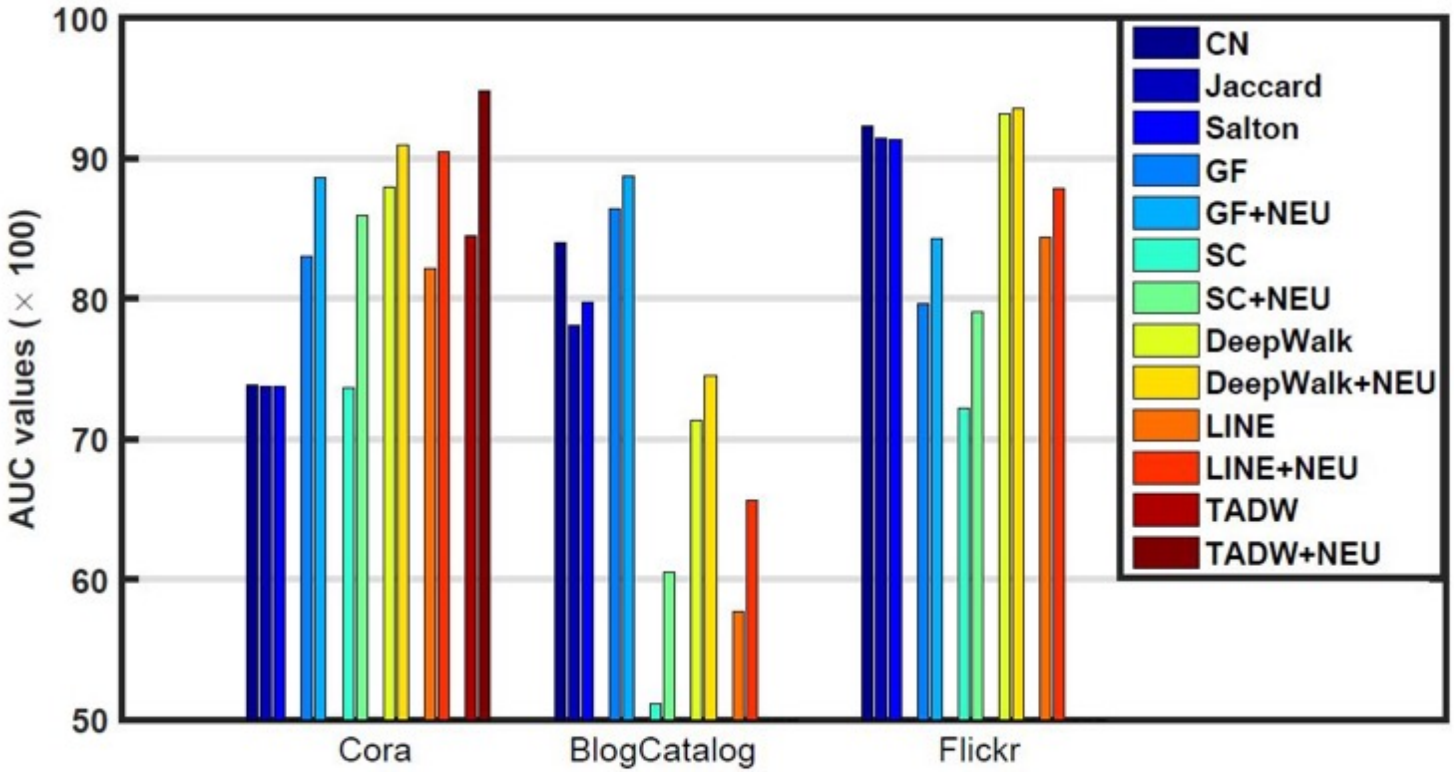
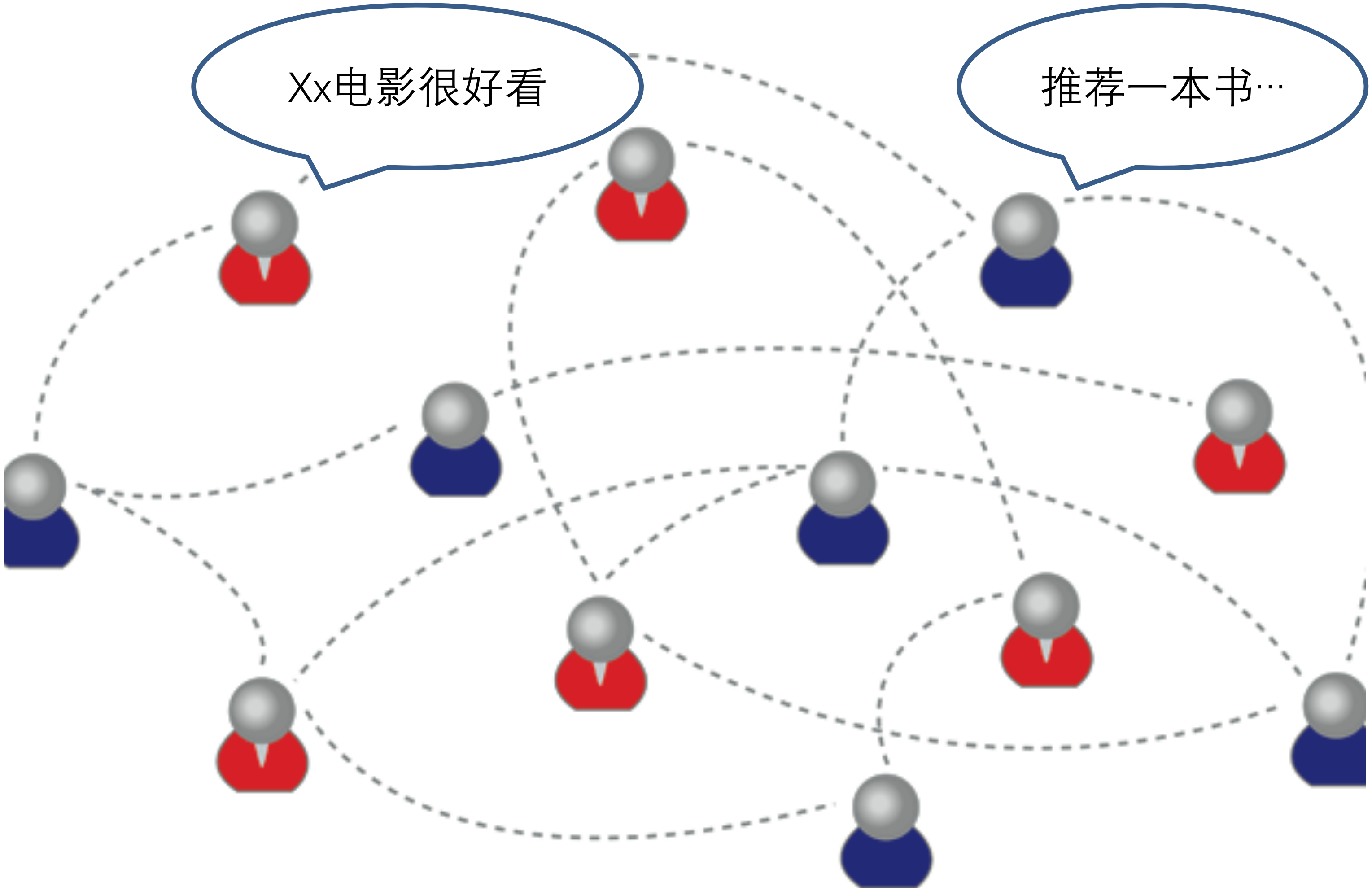


Figure 1: Experimental results on link prediction.

丰富文本信息下的网络表示学习

Cheng Yang, Zhiyuan Liu, Deli Zhao, Maosong Sun, Edward Y. Chang Network Representation Learning with Rich Text Information (IJCAI 2015)

网络结构外部信息



现实网络中往往存在着丰富的文本信息

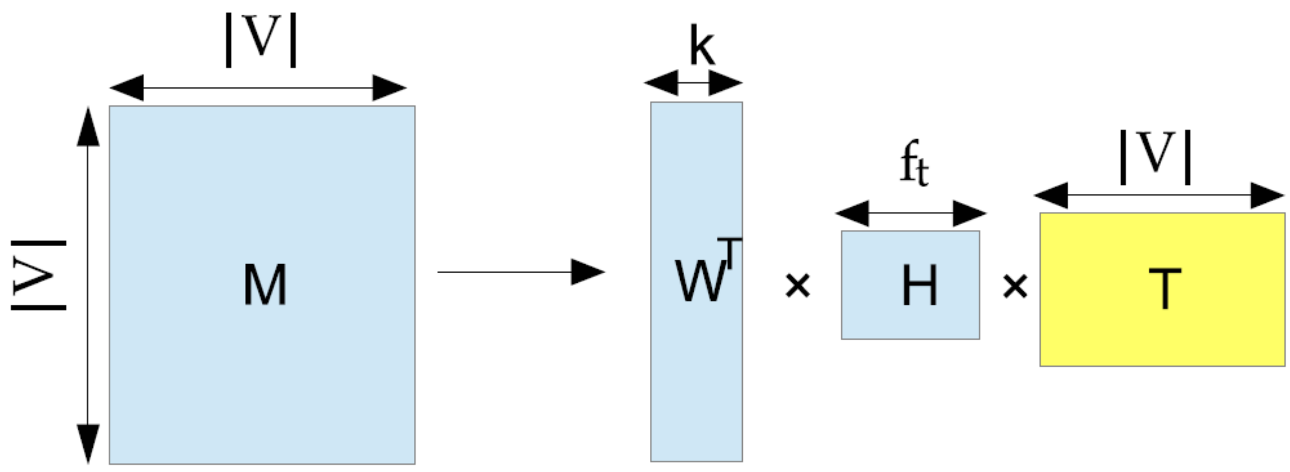
研究挑战

- 创新：将网络结构外的文本信息引入网络表示学习
- 挑战1：已有工作的框架中很难方便引入外部信息，如DeepWalk (Perozzi et al. KDD 2014)
- 挑战2：需要将网络结构信息和文本信息融为一个整体，统一建立模型

Text-Associated DeepWalk（TADW）

- 在矩阵分解的基础上，我们设计了Text-Associated DeepWalk算法：

$$\min_{W,H} ||M - W^T HT||_F^2 + \frac{\lambda}{2}(||W||_F^2 + ||H||_F^2).$$



其中T是收集文本信息得到的TFIDF矩阵，W、H是参数
TADW将W和HT拼接作为2k长度的网络表示

实验结果：Cora

- 我们将网络表示作为节点的特征(feature)，用SVM算法做分类，测试节点分类正确率

Table 1: Evaluation results on Cora dataset.

Classifier	Transductive SVM				SVM				
% Labeled Nodes	1%	3%	7%	10%	10%	20%	30%	40%	50%
DeepWalk	62.9	68.3	72.2	72.8	76.4	78.0	79.5	80.5	81.0
PLSA	47.7	51.9	55.2	60.7	57.0	63.1	65.1	66.6	67.6
Text Features	33.0	43.0	57.1	62.8	58.3	67.4	71.1	73.3	74.0
Naive Combination	67.4	70.6	75.1	77.4	76.5	80.4	82.3	83.3	84.1
NetPLSA	65.7	67.9	74.5	77.3	80.2	83.0	84.0	84.9	85.4
TADW	**72.1**	**77.0**	**79.1**	**81.3**	**82.4**	**85.0**	**85.6**	**86.0**	**86.7**

实验结果：Citeseer

Table 2: Evaluation results on Citeseer dataset.

Classifier	Transductive SVM				SVM				
% Labeled Nodes	1%	3%	7%	10%	10%	20%	30%	40%	50%
DeepWalk	-	-	49.0	52.1	52.4	54.7	56.0	56.5	57.3
PLSA	45.2	49.2	53.1	54.6	54.1	58.3	60.9	62.1	62.6
Text Features	36.1	49.8	57.7	62.1	58.3	66.4	69.2	71.2	72.2
Naive Combination	39.0	45.7	58.9	61.0	61.0	66.7	69.1	70.8	72.0
NetPLSA	45.4	49.8	52.9	54.9	58.7	61.6	63.3	64.0	64.7
TADW	**63.6**	**68.4**	**69.1**	**71.1**	**70.6**	**71.9**	**73.3**	**73.7**	**74.2**

“-” indicates TSVM can not converge in 12 hours because of low quality of representation (TSVM can always converge in 5 minutes for TADW).

实验结果

Table 3: Evaluation results on Wiki dataset.

Classifier	SVM						
% Labeled Nodes	3%	7%	10%	20%	30%	40%	50%
DeepWalk	48.4	56.6	59.3	64.3	66.2	68.1	68.8
PLSA	58.3	66.5	69.0	72.5	74.7	75.5	76.0
Text Features	46.7	60.8	65.1	72.9	75.6	77.1	77.4
Naive Combination	48.7	62.6	66.3	73.0	75.2	77.1	78.6
NetPLSA	56.3	64.6	67.2	70.6	71.7	71.9	72.3
TADW (k=100)	**59.8**	**68.2**	**71.6**	**75.4**	**77.3**	**77.7**	**79.2**
TADW (k=200)	**60.4**	**69.9**	**72.6**	**77.3**	**79.2**	**79.9**	**80.3**

举例

- Cora的样例数据
  - 论文标题: Irrelevant Features and the Subset Selection Problem
  - 论文分组: Theory

Top 5 nearest documents by DeepWalk	
Title	Class Label
Feature selection methods for classifications	Neural Network
Automated model selection	Rule Learning
Compression-Based Feature Subset Selection	Theory
Induction of Condensed Determinations	Case Based
MLC Tutorial A Machine Learning library of C classes	Theory

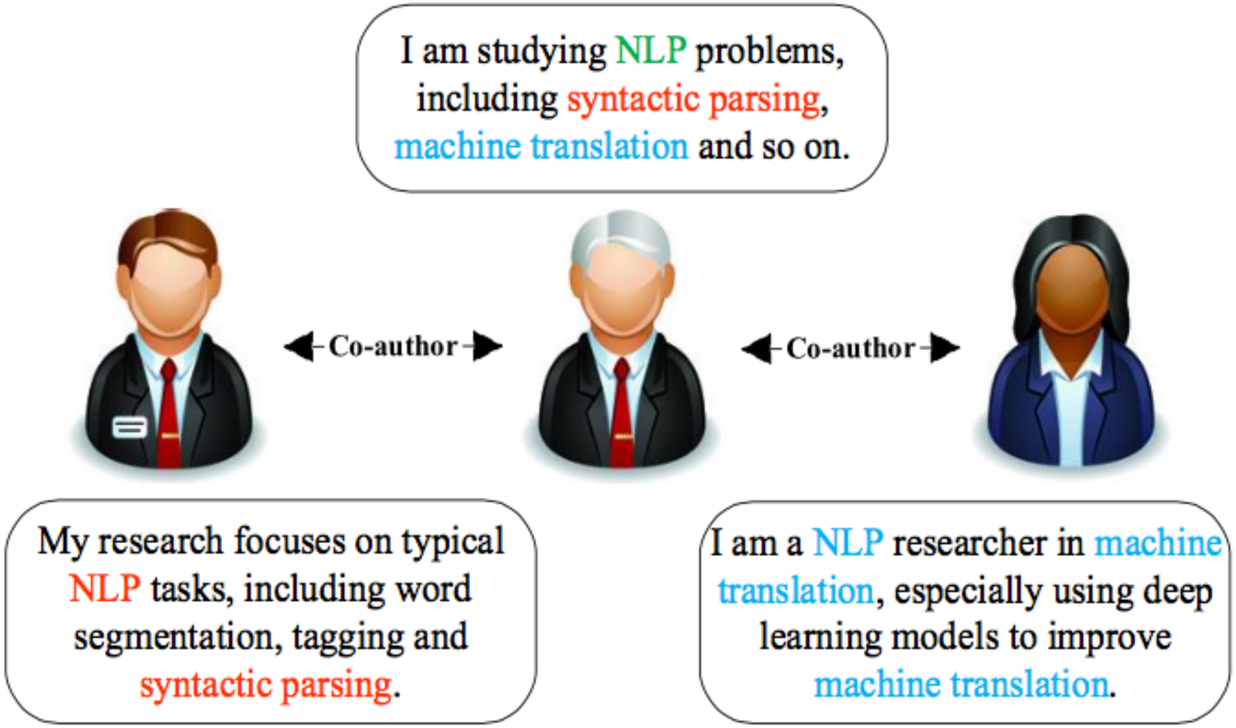
Top 5 nearest documents by TADW	
Title	Class Label
Feature subset selection as search with probabilistic estimates	Theory
Compression-Based Feature Subset Selection	Theory
Selection of Relevant Features in Machine Learning	Theory
NP-Completeness of Searches for Smallest Possible Feature Sets	Theory
Feature subset selection using a genetic algorithm	Genetic Algorithms

语境敏感的网络表示学习

Cunchao Tu, Han Liu, Zhiyuan Liu, Maosong Sun. CANE: Context-Aware Network Embedding for Relation Modeling. The 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017).

语境感知问题

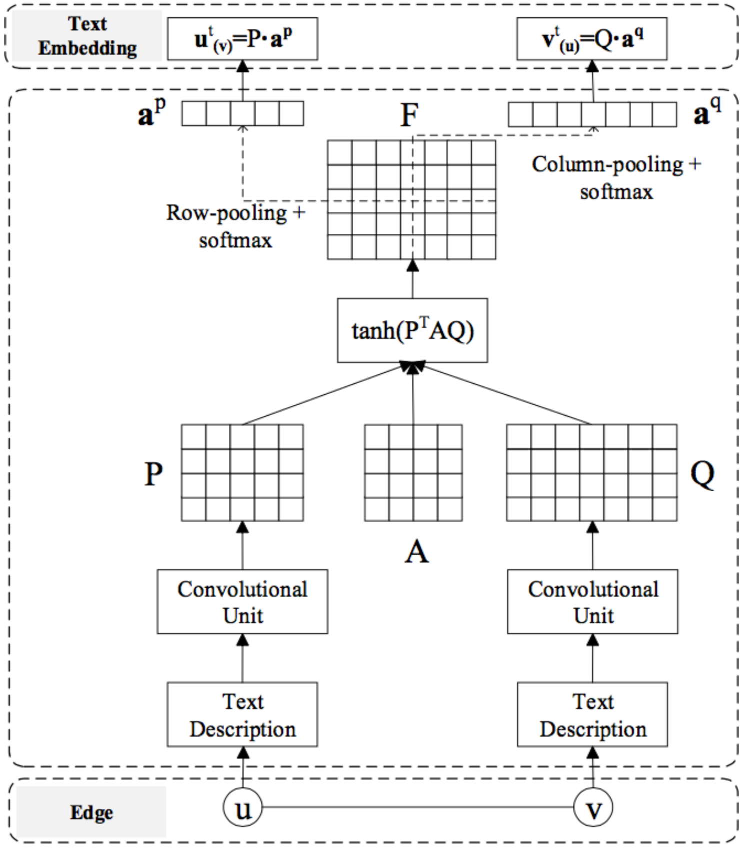
- 节点与邻居节点进行交互时，展现出不同方面



传统NRL简单使边上的两个节点表示相近
不能很好地对具体关系建模

Context-Aware Network Embedding

- 根据节点不同邻居，学习不同的向量表示
- 利用文本信息进行相互关注（mutual attention）



Context-Aware Network Embedding

- 链接预测结果

%Removed edges	15%	25%	35%	45%	55%	65%	75%	85%	95%
DeepWalk	55.2	66.0	70.0	75.7	81.3	83.3	87.6	88.9	88.0
LINE	53.7	60.4	66.5	73.9	78.5	83.8	87.5	87.7	87.6
node2vec	57.1	63.6	69.9	76.2	84.3	87.3	88.4	89.2	89.2
Naive Combination	78.7	82.1	84.7	88.7	88.7	91.8	92.1	92.0	92.7
TADW	87.0	89.5	91.8	90.8	91.1	92.6	93.5	91.9	91.7
CENE	86.2	84.6	89.8	91.2	92.3	91.8	93.2	92.9	93.2
CANE (text only)	83.8	85.2	87.3	88.9	91.1	91.2	91.8	93.1	93.5
CANE (w/o attention)	84.5	89.3	89.2	91.6	91.1	91.8	92.3	92.5	93.6
CANE	**90.0**	**91.2**	**92.0**	**93.0**	**94.2**	**94.6**	**95.4**	**95.7**	**96.3**

Table 3: AUC values on HepTh. ($\alpha = 0.7, \beta = 0.2, \gamma = 0.2$)

%Removed edges	15%	25%	35%	45%	55%	65%	75%	85%	95%
DeepWalk	56.6	58.1	60.1	60.0	61.8	61.9	63.3	63.7	67.8
LINE	52.3	55.9	59.9	60.9	64.3	66.0	67.7	69.3	71.1
node2vec	54.2	57.1	57.3	58.3	58.7	62.5	66.2	67.6	68.5
Naive Combination	55.1	56.7	58.9	62.6	64.4	68.7	68.9	69.0	71.5
TADW	52.3	54.2	55.6	57.3	60.8	62.4	65.2	63.8	69.0
CENE	56.2	57.4	60.3	63.0	66.3	66.0	70.2	69.8	73.8
CANE (text only)	55.6	56.9	57.3	61.6	63.6	67.0	68.5	70.4	73.5
CANE (w/o attention)	56.7	59.1	60.9	64.0	66.1	68.9	69.8	71.0	74.3
CANE	**56.8**	**59.3**	**62.9**	**64.5**	**68.9**	**70.4**	**71.4**	**73.6**	**75.4**

Table 4: AUC values on Zhihu. ($\alpha = 1.0, \beta = 0.3, \gamma = 0.3$)

Context-Aware Network Embedding

- Mutual Attention

Edge #1: (A, B)
Machine Learning research making great progress many directions This article summarizes four directions discusses current open problems The four directions improving classification accuracy learning ensembles classifiers methods scaling supervised learning algorithms reinforcement learning learning complex stochastic models

The problem making optimal decisions uncertain conditions central Artificial Intelligence If state world known times world modeled Markov Decision Process MDP MDPs studied extensively many methods known determining optimal courses action policies The realistic case state information partially observable Partially Observable Markov Decision Processes POMDPs received much less attention The best exact algorithms problems inefficient space time We introduce Smooth Partially Observable Value Approximation SPOVA new approximation method quickly yield good approximations improve time This method combined reinforcement learning methods combination effective test cases

Edge #2: (A, C)
Machine Learning research making great progress many directions This article summarizes four directions discusses current open problems The four directions improving classification accuracy learning ensembles classifiers methods scaling supervised learning algorithms reinforcement learning learning complex stochastic models

In context machine learning examples paper deals problem estimating quality attributes without dependencies among Kira Rendell developed algorithm called RELIEF shown efficient estimating attributes Original RELIEF deal discrete continuous attributes limited twoclass problems In paper RELIEF analysed extended deal noisy incomplete multiclass data sets The extensions verified various artificial one well known realworld problem

考虑标签的网络表示学习

Cunchao Tu, Weicheng Zhang, Zhiyuan Liu, Maosong Sun, Huanbo Luan. Max-Margin DeepWalk: Discriminative Learning of Network Representation. International Joint Conference on Artificial Intelligence (IJCAI 2016).

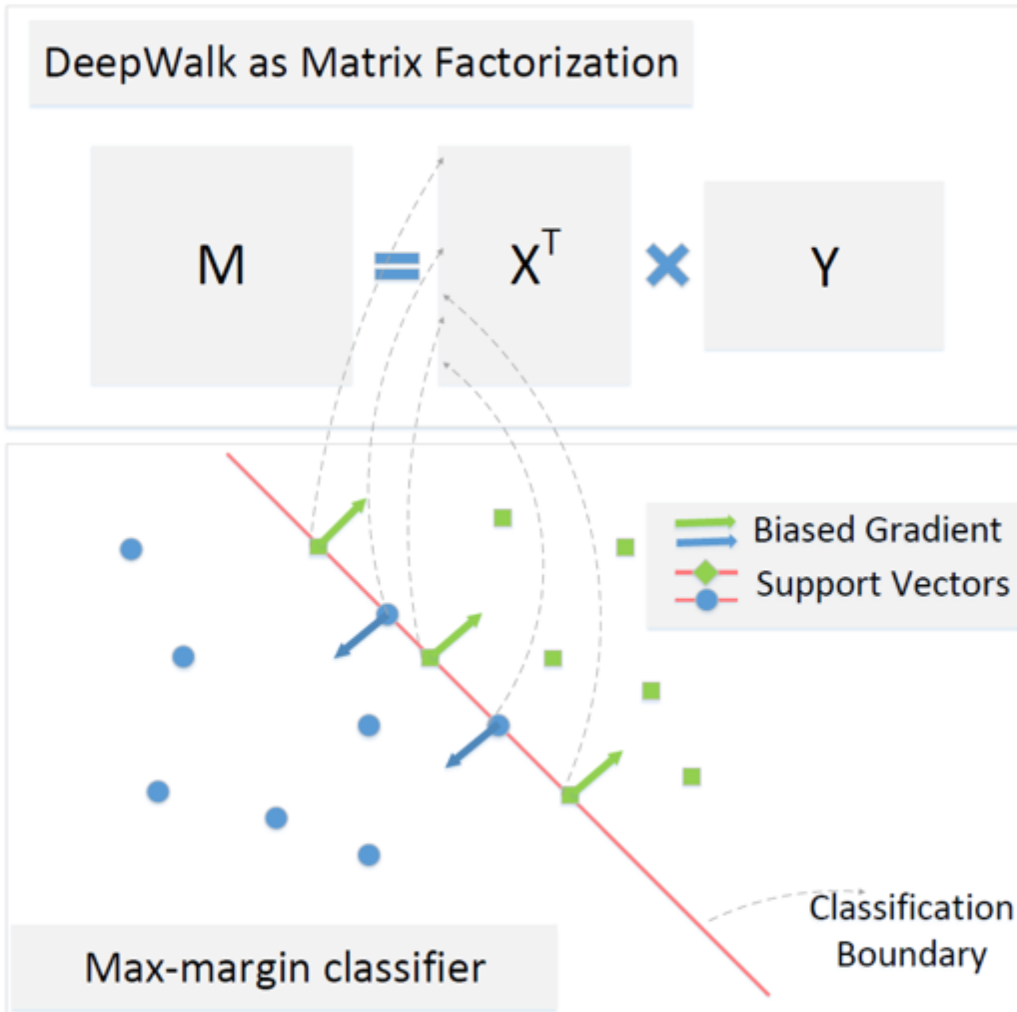
引入标签信息

- 真实世界网络节点往往被标注类别标签

传统NRL是无监督方法，无法考虑标签信息
在预测任务上效果差

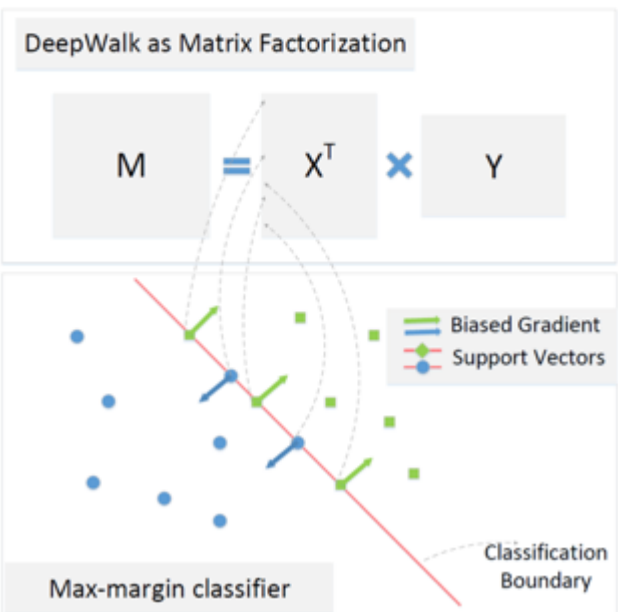
Max-Margin DeepWalk

- 共同训练DW+最大间隔分类器



Max-Margin DeepWalk

- Max-Margin DeepWalk (MMDW)
  - 利用MFDW初始化节点表示
  - 利用标注节点训练SVM
  - 对于标注节点计算其偏置向量
  - 重新训练MFDW



使边界支持向量向各自类别移动
让类别之间分类界限更加明显

Max-Margin DeepWalk

- 节点分类结果
  - >5%的提升
  - 仅用一半训练数据即可达到baseline的分类效果

Table 2: Accuracy (%) of vertex classification on Citeseer.

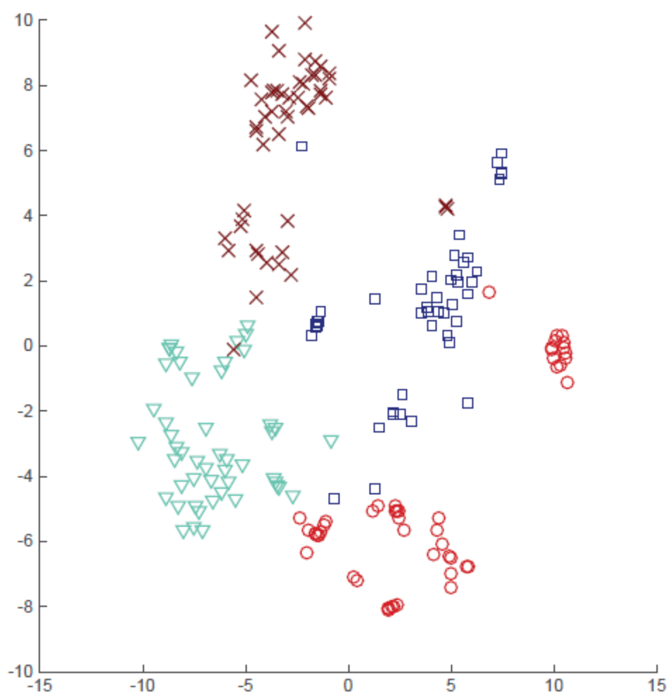
%Labeled Nodes	10%	20%	30%	40%	50%	60%	70%	80%	90%
DW	49.09	55.96	60.65	63.97	65.42	67.29	66.80	66.82	63.91
MFDW	50.54	54.47	57.02	57.19	58.60	59.18	59.17	59.03	55.35
LINE	39.82	46.83	49.02	50.65	53.77	54.2	53.87	54.67	53.82
MMDW($\eta = 10^{-2}$)	**55.60**	60.97	63.18	65.08	**66.93**	**69.52**	**70.47**	**70.87**	**70.95**
MMDW($\eta = 10^{-3}$)	55.56	**61.54**	**63.36**	**65.18**	66.45	69.37	68.84	70.25	69.73
MMDW($\eta = 10^{-4}$)	54.52	58.49	59.25	60.70	61.62	61.78	63.24	61.84	60.25

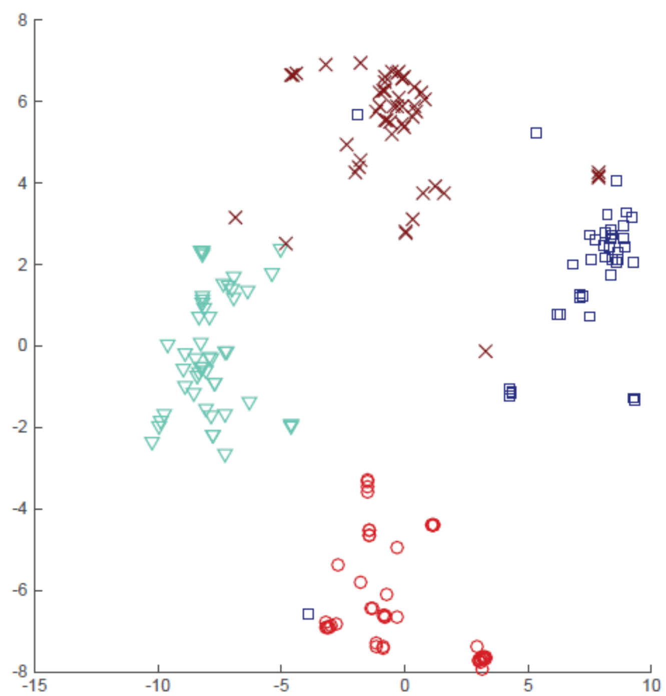
Table 3: Accuracy (%) of vertex classification on Wiki.

%Labeled Nodes	10%	20%	30%	40%	50%	60%	70%	80%	90%
DW	52.03	54.62	59.80	60.29	61.26	65.41	65.84	66.53	68.16
MFDW	56.40	60.28	61.90	63.39	62.59	62.87	64.45	62.71	61.63
LINE	52.17	53.62	57.81	57.26	58.94	62.46	62.24	66.74	67.35
MMDW($\eta = 10^{-2}$)	**57.76**	**62.34**	**65.76**	**67.31**	**67.33**	**68.97**	**70.12**	**72.82**	**74.29**
MMDW($\eta = 10^{-3}$)	54.31	58.69	61.24	62.63	63.18	63.58	65.28	64.83	64.08
MMDW($\eta = 10^{-4}$)	53.98	57.48	60.10	61.94	62.18	62.36	63.21	62.29	63.67

Max-Margin DeepWalk

- 节点表示可视化（t-SNE）
  - DeepWalk与MMDW



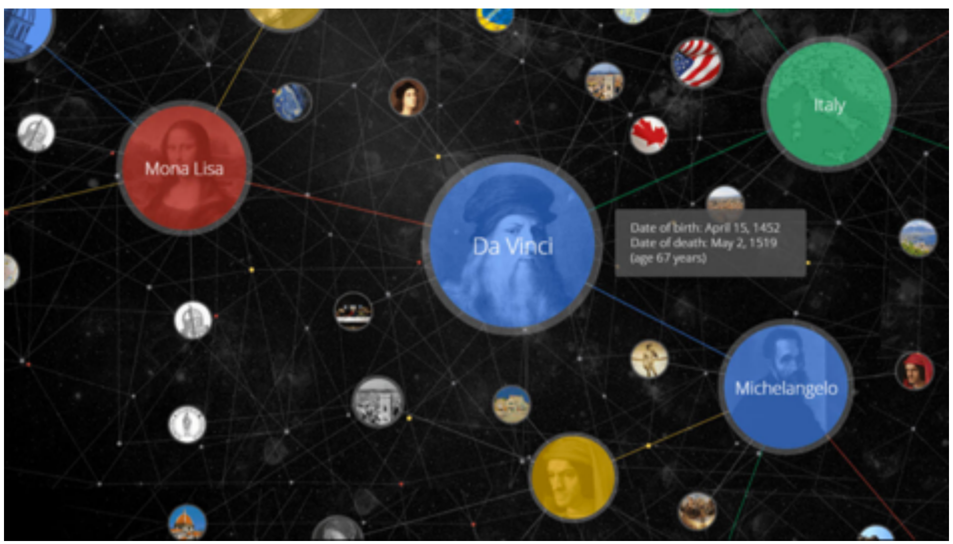


面向社会关系抽取的网络表示学习

Cunchao Tu, Zhengyan Zhang, Zhiyuan Liu, Maosong Sun. TransNet: Translation-Based Network Representation Learning for Social Relation Extraction. International Joint Conference on Artificial Intelligence (IJCAI 2017).

TransNet: Translation-based NRL

- 目前大部分社会网络分析技术忽略了用户间的关系标签，只考虑关系强弱
- 如何将边上具体的语义信息融合到NRL?
- 如何抽取节点之间具体的关系信息？

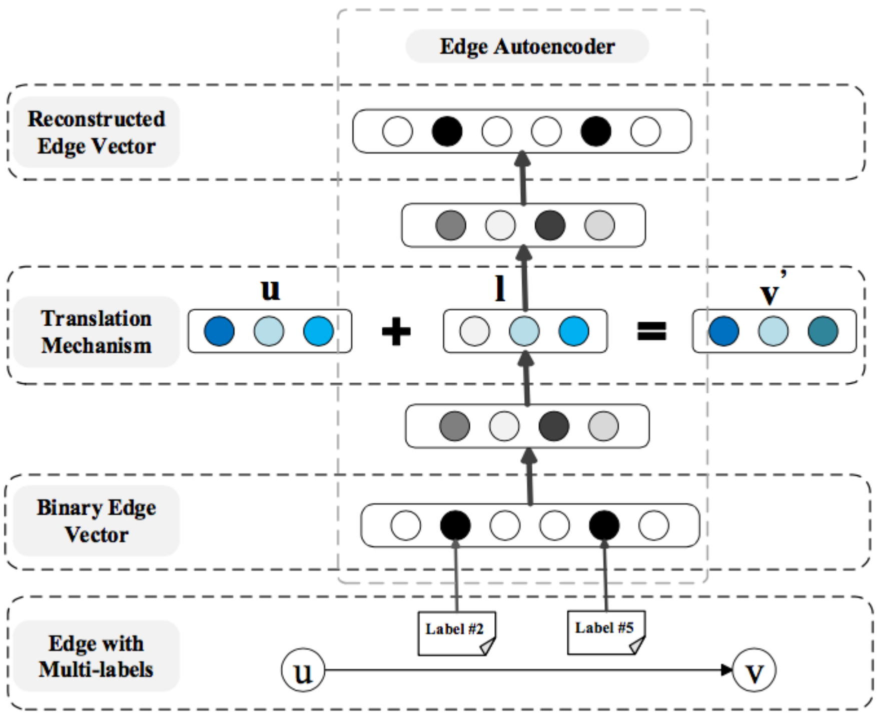


TransNet: Translation-based NRL

- TransNet
  - 利用知识图谱的概念/实体作为语义标签，对用户关系进行语义标注。
  - 基于平移思想自动学习语义标签和用户表示，实现社会关系抽取

TransNet: Translation-based NRL

- TransNet
  - 节点向量与关系向量之间的平移：u+l=v’
  - 通过Deep-Autoencoder对关系标签集合进行表示及预测



TransNet: Translation-based NRL

- 社会关系抽取结果

Table 2: SRE results on Arnet-S. ($\times 100$ for $hits@k$, $\alpha = 0.5$ and $\beta = 20$)

Metric	$hits@1$	$hits@5$	$hits@10$	MeanRank	$hits@1$	$hits@5$	$hits@10$	MeanRank
DeepWalk	13.88	36.80	50.57	19.69	18.78	39.62	52.55	18.76
LINE	11.30	31.70	44.51	23.49	15.33	33.96	46.04	22.54
node2vec	13.63	36.60	50.27	19.87	18.38	39.41	52.22	18.92
TransE	39.16	78.48	88.54	5.39	57.48	84.06	90.60	4.44
TransNet	**47.67**	**86.54**	**92.27**	**5.04**	**77.22**	**90.46**	**93.41**	**4.09**

Table 3: SRE results on Arnet-M. ($\times 100$ for $hits@k$, $\alpha = 0.5$ and $\beta = 50$)

Metric	$hits@1$	$hits@5$	$hits@10$	MeanRank	$hits@1$	$hits@5$	$hits@10$	MeanRank
DeepWalk	7.27	21.05	29.49	81.33	11.27	23.27	31.21	78.96
LINE	5.67	17.10	24.72	94.80	8.75	18.98	26.14	92.43
node2vec	7.29	21.12	29.63	80.80	11.34	23.44	31.29	78.43
TransE	19.14	49.16	62.45	25.52	31.55	55.87	66.83	23.15
TransNet	**27.90**	**66.30**	**76.37**	**25.18**	**58.99**	**74.64**	**79.84**	**22.81**

Table 4: SRE results on Arnet-L. ($\times 100$ for $hits@k$, $\alpha = 0.5$ and $\beta = 50$)

Metric	$hits@1$	$hits@5$	$hits@10$	MeanRank	$hits@1$	$hits@5$	$hits@10$	MeanRank
DeepWalk	5.41	16.17	23.33	102.83	7.59	17.71	24.58	100.82
LINE	4.28	13.44	19.85	114.95	6.00	14.60	20.86	112.93
node2vec	5.39	16.23	23.47	102.01	7.53	17.76	24.71	100.00
TransE	15.38	41.87	55.54	32.65	23.24	47.07	59.33	30.64
TransNet	**28.85**	**66.15**	**75.55**	**29.60**	**56.82**	**73.42**	**78.60**	**27.40**

TransNet: Translation-based NRL

- 示例
  - 对于“A. Swami”不同邻居的关系标签推荐结果

Table 6: Recommended top-3 labels for each neighbor.

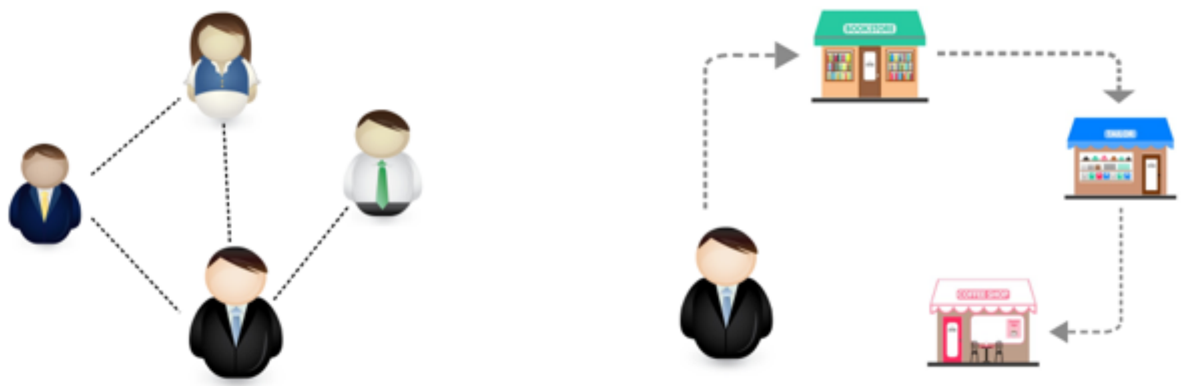
Neighbors	TransE	TransNet
Matthew Duggan	**ad hoc network**; wireless sensor network; wireless sensor networks	**management system**; **ad hoc network**; wireless sensor
K. Pelechrinis	**wireless network**; wireless networks; ad hoc network	**wireless network**; wireless sensor network; **routing protocol**
Oleg Korobkin	**wireless network**; wireless networks; **wireless communication**	**resource management**; **system design**; **wireless network**

网络表示学习的应用

Cheng Yang, Maosong Sun, Wayne Xin Zhao, Zhiyuan Liu, Edward Chang. A Neural Network Approach to Joint Modeling Social Networks and Mobile Trajectories. ACM Transactions on Information Systems (ACM TOIS), 2017.

社交网络和用户轨迹的联合神经模型

- 将社交网络和用户的移动轨迹联合建模



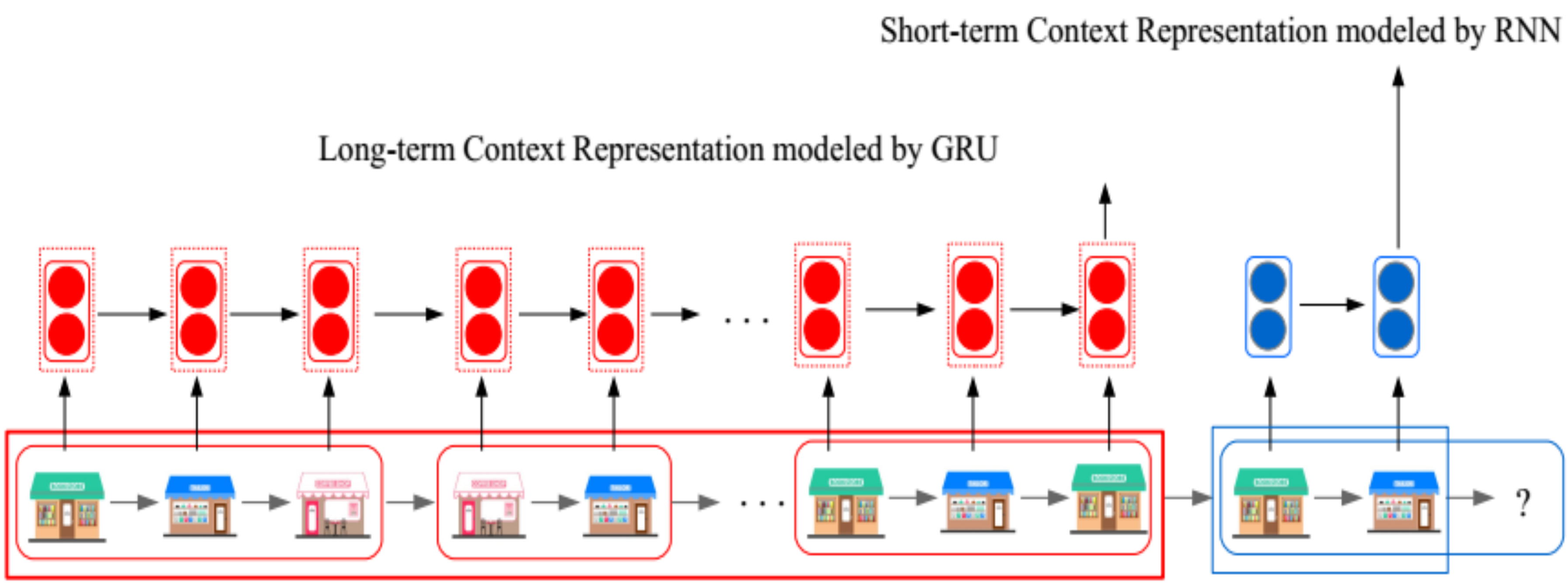
(a) Friendship Network

(b) User Trajectory

Fig. 1. An illustrative example for the data in LBSNs: (a) Link connections represent the friendship between users. (b) A trajectory generated by a user is a sequence of chronologically ordered check-in records.

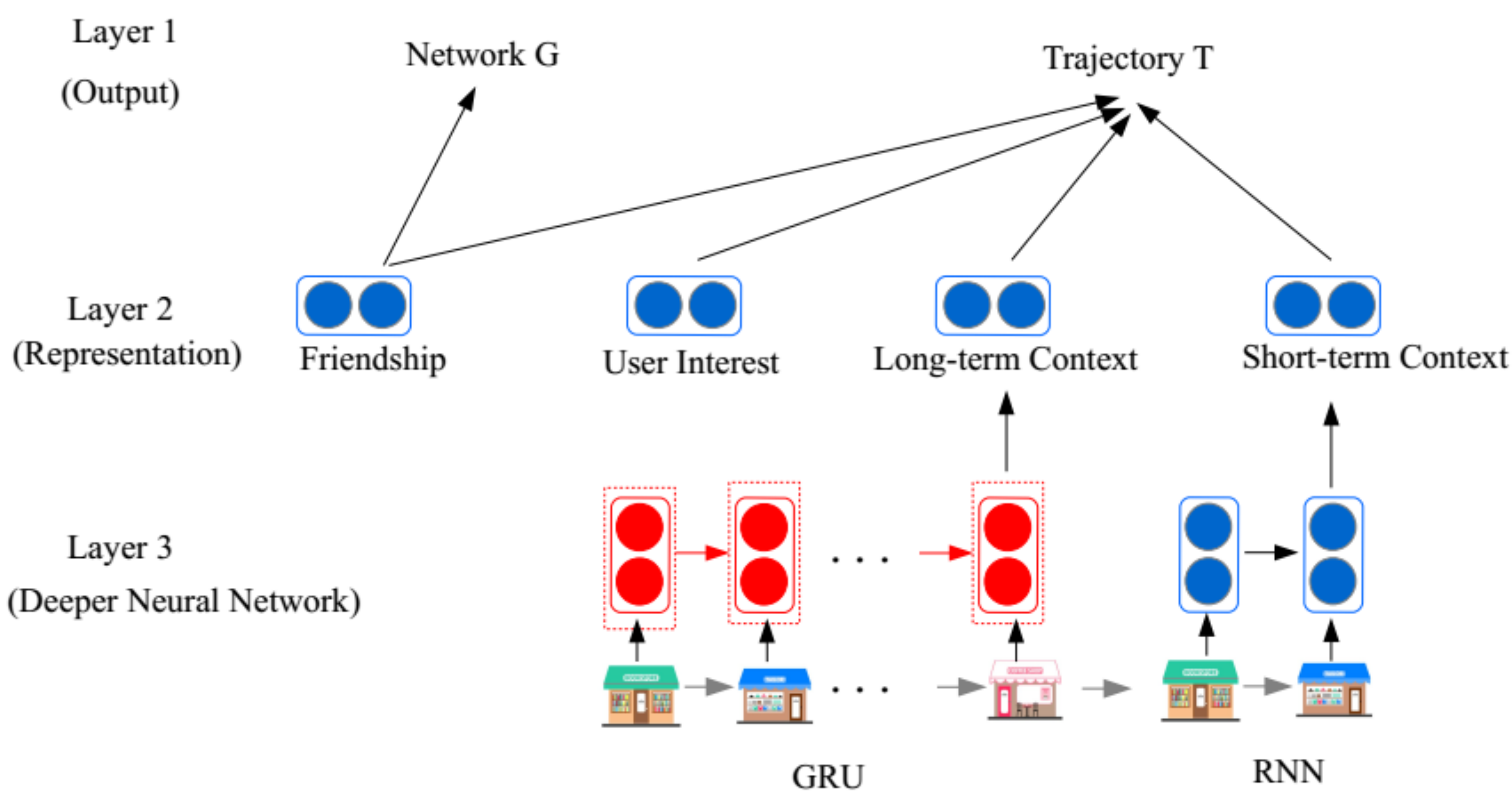
社交网络和用户轨迹的联合神经模型

- 使用循环神经网络对用户轨迹建模



社交网络和用户轨迹的联合神经模型

- 以用户表示为基础的整体神经网络模型



社交网络和用户轨迹的联合神经模型

- 下个地点预测任务实验结果

Dataset	Brightkite			Gowalla		
Metric (%)	R@1	R@5	R@10	R@1	R@5	R@10
PV	18.5	44.3	53.2	9.9	27.8	36.3
FBC	16.7	44.1	54.2	13.3	34.4	42.3
FPMC	20.6	45.6	53.8	10.1	24.9	31.6
PRME	15.4	44.6	53.0	12.2	31.9	38.2
HRM	17.4	46.2	56.4	7.4	26.2	37.0
JNTM	**22.1**	**51.1**	**60.3**	**15.4**	**38.8**	**48.1**

社交网络和用户轨迹的联合神经模型

- 下个新地点预测任务实验结果

Dataset	Brightkite			Gowalla		
Metric (%)	R@1	R@5	R@10	R@1	R@5	R@10
PV	0.5	1.5	2.3	1.0	3.3	5.3
FBC	0.5	1.9	3.0	1.0	3.1	5.1
FPMC	0.8	2.7	4.3	2.0	6.2	9.9
PRME	0.3	1.1	1.9	0.6	2.0	3.3
HRM	1.2	3.5	5.2	1.7	5.3	8.2
JNTM	**1.3**	**3.7**	**5.5**	**2.7**	**8.1**	**12.1**

社交网络和用户轨迹的联合神经模型

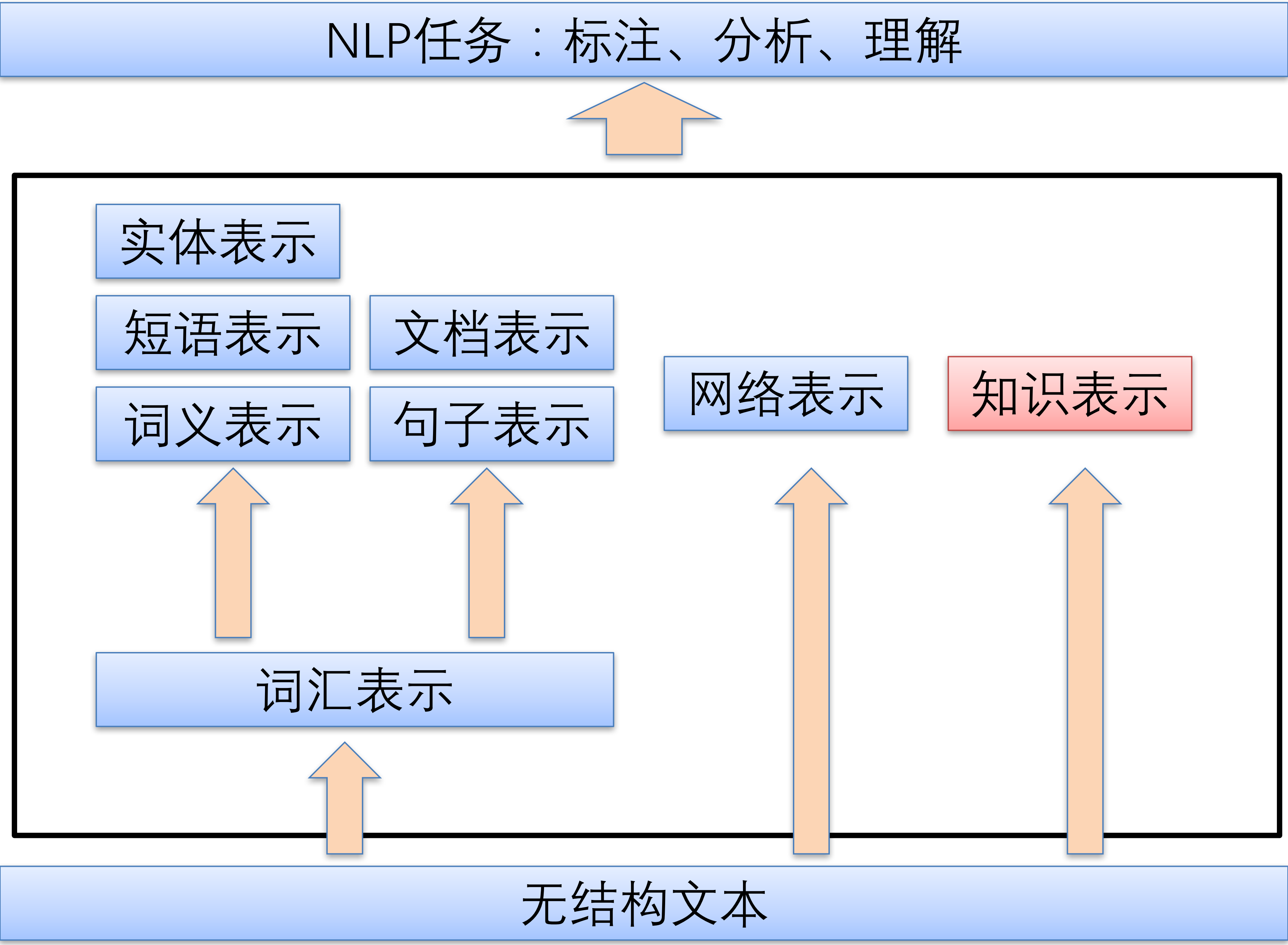
- 朋友推荐任务实验结果

Training Ratio	20%		30%		40%		50%	
Metric (%)	R@5	R@10	R@5	R@10	R@5	R@10	R@5	R@10
DeepWalk	2.3	3.8	3.9	6.7	5.5	9.2	7.4	12.3
PMF	2.1	3.6	2.1	3.7	2.3	3.4	2.3	3.8
PTE	1.5	2.5	3.8	4.7	4.0	6.6	5.1	8.3
TADW	2.2	3.4	3.6	3.9	2.9	4.3	3.2	4.5
JNTM	**3.7**	**6.0**	**5.4**	**8.7**	**6.7**	**11.1**	**8.4**	**13.9**

Training Ratio	20%		30%		40%		50%	
Metric (%)	R@5	R@10	R@5	R@10	R@5	R@10	R@5	R@10
DeepWalk	2.6	3.9	5.1	8.1	**7.9**	**12.1**	**10.5**	**15.8**
PMF	1.7	2.4	1.8	2.5	1.9	2.7	1.9	3.1
PTE	1.1	1.8	2.3	3.6	3.6	5.6	4.9	7.6
TADW	2.1	3.1	2.6	3.9	3.2	4.7	3.6	5.4
JNTM	**3.8**	**5.5**	**5.9**	**8.9**	**7.9**	11.9	10.0	15.1

网络表示的研究趋势

- 探索特殊社会网络表示学习
  - Bipartite Networks, Signed Networks, Hetereogeneous Networks, …
- 大规模网络表示学习
- 探索动态网络下的表示学习
- 改进社会计算典型任务
  - 链接预测，社区发现
  - 影响力分析，传播预测
  - 用户建模，个性推荐
- 知识驱动的社会计算
  - 为社会计算引入推理能力
  - 提高社会计算的可解释性


NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
词义表示
句子表示
网络表示
知识表示
词汇表示
无结构文本

结构知识

write

(***William Shakespeare***, book/author/works_written, ***Romeo and Juliet***)

head entity **relation** **tail entity**

大规模知识图谱

DBpedia

250概念
4M实例
6000属性
500M三元组
在线更新

yago select knowledge

350K概念
10M实例
100属性
120M三元组

WolframAlpha

NELL

OpenIE
(Reverb, OLLIE)

XLORE

850K概念
8M实例
70K属性

搜狗搜索

Bai du 百度

Freebase

15K概念
40M实例
4000属性
1B三元组
Google KB核心

Google KG

15K概念
600M实例
20B三元组

BabelNet 1.1.1
A very large multilingual ontology

50M义项
50+种语言
262M三元组

WordNet
7种欧洲语言
跨语言链接

EuroWordNet

维基百科富知识文本

Cyc

WordNet

WordNet

HowNet

知网

Knowledge Database

HowNet

keenage.com

Wikipedia

DBpedia

yago

select knowledge

freebase™

High-Quality Data Source

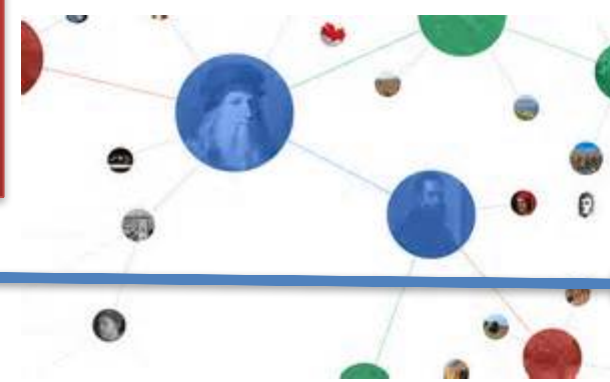
5 million concepts

muilti-lingual

rich structural data:

Infobox， table，

list， category…

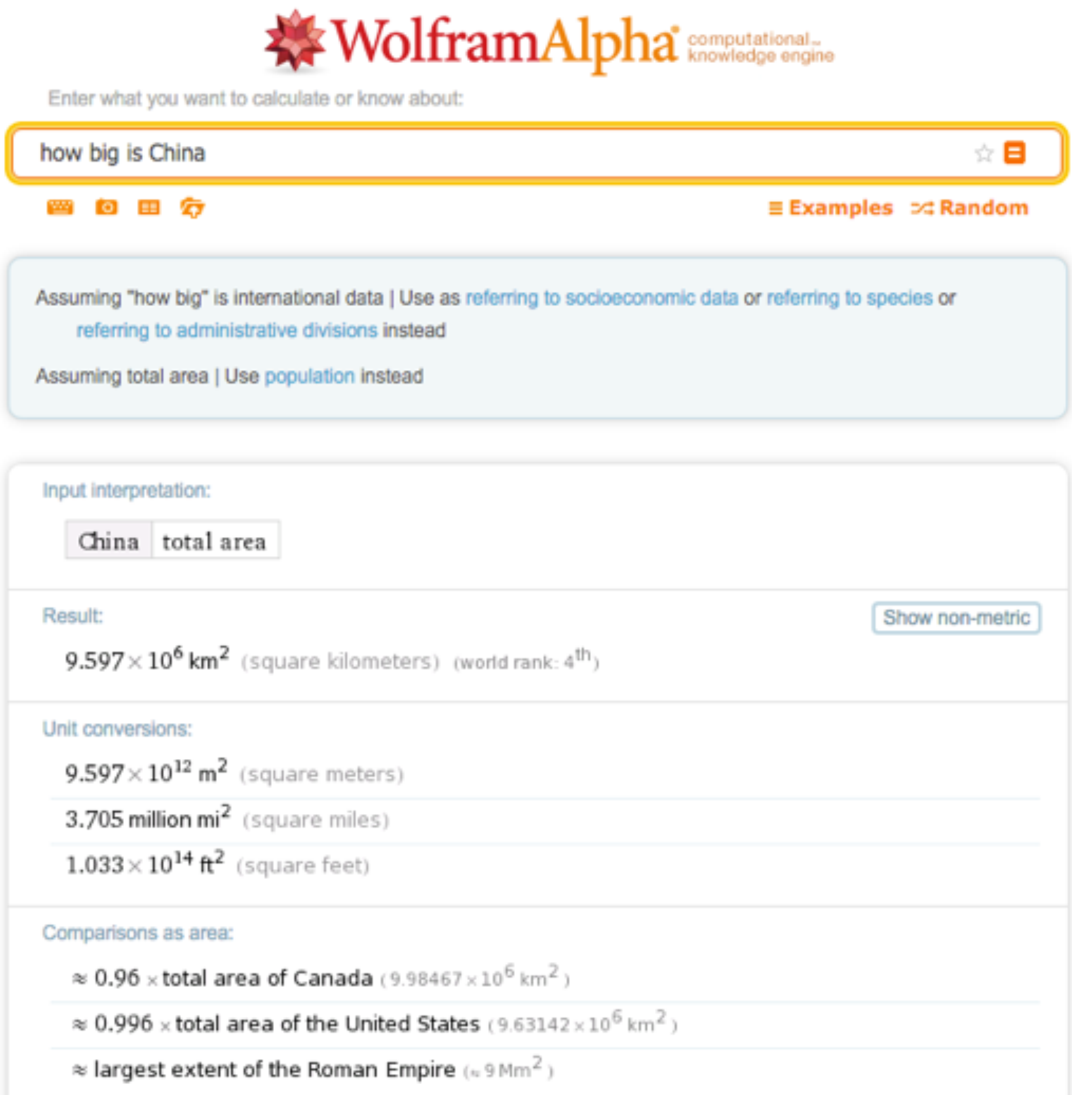


1985

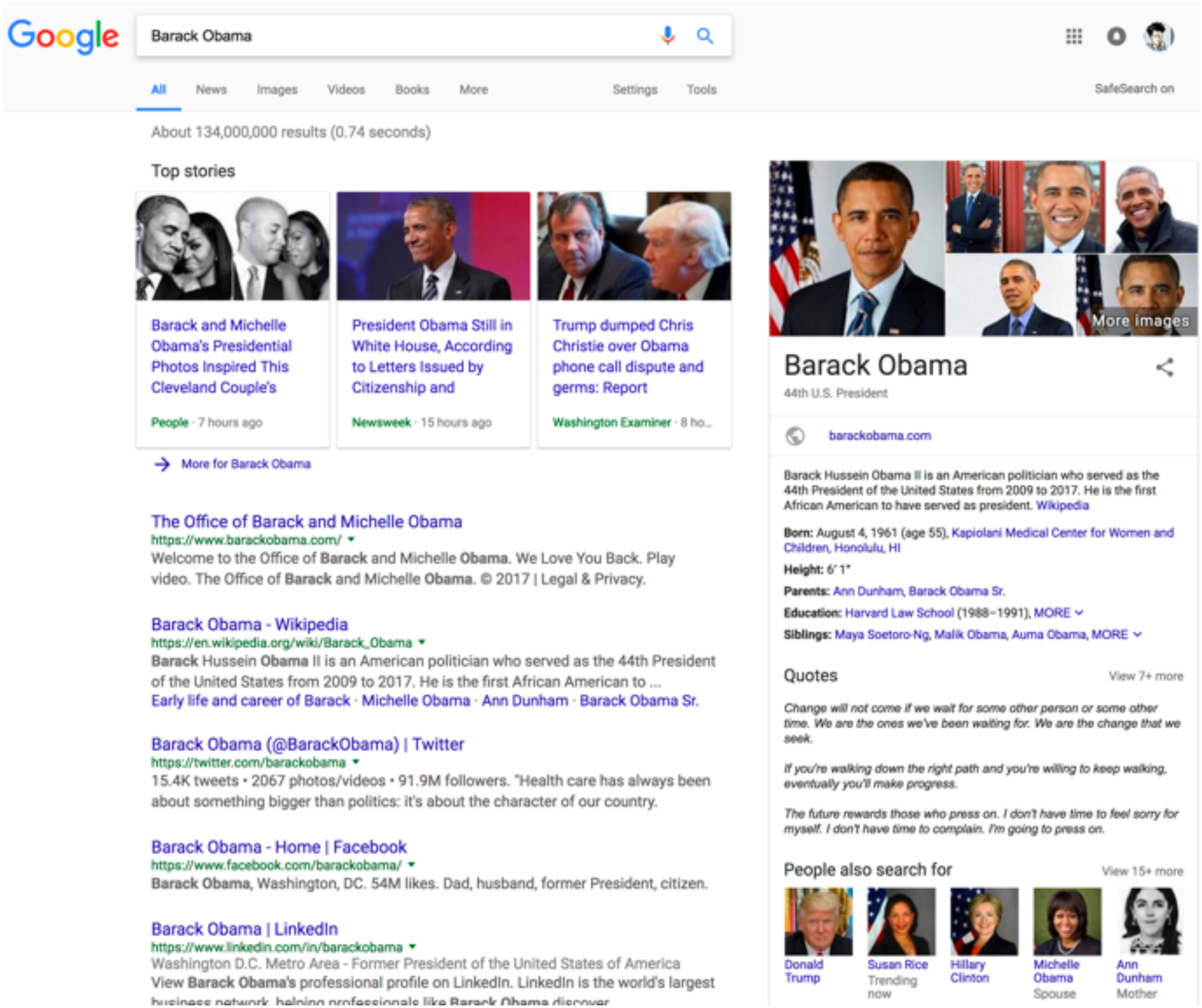
1990

2005-2010

知识图谱应用：智能问答


WolframAlpha computational knowledge engine
Enter what you want to calculate or know about:
how big is China
Examples
Random
Assuming "how big" is international data | Use as referring to socioeconomic data or referring to species or referring to administrative divisions instead
Assuming total area | Use population instead
Input interpretation:
China total area
Result:
Show non-metric
9.597 × 10^6 km^2 (square kilometers) (world rank: 4th)
Unit conversions:
9.597 × 10^12 m^2 (square meters)
3.705 million mi^2 (square miles)
1.033 × 10^14 ft^2 (square feet)
Comparisons as area:
≈ 0.96 × total area of Canada (9.98467 × 10^6 km^2)
≈ 0.996 × total area of the United States (9.63142 × 10^6 km^2)
≈ largest extent of the Roman Empire (≈ 9 Mm^2)

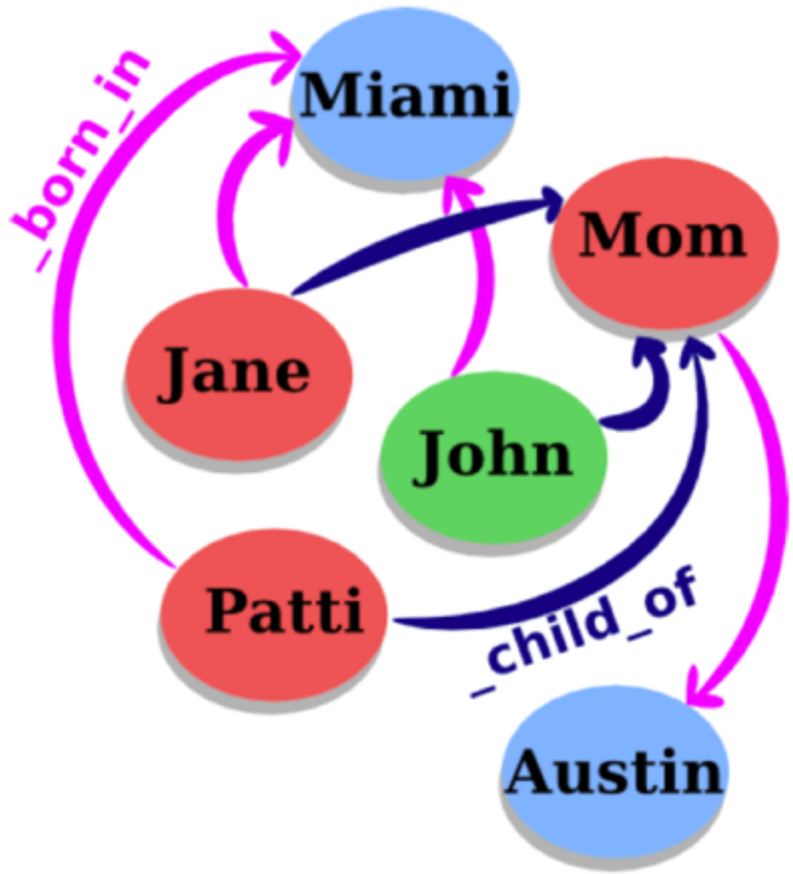
知识图谱应用：搜索引擎



Google
Barack Obama
All
News
Images
Videos
Books
More
Settings
Tools
SafeSearch on
About 134,000,000 results (0.74 seconds)
Top stories
Barack and Michelle Obama's Presidential Photos Inspired This Cleveland Couple's
People · 7 hours ago
President Obama Still in White House, According to Letters Issued by Citizenship and
Newsweek · 15 hours ago
Trump dumped Chris Christie over Obama phone call dispute and germs: Report
Washington Examiner · 8 ho...
More for Barack Obama
The Office of Barack and Michelle Obama
https://www.barackobama.com/
Welcome to the Office of Barack and Michelle Obama. We Love You Back. Play video. The Office of Barack and Michelle Obama. © 2017 | Legal & Privacy.
Barack Obama - Wikipedia
https://en.wikipedia.org/wiki/Barack_Obama
Barack Hussein Obama II is an American politician who served as the 44th President of the United States from 2009 to 2017. He is the first African American to ...
Early life and career of Barack · Michelle Obama · Ann Dunham · Barack Obama Sr.
Barack Obama (@BarackObama) | Twitter
https://twitter.com/barackobama
15.4K tweets • 2067 photos/videos • 91.9M followers. "Health care has always been about something bigger than politics: it's about the character of our country.
Barack Obama - Home | Facebook
https://www.facebook.com/barackobama/
Barack Obama, Washington, DC. 54M likes. Dad, husband, former President, citizen.
Barack Obama | LinkedIn
https://www.linkedin.com/in/barackobama
Washington D.C. Metro Area - Former President of the United States of America
View Barack Obama's professional profile on LinkedIn. LinkedIn is the world's largest
More images
Barack Obama
44th U.S. President
barackobama.com
Barack Hussein Obama II is an American politician who served as the 44th President of the United States from 2009 to 2017. He is the first African American to have served as president. Wikipedia
Born: August 4, 1961 (age 55), Kapiolani Medical Center for Women and Children, Honolulu, HI
Height: 6′ 1″
Parents: Ann Dunham, Barack Obama Sr.
Education: Harvard Law School (1988–1991), MORE
Siblings: Maya Soetoro-Ng, Malik Obama, Auma Obama, MORE
Quotes
View 7+ more
Change will not come if we wait for some other person or some other time. We are the ones we've been waiting for. We are the change that we seek.
If you're walking down the right path and you're willing to keep walking, eventually you'll make progress.
The future rewards those who press on. I don't have time to feel sorry for myself. I don't have time to complain. I'm going to press on.
People also search for
View 15+ more
Donald Trump
Susan Rice
Trending now
Hillary Clinton
Michelle Obama
Spouse
Ann Dunham
Mother

研究问题：知识图谱实体与关系的表示

- 知识图谱包括实体与关系
- 节点代表实体
- 连边代表关系
- 事实可以用三元组表示
  - (head, relation, tail):
- 代表知识库
  - WordNet: 语言知识
  - Freebase: 世界知识



知识图谱研究的挑战问题

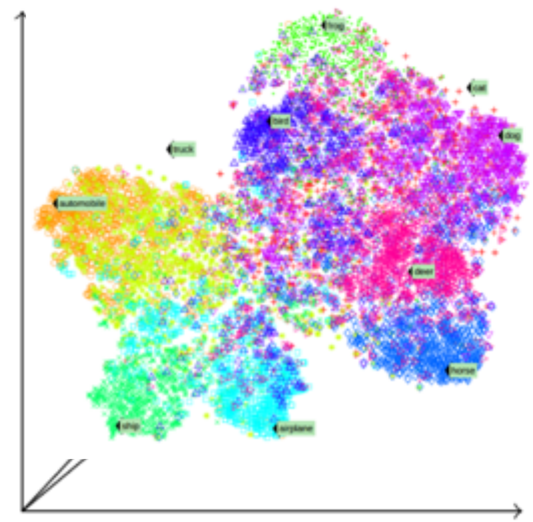
- 知识图谱非常稀疏，亟需关系抽取补充知识

- 从文本数据抽取关系：信息抽取任务
- 从知识图谱抽取关系：知识图谱补全

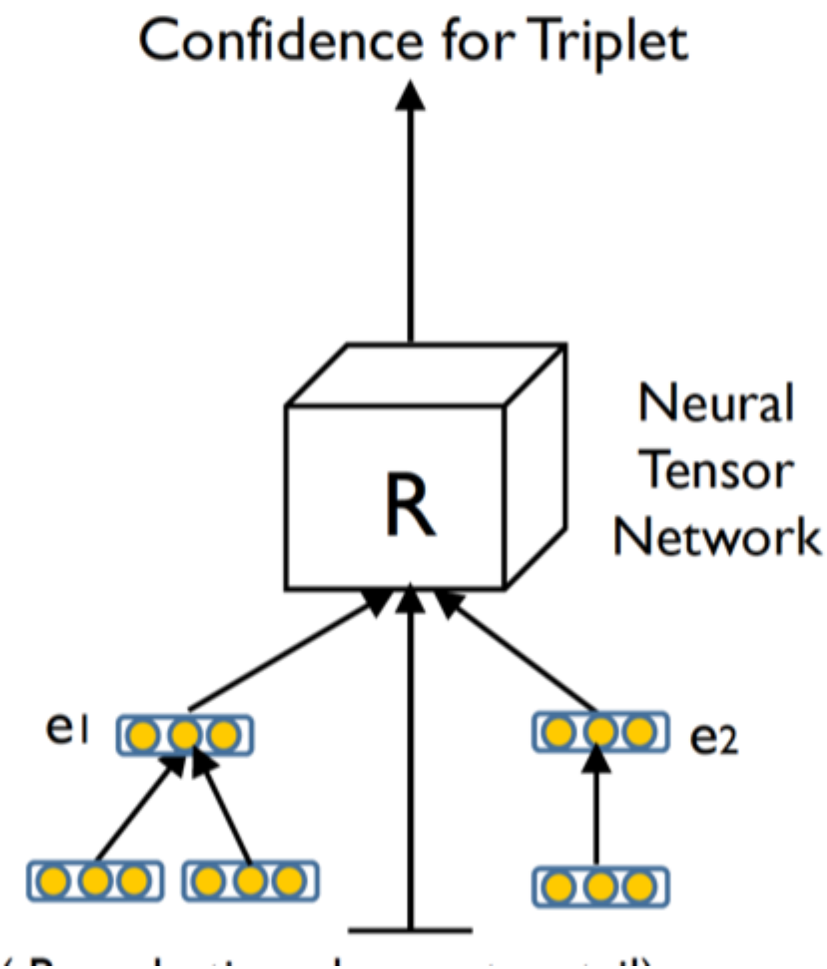
- 研究挑战：如何表示与利用知识图谱中的信息
  - 高维: 10^5~10^8个实体, 10^7~10^9种关系
  - 稀疏
  - 高噪音、不完整
- 研究思路：将知识图谱嵌入到低维向量空间

知识表示学习

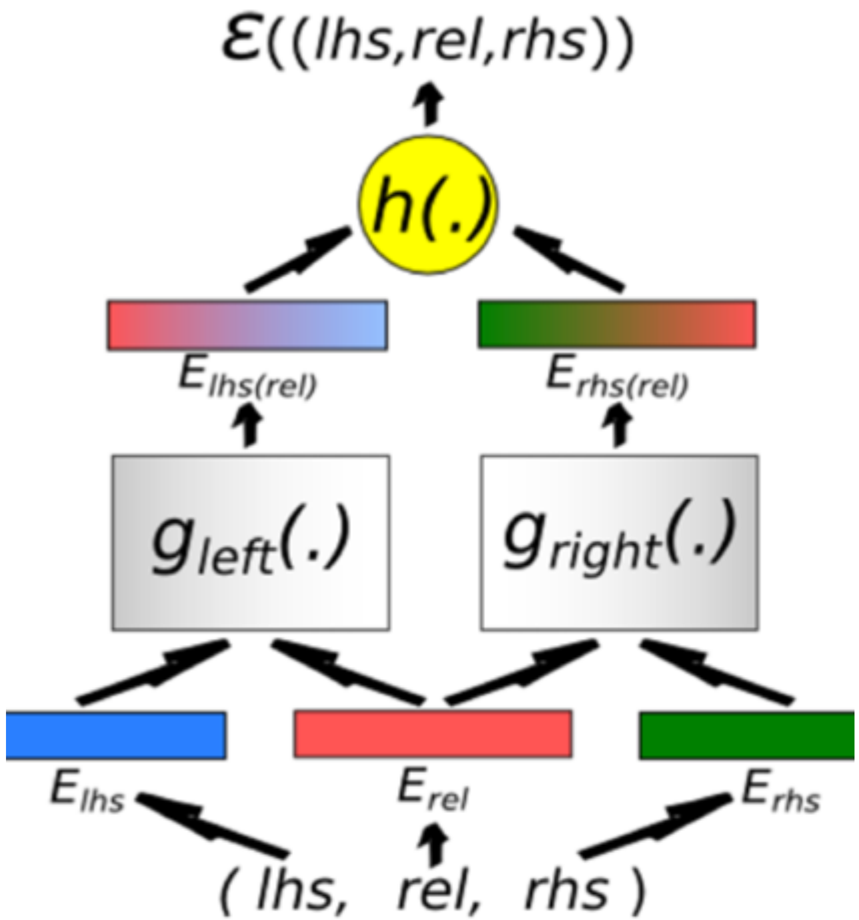
- 将知识映射到低维向量空间



知识表示代表模型



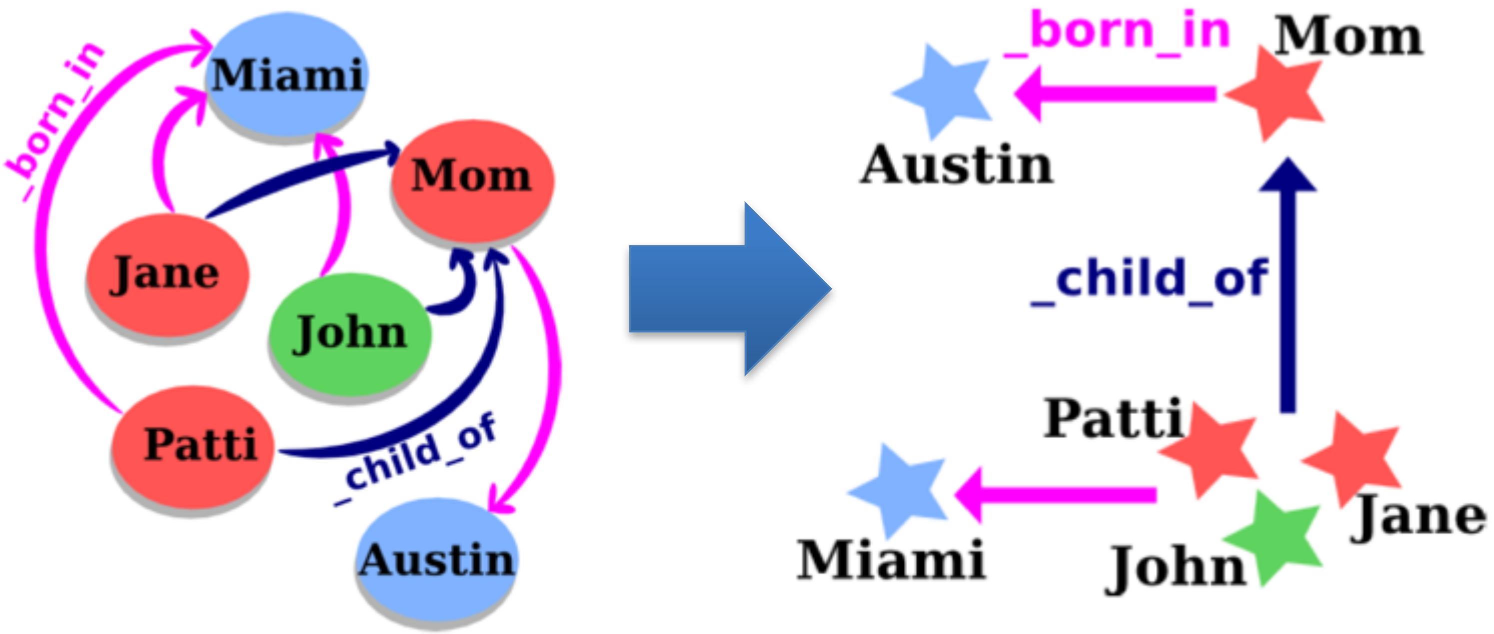
Neural Tensor Network (NTN)



Energy Model

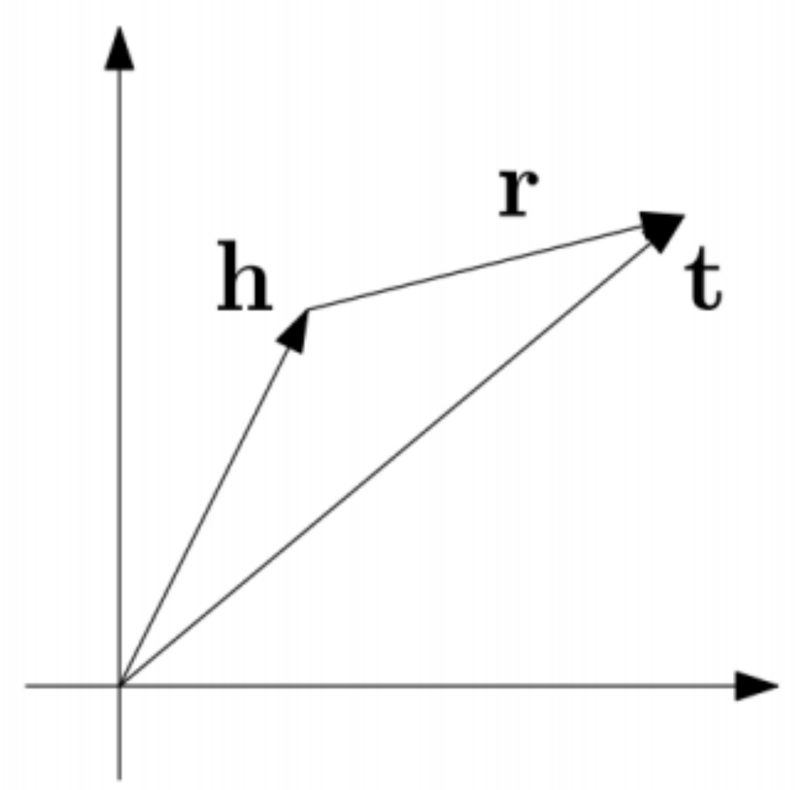
TransE: 将关系表示为平移

- 对每个事实 (head, relation, tail)， 将其中的relation作为从head到tail的平移操作



TransE: 将关系表示为平移

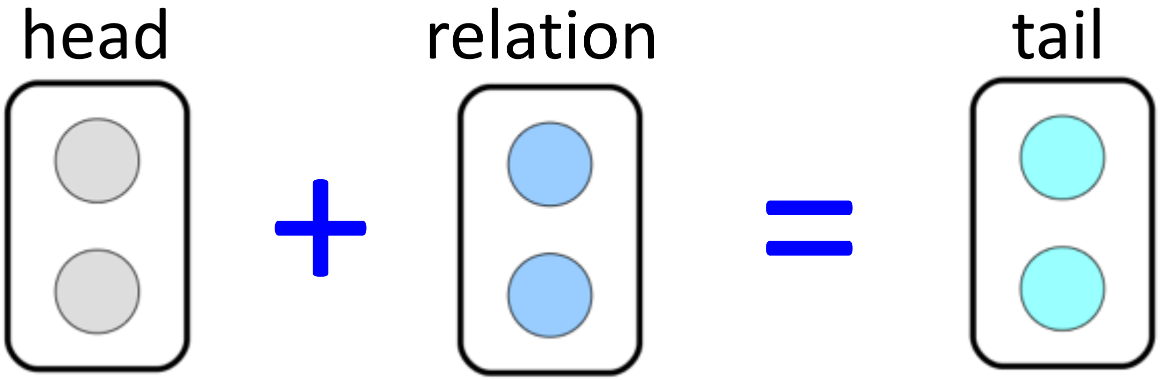
- 对每个事实 (head, relation, tail)，将relation作为从head到tail的平移操作



优化目标: **h + r = t**

TransE: 将关系表示为平移

- 对每个事实(head, relation, tail)，保证h + r = t



评价任务：链接预测

Which genre is the movie WALL-E?

WALL-E _has_genre ?

评价任务：链接预测

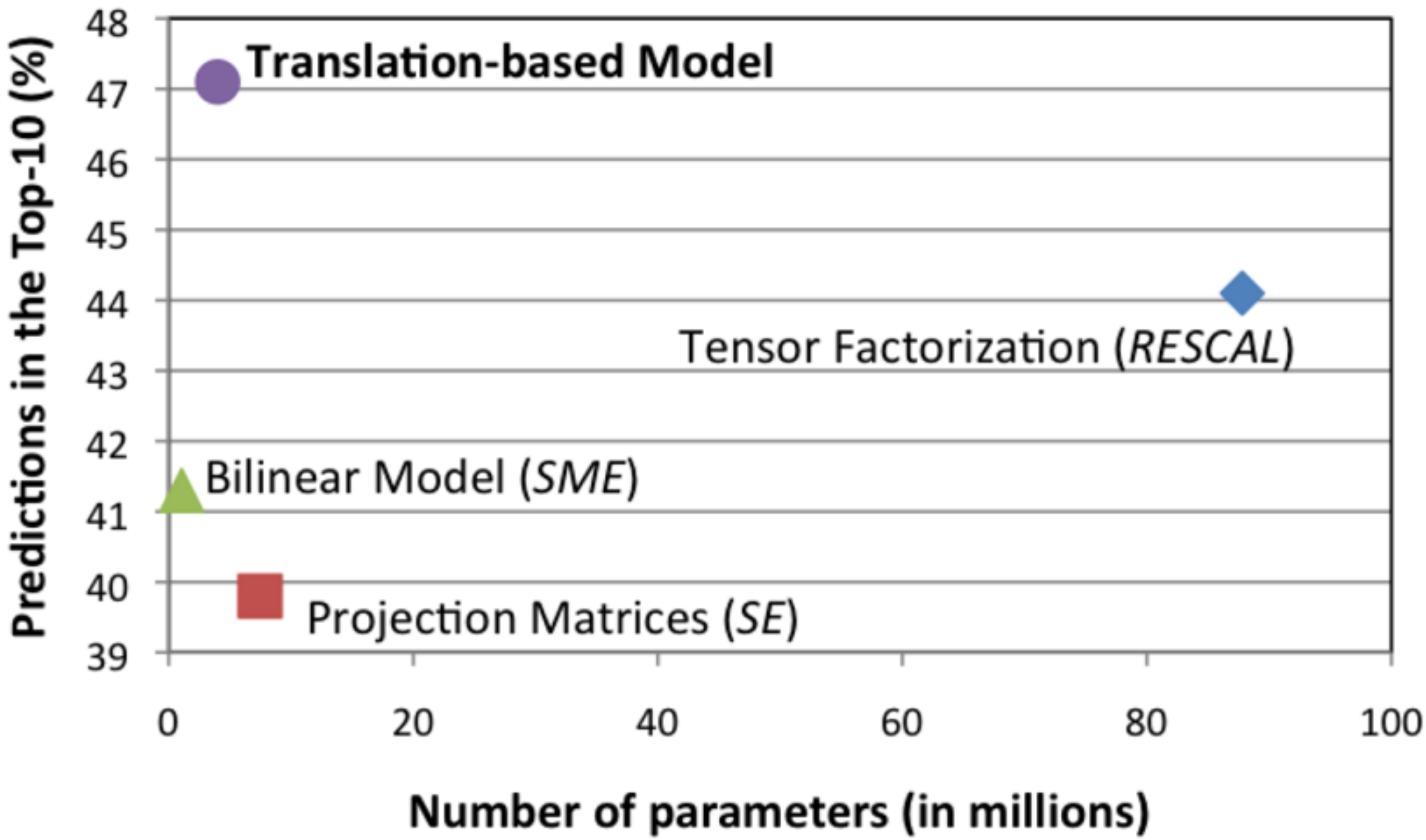
Which genre is the movie WALL-E?

WALL-E _has_genre

Animation
Computer animation
Comedy film
Adventure film
Science Fiction
Fantasy
Stop motion
Satire
Drama
Connecting

链接预测性能比较

Freebase15K



TransE Examples

Entity	**Tsinghua_University**	**A.C._Milan**
1	University_of_Victoria	Inter_Milan
2	St._Stephen's_College,_Delhi	Celtic_F.C.
3	University_of_Ottawa	FC_Barcelona
4	University_of_British_Columbia	Genoa_C.F.C.
5	Peking_University	Udinese_Calcio
6	Utrecht_University	Real_Madrid_C.F.
7	Dalhousie_University	FC_Bayern_Munich
8	Brasenose_College,_Oxford	Bolton_Wanderers_F.C.
9	Cardiff_University	Borussia_Dortmund
10	Memorial_University_of_Newfoundland	Hertha_BSC_Berlin

TransE Examples

Entity	China	Barack_Obama	Apple
1	Japan	George_W._Bush	Onion
2	Taiwan	Nancy_Pelosi	Strawberries
3	South_Korea	John_Kerry	Avocado
4	Argentina	Hillary_Rodham_Clinton	Pear
5	North_Korea	Al_Gore	Cabbage
6	Hungary	George_H._W._Bush	Broccoli
7	Israel	John_McCain	Egg
8	Australia	Colin_Powell	Cheese
9	Iceland	Bill_Clinton	Bread
10	Hong_Kong	Charles_B._Rangel	Tomato

TransE Examples

Relation	/people/person/nationality	/location/location/contains
1	/people/person/places_lived	/base/aareas/schema/administrative_area/administrative_children
2	/people/person/place_of_birth	/location/country/administrative_divisions
3	/people/person/spouse_s	/location/country/first_level_divisions
4	/base/popstra/celebrity/vacations_in	/location/country/capital
5	/government/politician/government_positions_held	/award/award_nominee/award_nominations
6	/people/deceased_person/place_of_death	/location/administrative_division/capital
7	/olympics/olympic_athlete/country	/location/us_county/county_seat
8	/olympics/olympic_athlete/medals_won	/base/aareas/schema/administrative_area/capital
9	/music/artist/origin	/location/us_county/hud_county_place
10	/people/person/employment_history	/award/award_winner/awards_won

TransE Examples

Head	China	Barack_Obama
Relation	**/location/location/adjoin**	**/education/education/institution**
1	Japan	Harvard_College
2	Taiwan	Massachusetts_Institute_of_Technology
3	Israel	American_University
4	South_Korea	University_of_Michigan
5	Argentina	Columbia_University
6	France	Princeton_University
7	Philippines	Emory_University
8	Hungary	Vanderbilt_University
9	North_Korea	University_of_Notre_Dame
10	Hong_Kong	Texas_A&M_University

TransE Examples

Head	Stanford_University	Apple	Titanic
Relation	/education/educational_institution/students_graduates	/food/food/nutrients	/film/film/genre
1	Steven_Spielberg	Lipid	War_film
2	Ron_Howard	Protein	Period_piece
3	Stan_Lee	Valine	Drama
4	Barack_Obama	Tyrosine	History
5	Milton_Friedman	Serine	Biography
6	Walter_F._Parkes	Iron	Film_adaptation
7	Michael_Cimino	Cystine	Adventure_Film
8	Gale_Anne_Hurd	Pantothenic_acid	Action_Film
9	Bryan_Singer	Vitamin_A	Political_drama
10	Aaron_Sorkin	Sugar	Costume_drama

TransE Examples

Head	Barack_Obama
Tail	**Columbia_University**
1	/education/education/institution
2	/business/employment_tenure/company
3	/organization/leadership/organization
4	/base/popstra/paid_support/company
5	/location/location/contains
6	/education/education/institution
7	/american_football/player_game_statistics/team
8	/organization/organization_board_membership/organization
9	/music/artist/album
10	/baseball/batting_statistics/team

TransE Examples

Head	Barack_Obama
Tail	**United_States_of_America**
1	/people/person/nationality
2	/government/politician/government_positions_held./government/government_position_held/jurisdiction_of_office
3	/people/person/spouse_s./people/marriage/location_of_ceremony
4	/government/politician/government_positions_held./government/government_position_held/district_represented
5	/people/person/places_lived
6	/base/popstra/celebrity/vacations_in
7	/people/person/place_of_birth
8	/government/political_appointer/appointees./government/government_position_held/jurisdiction_of_office
9	/royalty/monarch/kingdom
10	/people/person/languages

TransE Examples

Head	Titanic
Tail	Drama
1	/film/film/genre
2	/media_common/netflix_title/netflix_genres
3	/film/film/subjects
4	/film/film_cut/type_of_film_cut
5	/film/film_regional_release_date/film_release_distribution_medium
6	/law/inventor/inventions
7	/government/government_position_held/jurisdiction_of_office
8	/tv/tv_program/genre
9	/film/dubbing_performance/language
10	/film/film_film_distributor_relationship/film_distribution_medium

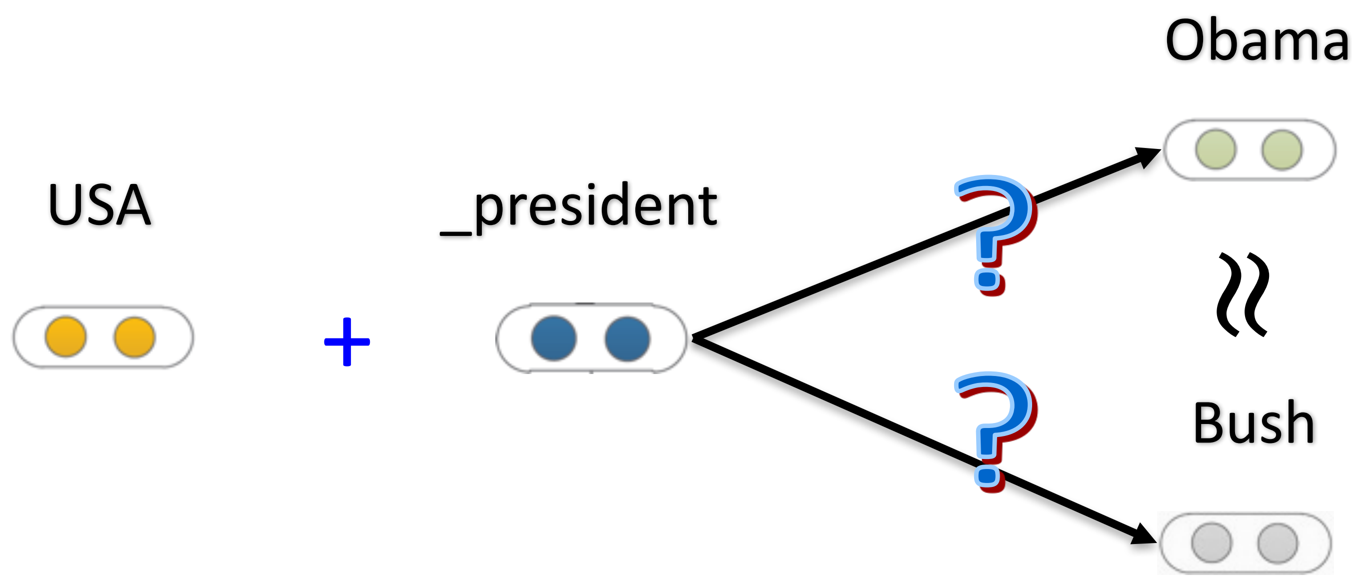
RL4KG面临的三个挑战

- 复杂关系的建模
- 文本与KG的融合
- 对关系路径建模

面向复杂关系建模

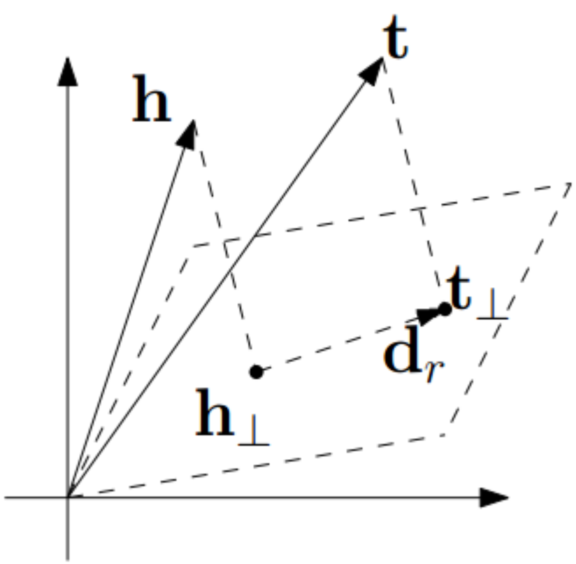
TransE面临挑战

- 1-to-N, N-to-1, N-to-N关系
  - (USA, _president, Obama)
  - (USA, _president, Bush)

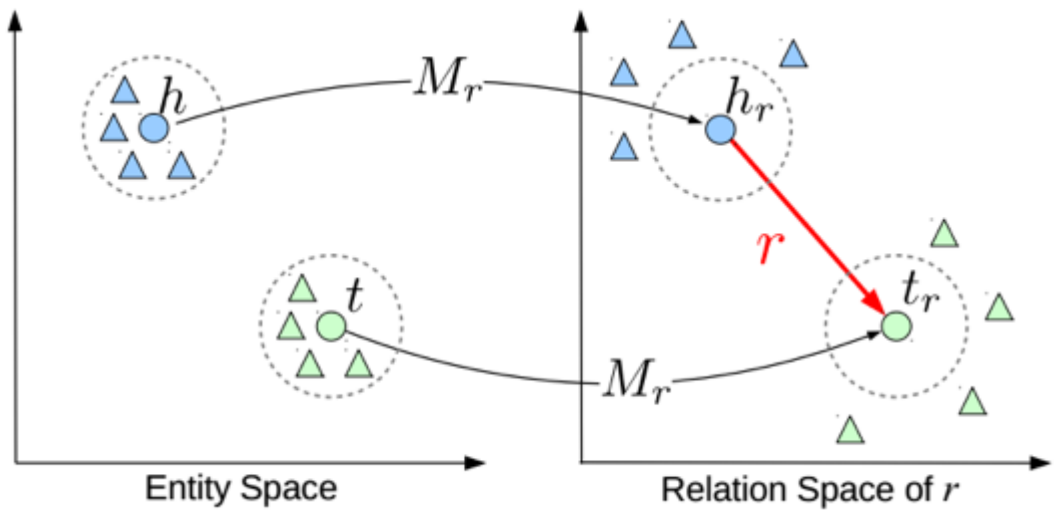


对复杂关系建模

- 建立与特定关系有关的实体表示



TransH

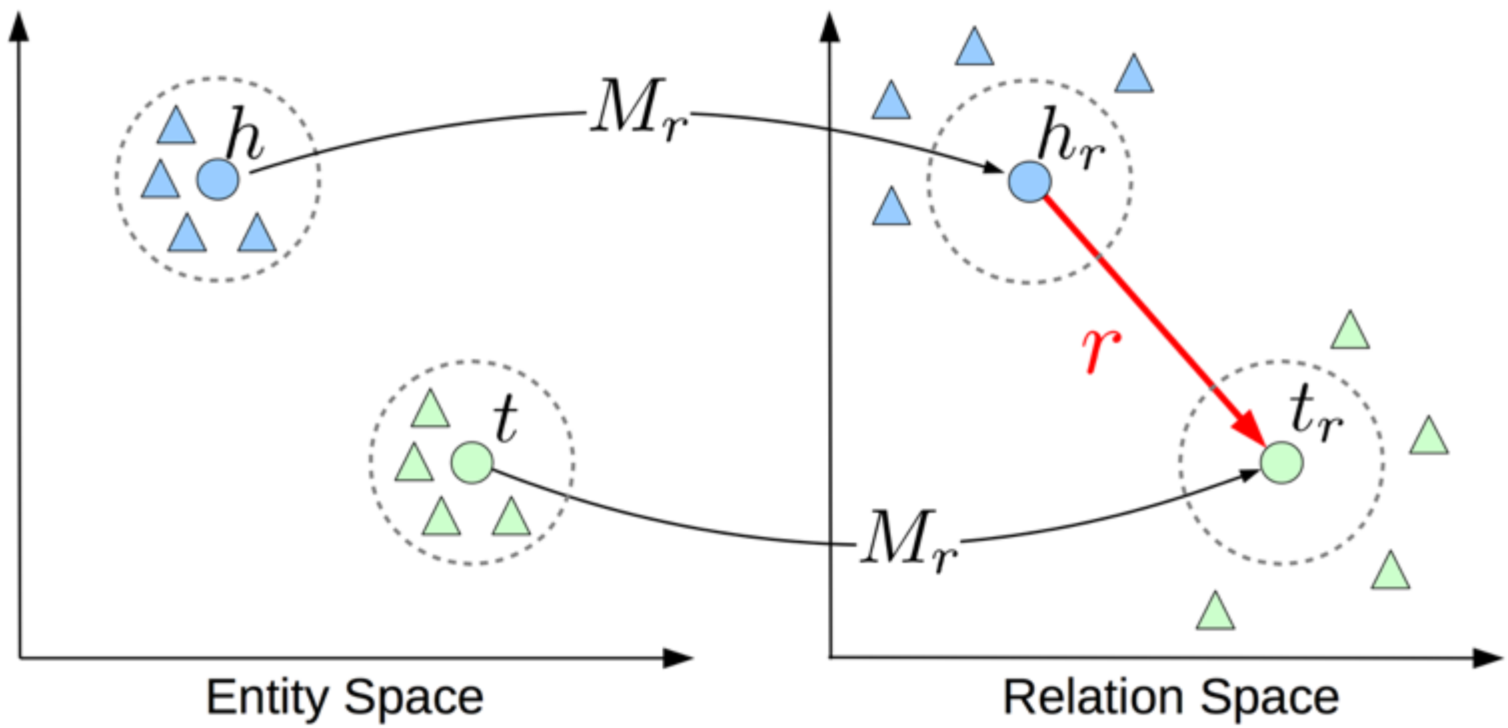


TransR

Wang, et al. (2014). Knowledge graph embedding by translating on hyperplanes. AAAI.
Lin, et al. (2015). Learning entity and relation embeddings for knowledge graph completion. AAAI.

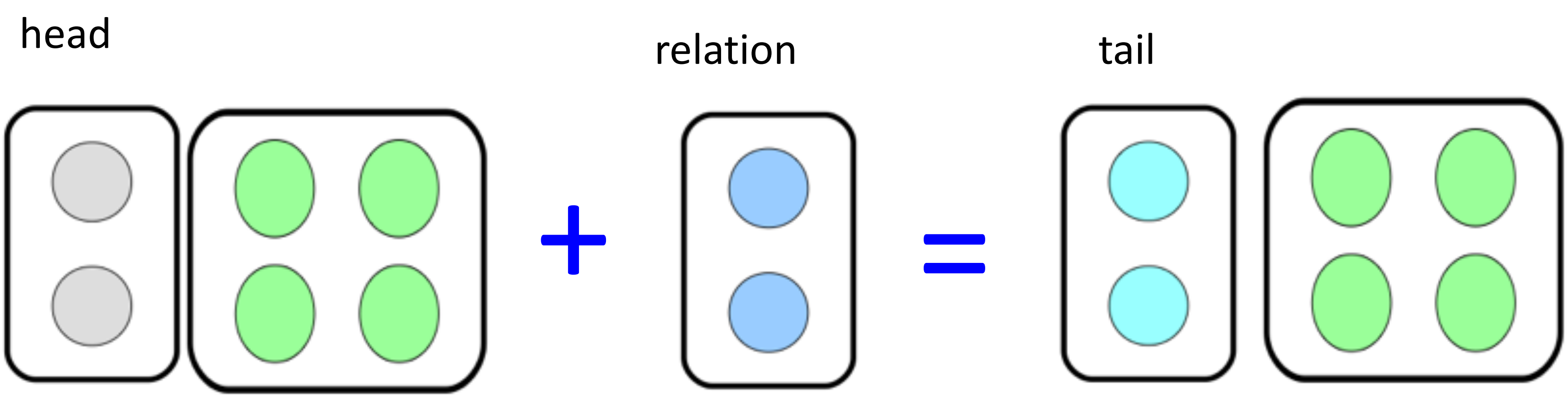
TransR：将实体与关系在不同空间中建模

- 对每个事实(head, relation, tail) 保证
  - h x W_r + r = t x W_r
- 将实体映射到不同关系的空间，建立平移关系

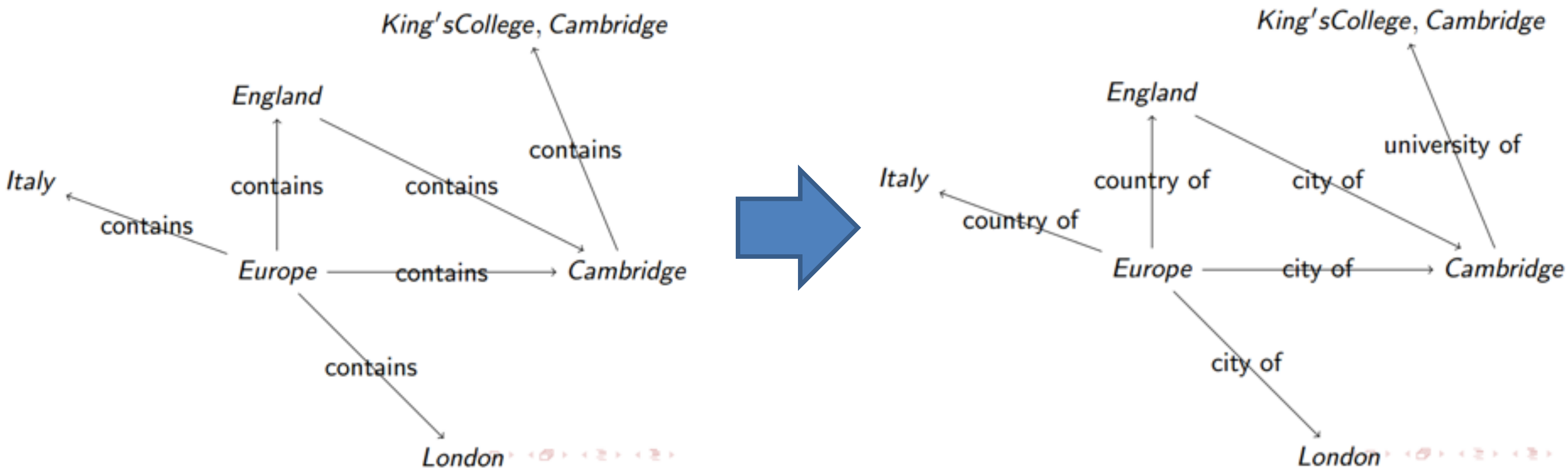


TransR：将实体与关系在不同空间中建模

- 对每个事实(head, relation, tail) 保证
  - h x W_r + r = t x W_r
- 将实体映射到不同关系的空间，建立平移关系



基于聚类的关系划分


King′sCollege, Cambridge
England
contains
Italy
contains
contains
contains
Europe
contains
Cambridge
contains
London
King′sCollege, Cambridge
England
university of
Italy
country of
city of
country of
Europe
city of
Cambridge
city of
London

基于聚类的关系划分

- location/location/contains 的不同类型

	⟨Head, Tail⟩
1	⟨Africa, Congo⟩, ⟨Asia, Nepal⟩, ⟨Americas, Aruba⟩, ⟨Oceania, Federated States of Micronesia⟩
2	⟨United States of America, Kankakee⟩, ⟨England, Bury St Edmunds⟩, ⟨England, Darlington⟩, ⟨Italy, Perugia⟩
3	⟨Georgia, Chatham County⟩, ⟨Idaho, Boise⟩, ⟨Iowa, Polk County⟩, ⟨Missouri, Jackson County⟩, ⟨Nebraska, Cass County⟩
4	⟨Sweden, Lund University⟩, ⟨England, King's College, Cambridge⟩, ⟨Fresno, California State University, Fresno⟩, ⟨Italy, Milan Conservatory⟩

实验结果：实体预测

Data Sets	WN18				FB15K			
Metric	Mean Rank		Hits@10 (%)		Mean Rank		Hits@10 (%)	
	Raw	Filter	Raw	Filter	Raw	Filter	Raw	Filter
Unstructured (Bordes et al. 2012)	315	304	35.3	38.2	1,074	979	4.5	6.3
RESCAL (Nickel, Tresp, and Kriegel 2011)	1,180	1,163	37.2	52.8	828	683	28.4	44.1
SE (Bordes et al. 2011)	1,011	985	68.5	80.5	273	162	28.8	39.8
SME (linear) (Bordes et al. 2012)	545	533	65.1	74.1	274	154	30.7	40.8
SME (bilinear) (Bordes et al. 2012)	526	509	54.7	61.3	284	158	31.3	41.3
LFM (Jenatton et al. 2012)	469	456	71.4	81.6	283	164	26.0	33.1
TransE (Bordes et al. 2013)	263	251	75.4	89.2	243	125	34.9	47.1
TransH (unif) (Wang et al. 2014)	318	303	75.4	86.7	211	84	42.5	58.5
TransH (bern) (Wang et al. 2014)	401	388	73.0	82.3	212	87	45.7	64.4
TransR (unif)	232	219	78.3	91.7	226	78	43.8	65.5
TransR (bern)	238	225	**79.8**	92.0	**198**	77	48.2	68.7
CTransR (unif)	243	230	78.9	**92.3**	233	82	44	66.3
CTransR (bern)	**231**	**218**	79.4	**92.3**	199	**75**	**48.4**	**70.2**

+20%

实验结果：关系预测

knowledge_completion.pdf

Data Sets	WN11	FB13	FB15K
SE	53.0	75.2	-
SME (bilinear)	70.0	63.7	-
SLM	69.9	85.3	-
LFM	73.8	84.3	-
NTN	70.4	**87.1**	68.5
TransE (unif)	75.9	70.9	79.6
TransE (bern)	75.9	81.5	79.2
TransH (unif)	77.7	76.5	79.0
TransH (bern)	78.8	83.3	80.2
TransR (unif)	85.5	74.7	81.7
TransR (bern)	**85.9**	82.5	83.9
CTransR (bern)	85.7	-	**84.5**

Examples

Head Entity	Titanic		
Relation	/film/film/genre		
Model	TransE	TransH	TransR
1	War_film	Drama	Costume_drama
2	Period_piece	Romance_Film	Drama
3	Drama	Costume_drama	Romance_Film
4	History	Film_adaptation	Period_piece
5	Biography	Period_piece	Epic_film
6	Film_adaptation	Adventure_Film	Adventure_Film
7	Adventure_Film	LGBT	LGBT
8	Action_Film	Existentialism	Film_adaptation
9	Political_drama	Epic_film	Existentialism
10	Costume_drama	War_film	War_film

Examples

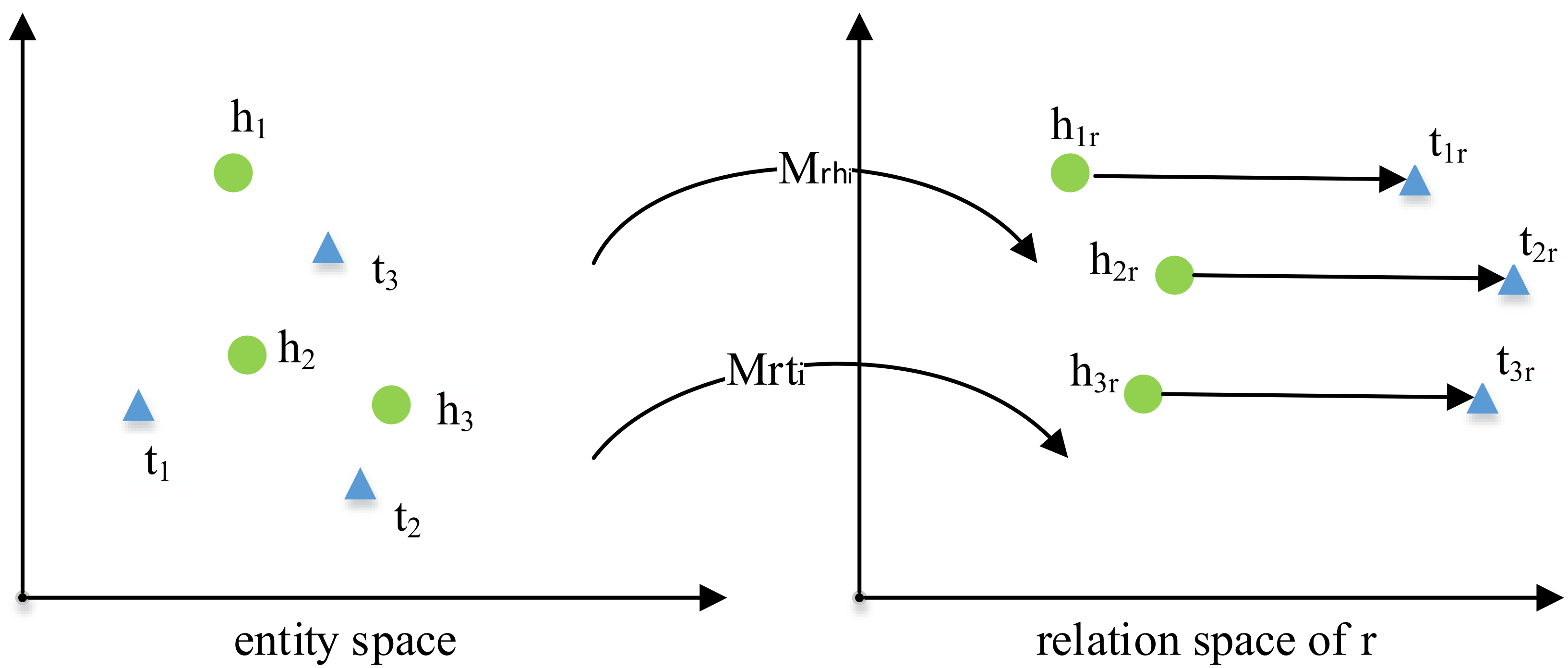
Head	**Barack_Obama**		
Relation	**/people/person/education**		
Model	TransE	TransH	TransR
1	Harvard_College	University_of_Virginia	University_of_Virginia
2	Massachusetts_Institute_of_Technology	George_Washington_University	George_Washington_University
3	American_University	University_of_Michigan	Stanford_University
4	University_of_Michigan	Harvard_College	Harvard_College
5	Columbia_University	Princeton_University	Purdue_University
6	Princeton_University	University_of_Washington	Princeton_University
7	Emory_University	Yale_University	University_of_Michigan
8	Vanderbilt_University	Stanford_University	Occidental_College
9	University_of_Notre_Dame	Purdue_University	University_of_Maryland,_College_Park
10	Texas_A&M_University	Columbia_University	Columbia_University

Examples

Head	University_of_Cambridge		
Relation	/education/education/student		
Model	TransE	TransH	TransR
1	John_Cleese	Stephen_Fry	David_Attenborough
2	Samuel_Beckett	David_Attenborough	Stephen_Fry
3	Harold_Pinter	Ralph_Vaughan_Williams	Stephen_Hawking
4	Virginia_Woolf	Alan_Bennett	Ralph_Vaughan_Williams
5	Graham_Chapman	Francis_Bacon	Alan_Bennett
6	Philip_Pullman	Julian_Fellowes	Julian_Fellowes
7	Ian_McEwan	Hugh_Bonneville	Ernest_Rutherford
8	Douglas_Adams	Graham_Chapman	Jonathan_Lynn
9	Terry_Gilliam	Miriam_Margolyes	Tom_Hollander
10	Richard_Dawkins	Stephen_Hawking	Chris_Weitz

TransD

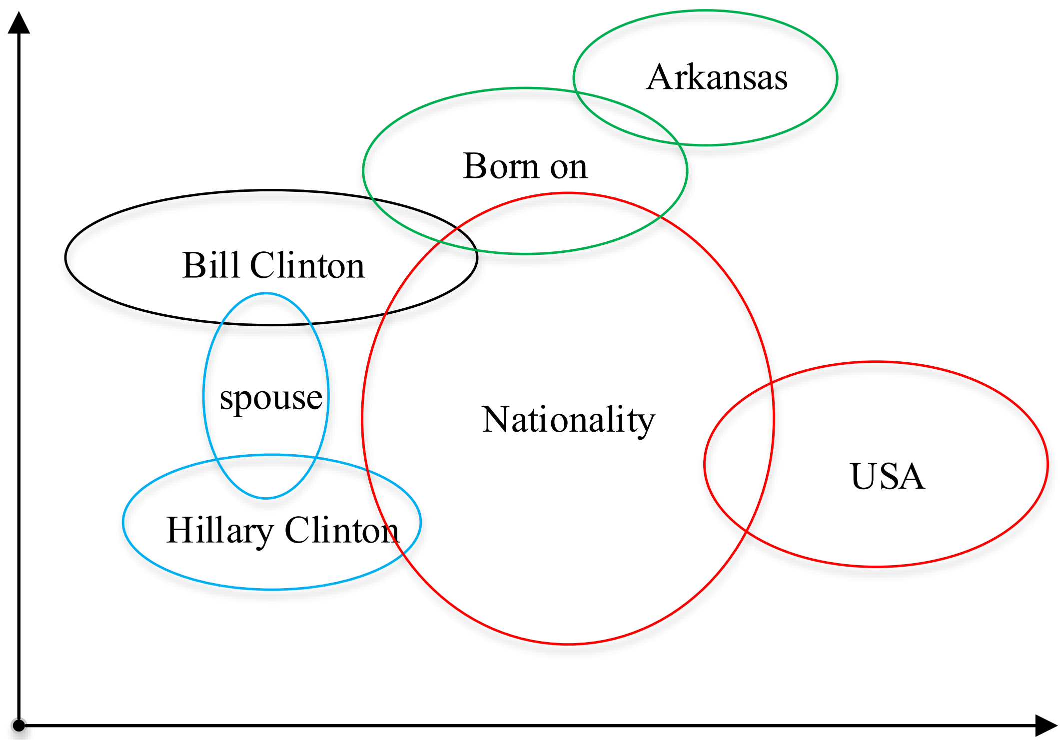
- Projection matrices not only related to relation but also head/tail entities



Ji, et al. (2015) Knowledge Graph Embedding via Dynamic Mapping Matrix. ACL-IJCNLP.

KG2E

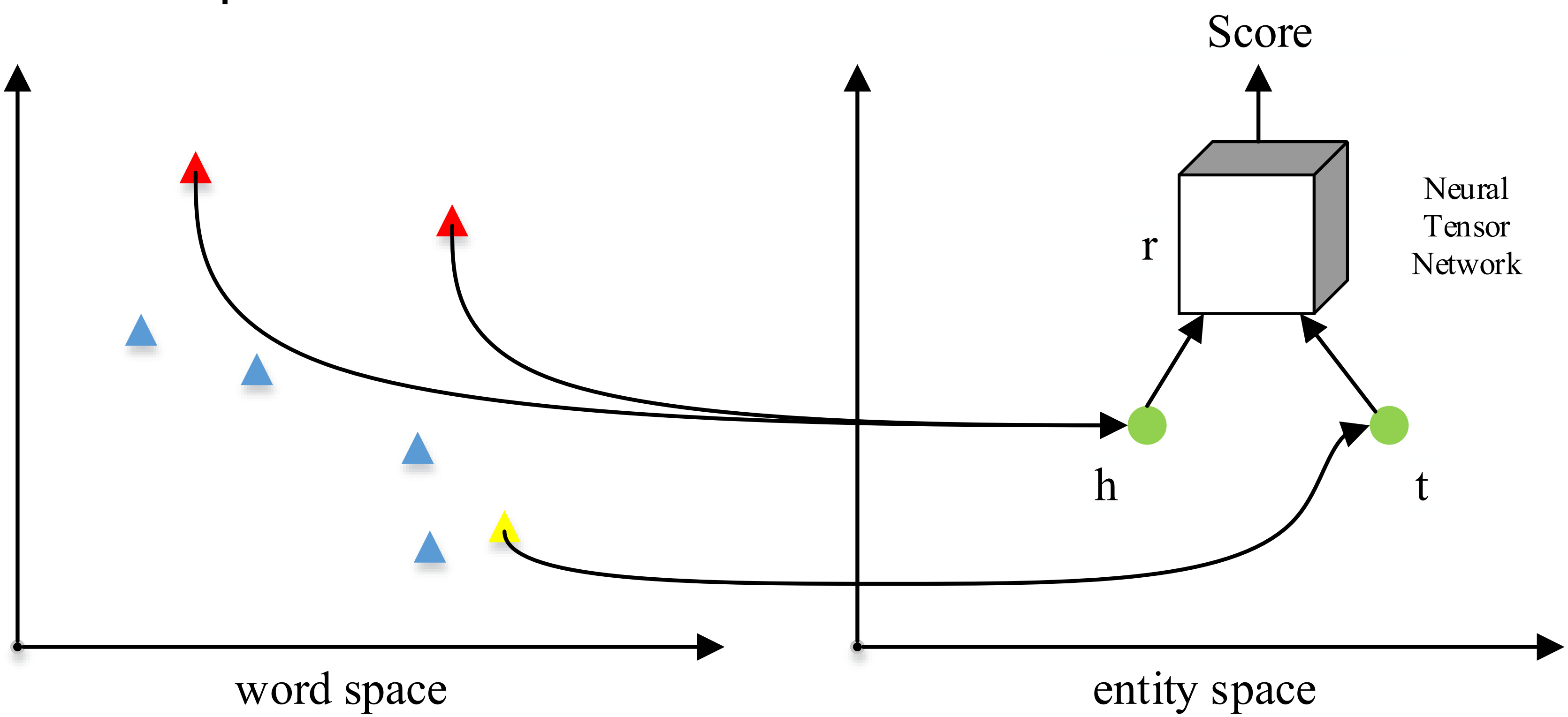
- Represent relations / entities with Gaussian distribution
- Consider (un)certainties of entities and relations



He, et al. (2015) Learning to Represent Knowledge Graphs with Gaussian Embedding. CIKM.

NTN

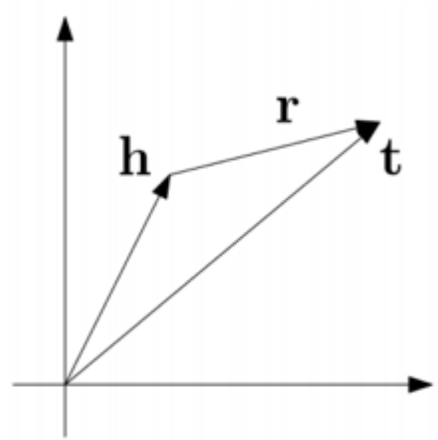
- NTN models KG with a Neural Tensor Network and represents entities via word vectors

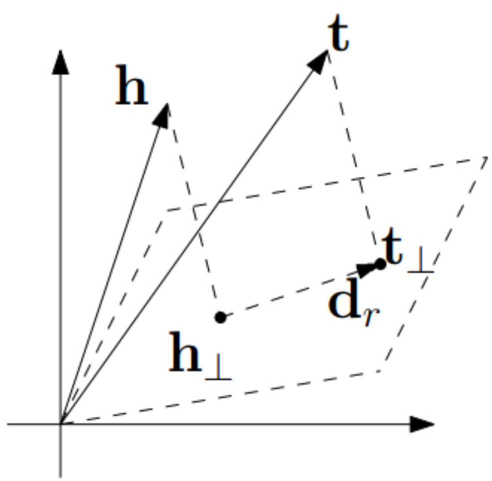


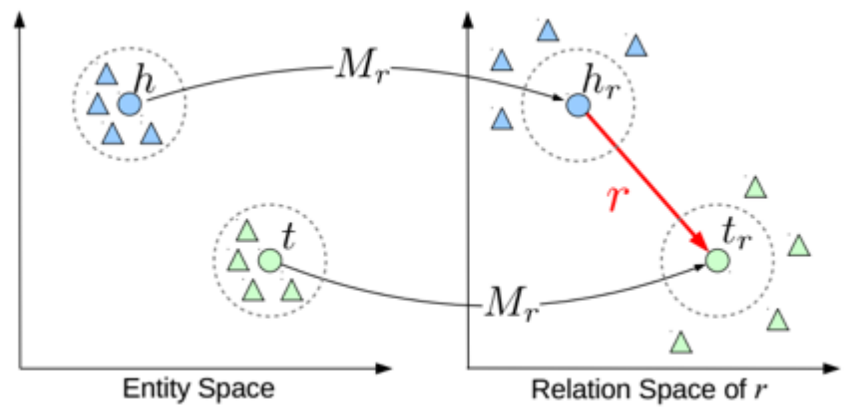
Socher, et al. (2013) Reasoning with neural tensor networks for knowledge base completion. NIPS.

小结

- TransE能够很好地处理1-1关系，但无法处理复杂关系
  - 1-N, N-1, N-N
  - TransA，TransD，TransE，TransG，TransH，TransR, KG2E, TranSparse, Hole







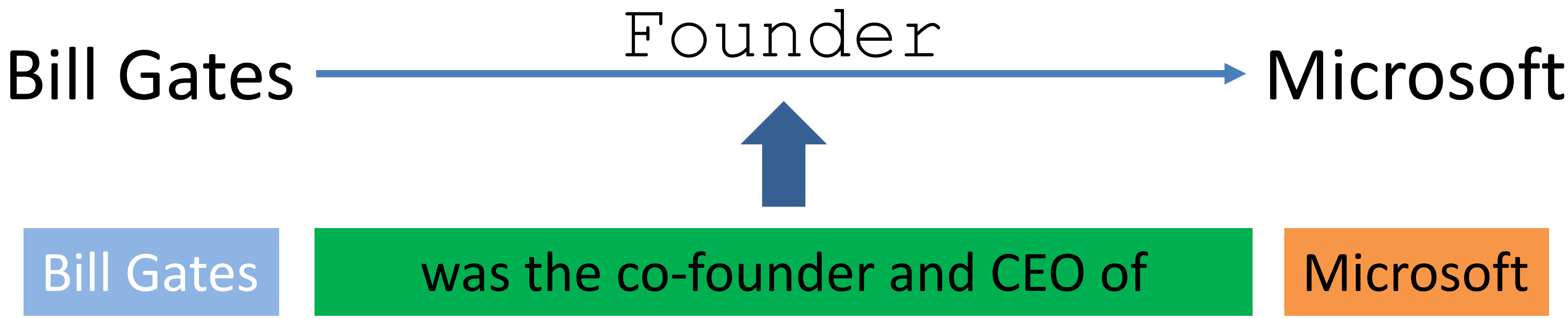
文本与KG融合

融合文本与知识进行关系抽取

- 基于知识图谱的关系预测

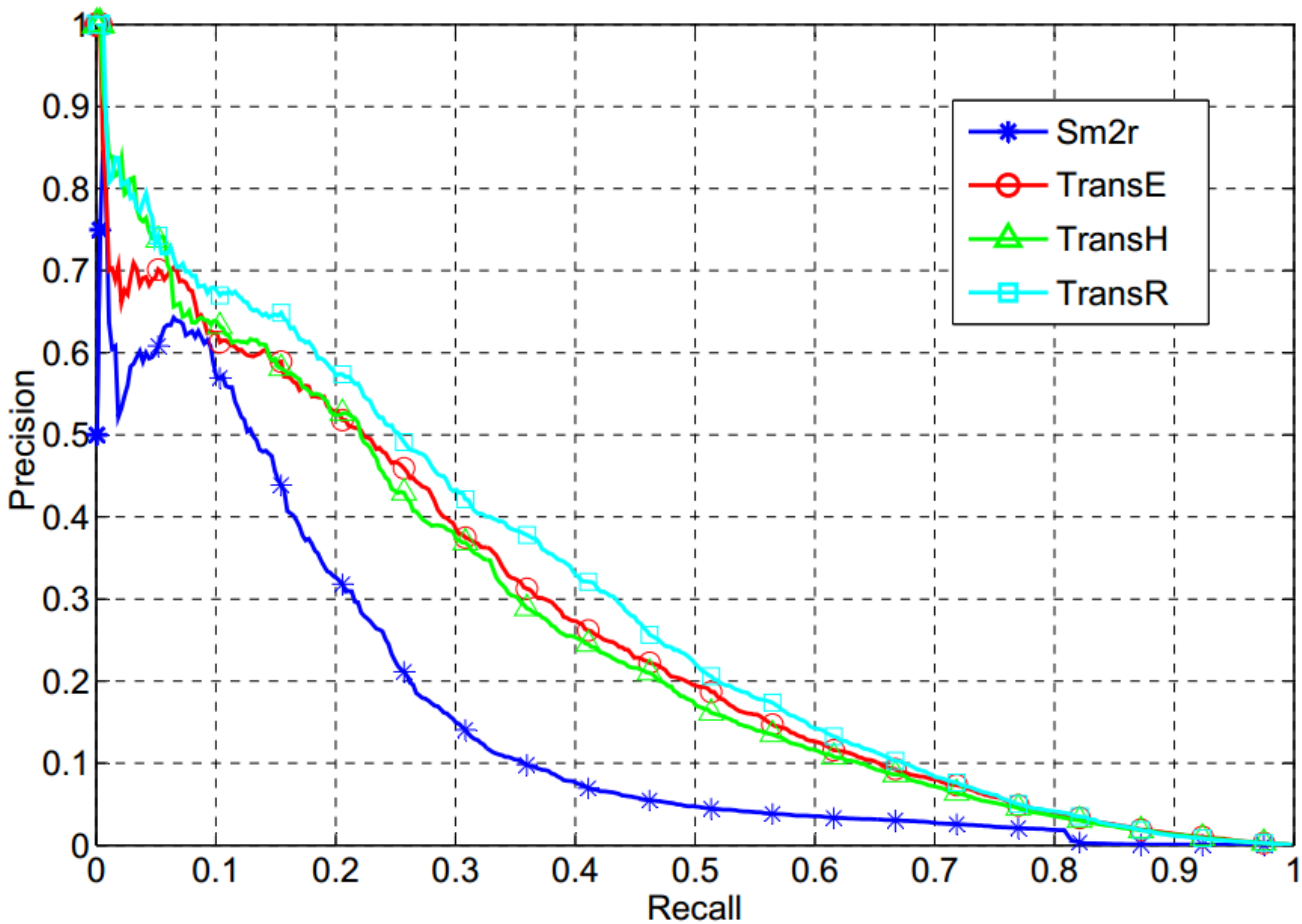
$$\mathbf{r \sim t\text{-}h}$$

- 基于文本信息的关系预测

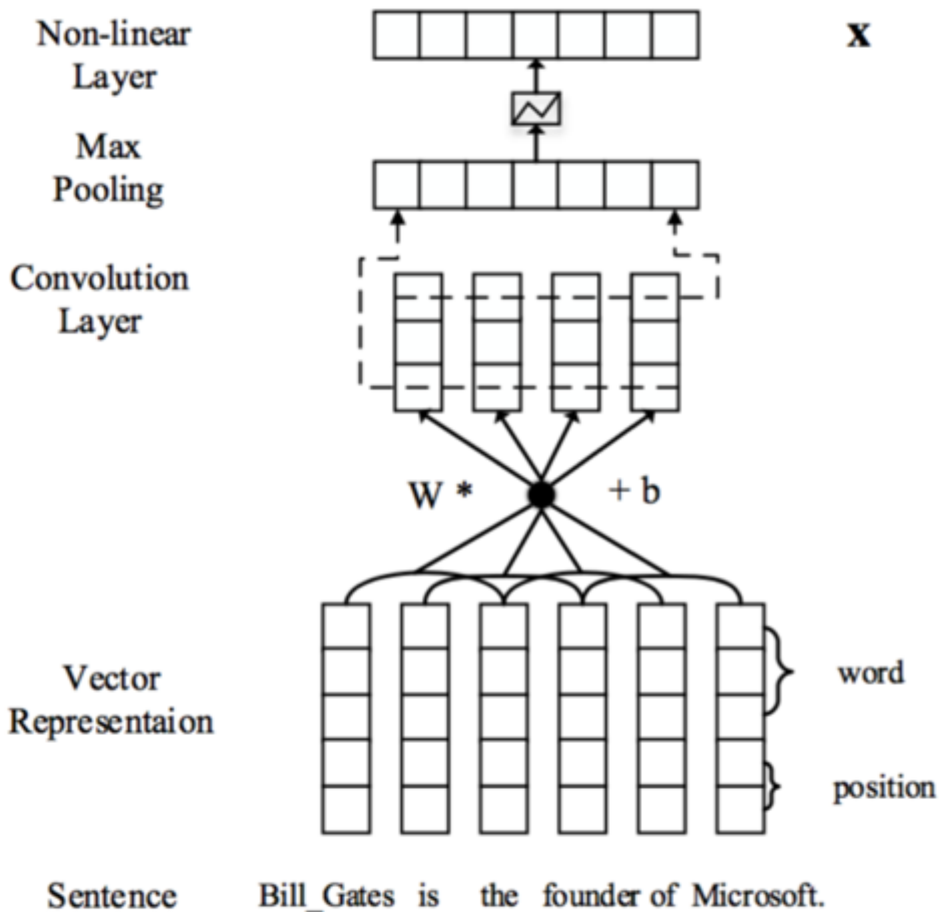


关系抽取效果

- 数据 NYT+FB (Weston et al.2013)



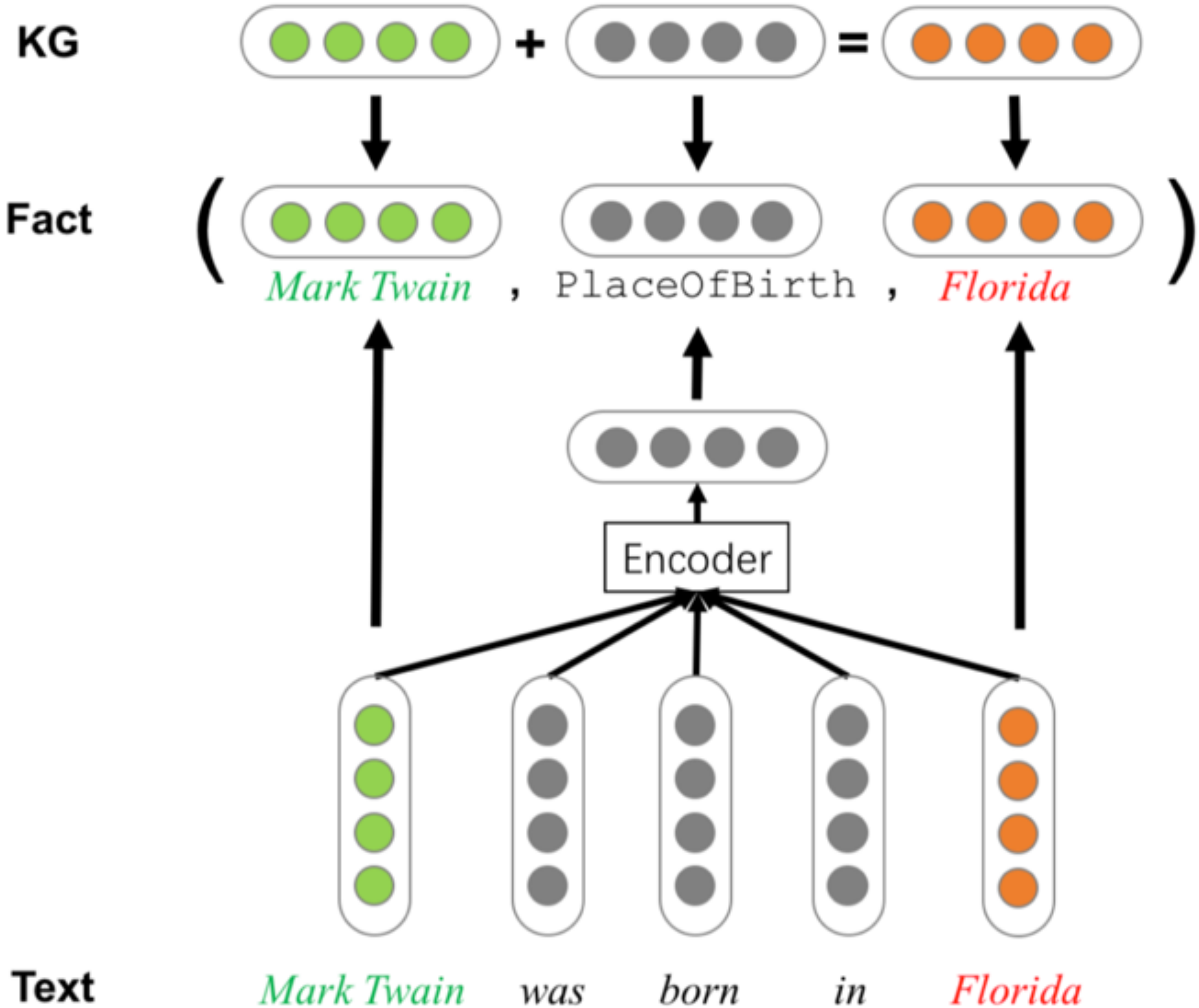
利用CNN进行文本关系抽取



Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. Relation classification via convolutional deep neural network. In Proceedings of COLING.

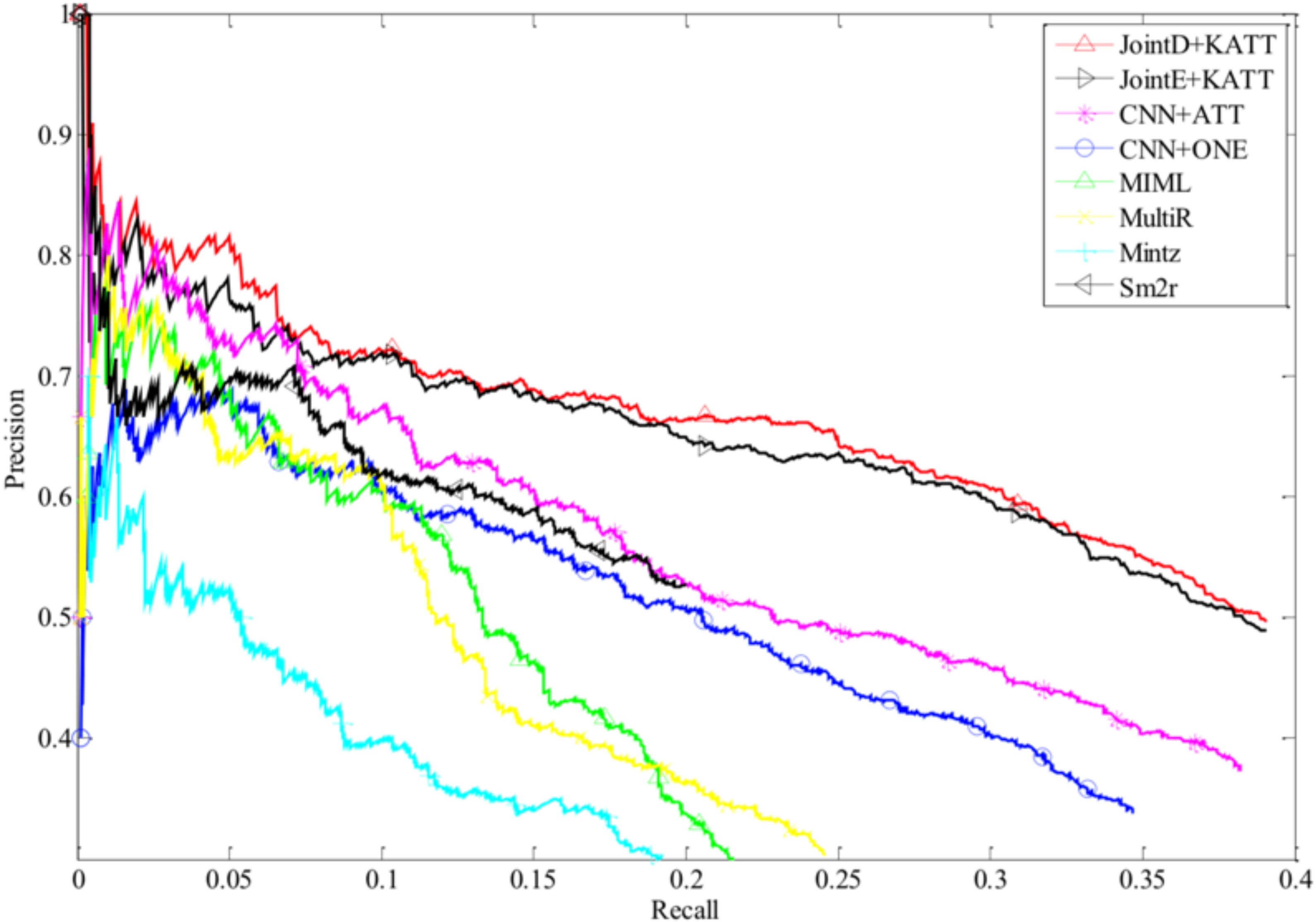
Daojian Zeng,Kang Liu,Yubo Chen,and Jun Zhao. Distant supervision for relation extraction via piecewise convolutional neural networks. In Proceedings of EMNLP.

融合文本与知识进行关系抽取



Xu Han, Zhiyuan Liu, Maosong Sun. Neural Knowledge Acquisition via Mutual Attention between Knowledge Graph and Text. The 32th AAAI Conference on Artificial Intelligence (AAAI 2018).

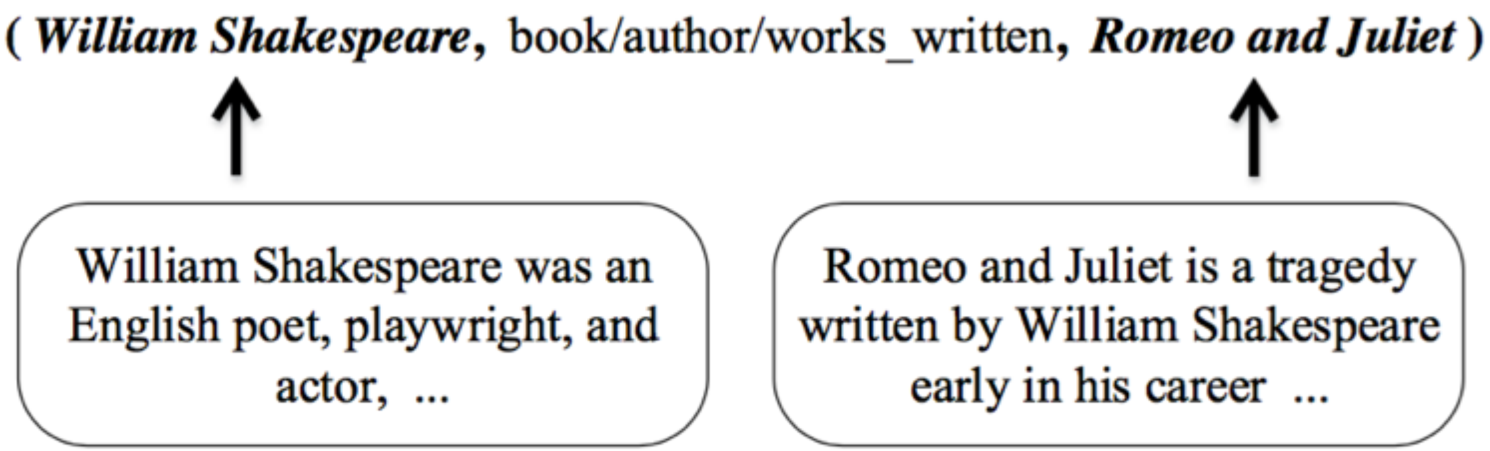
关系抽取结果



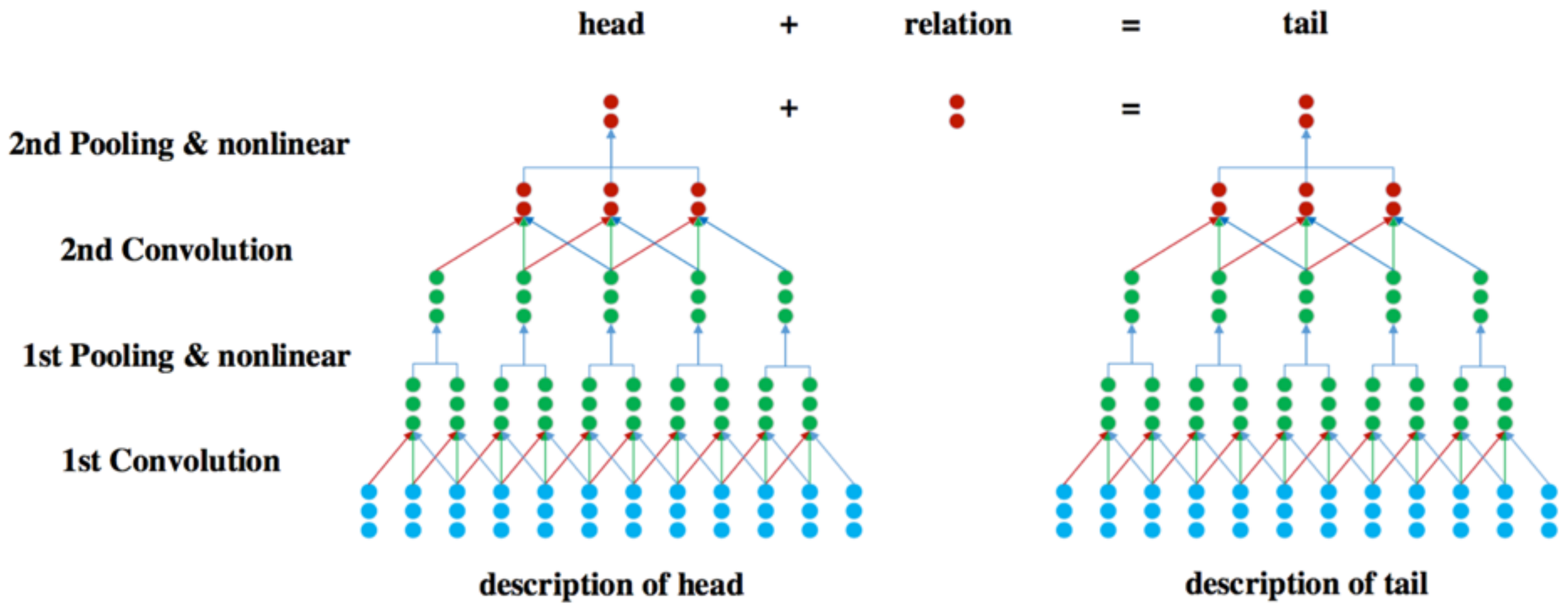
Precision
Recall
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
1
0.9
0.8
0.7
0.6
0.5
0.4
JointD+KATT
JointE+KATT
CNN+ATT
CNN+ONE
MIML
MultiR
Mintz
Sm2r

融合实体描述的知识表示

- 利用实体描述信息提供关于实体的语义信息



融合实体描述的知识表示



Xie, et al. (2016). Representation Learning of Knowledge Graphs with Entity Descriptions. AAAI.

Zero-shot场景下的关系预测

- 对于新实体，根据描述信息有效得到实体表示

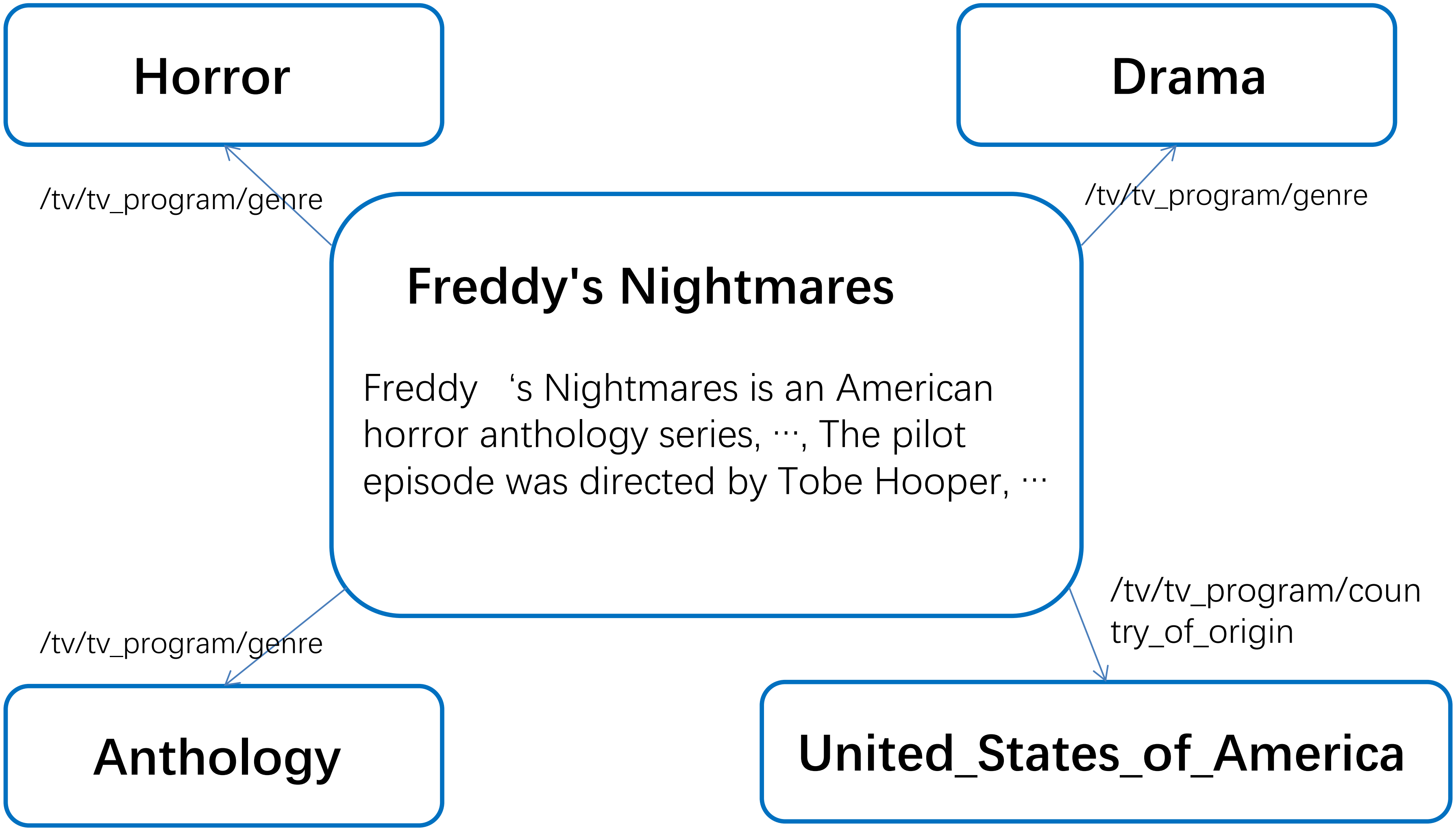
Table 5: Evaluation results on entity prediction in zero-shot scenario

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**

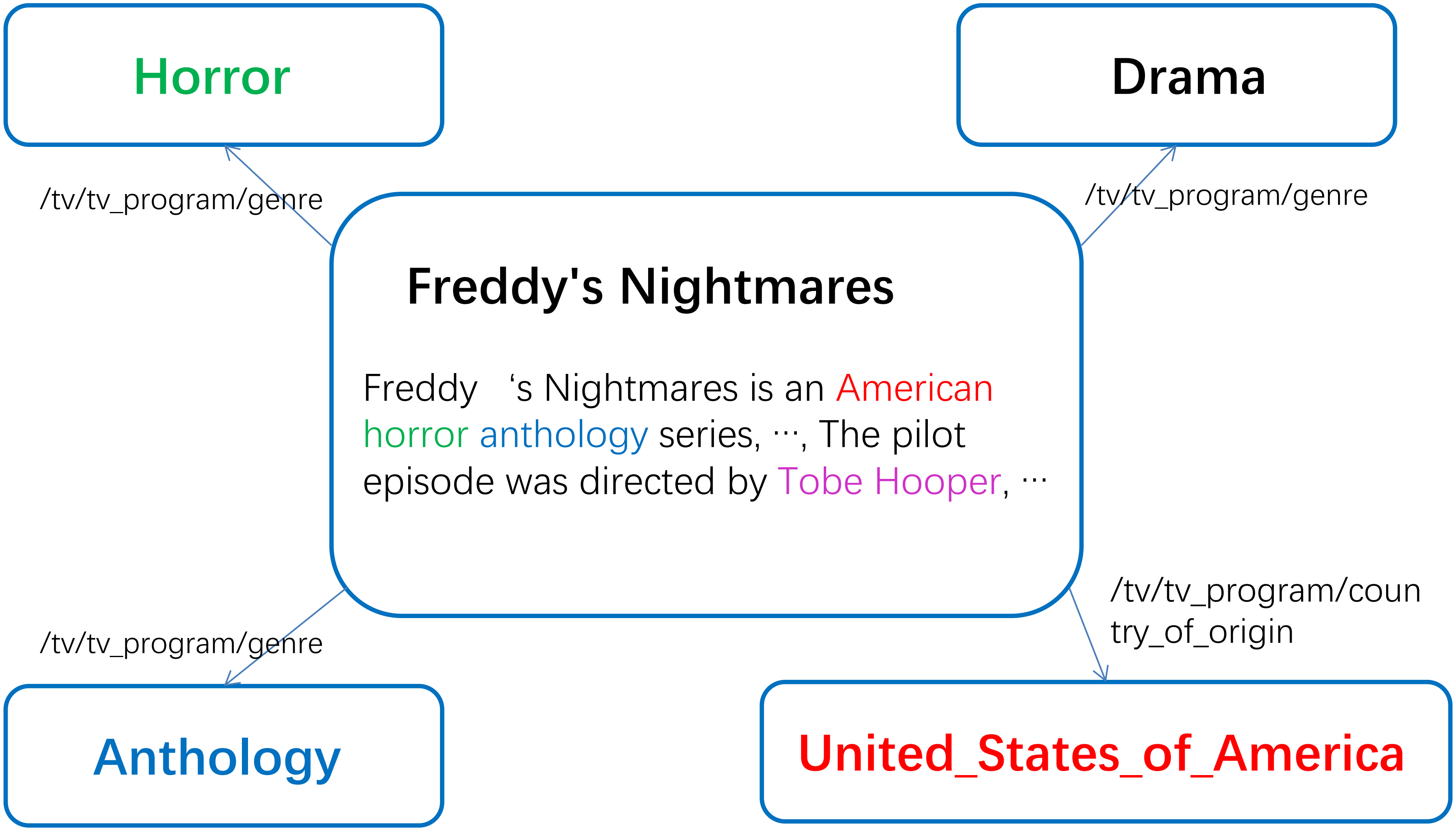
Table 6: Evaluation results on relation prediction in zero-shot scenario

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**

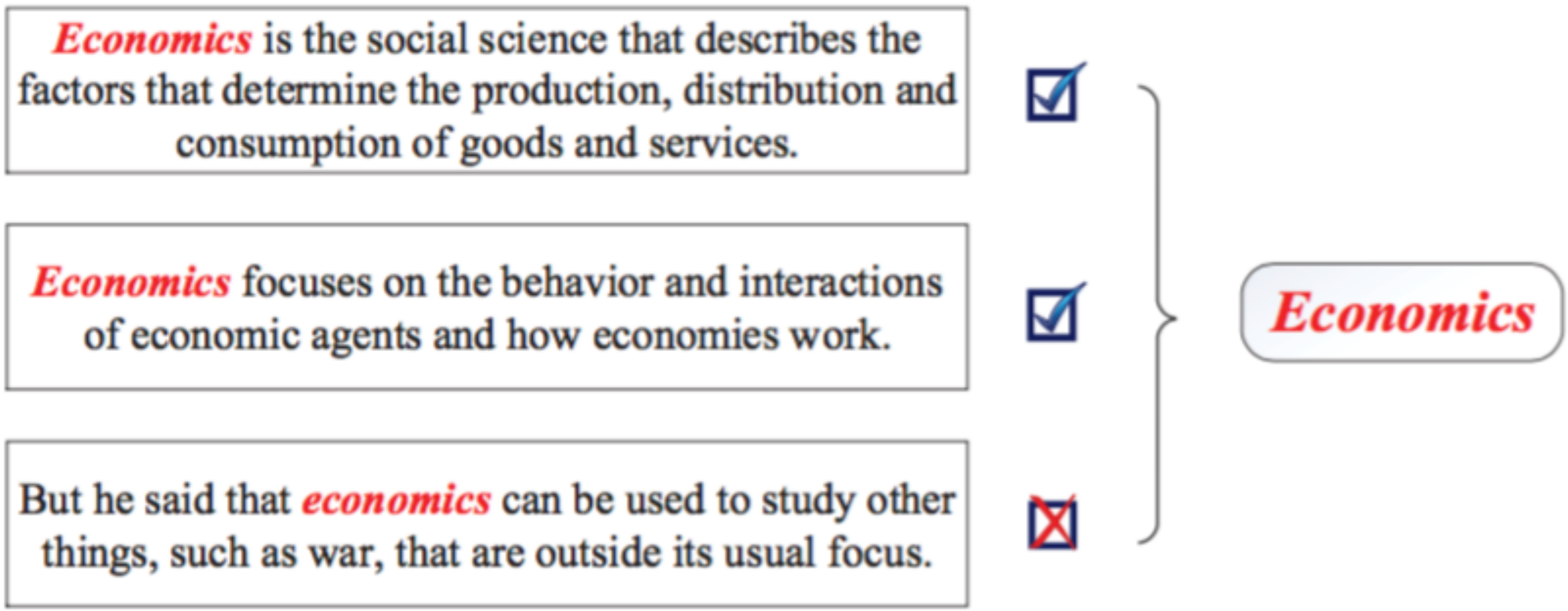
样例


Horror
Drama
/tv/tv_program/genre
/tv/tv_program/genre
Freddy's Nightmares
Freddy 's Nightmares is an American horror anthology series, …, The pilot episode was directed by Tobe Hooper, …
/tv/tv_program/genre
/tv/tv_program/country_of_origin
Anthology
United_States_of_America

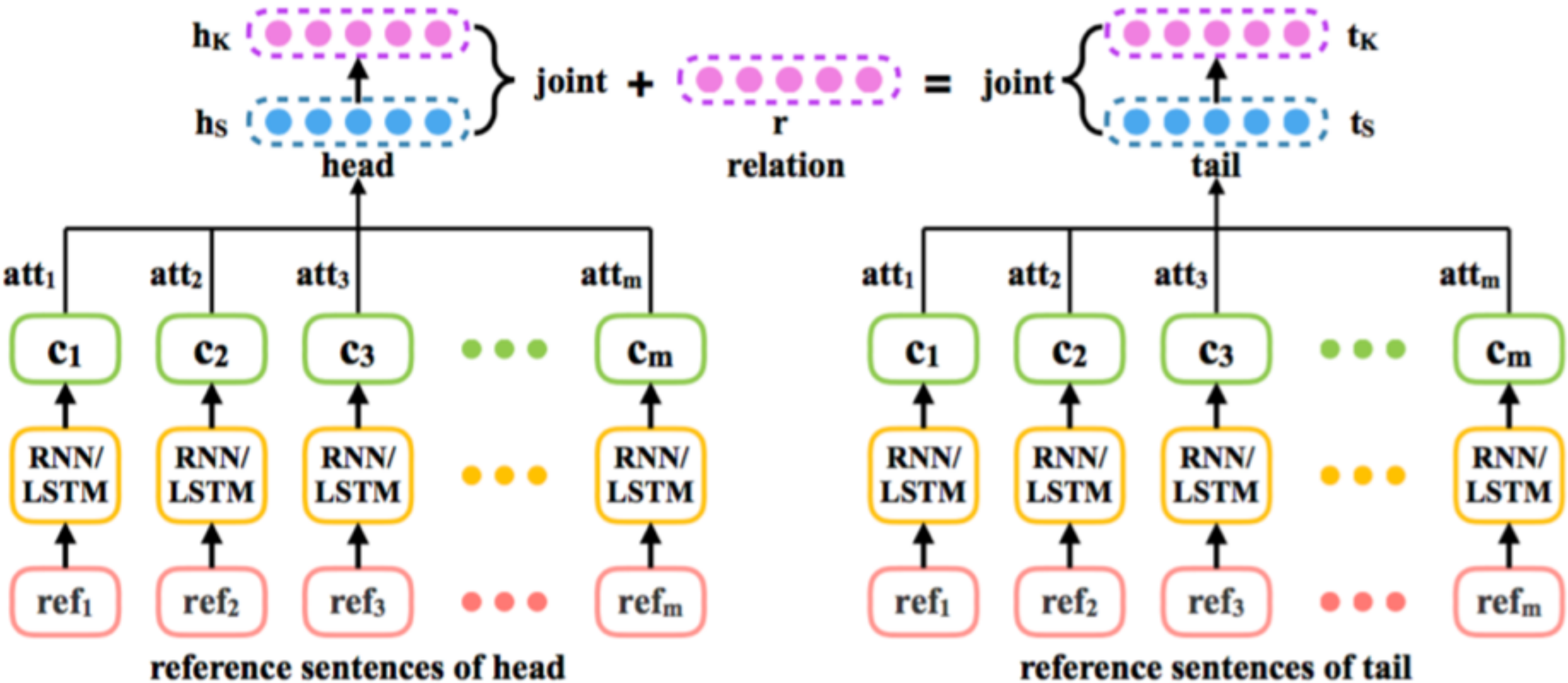
样例


Horror
Drama
/tv/tv_program/genre
/tv/tv_program/genre
Freddy's Nightmares
Freddy 's Nightmares is an American horror anthology series, …, The pilot episode was directed by Tobe Hooper, …
/tv/tv_program/genre
/tv/tv_program/country_of_origin
Anthology
United_States_of_America

融合实体所在句子的知识表示



融合实体所在句子的知识表示



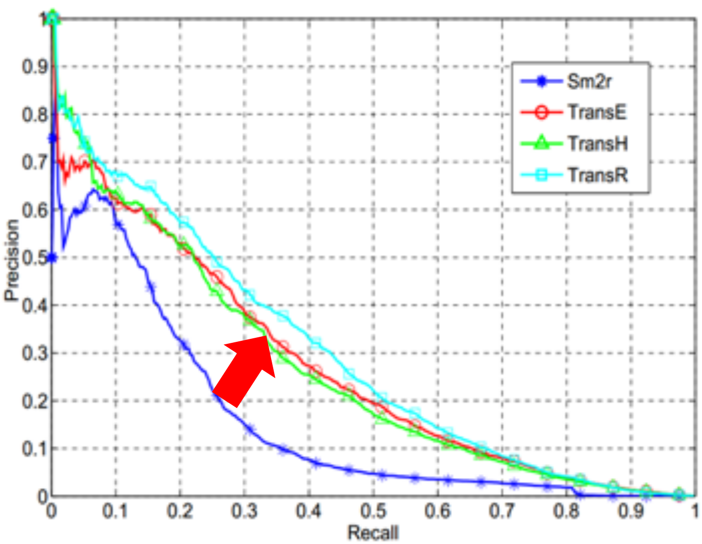
样例

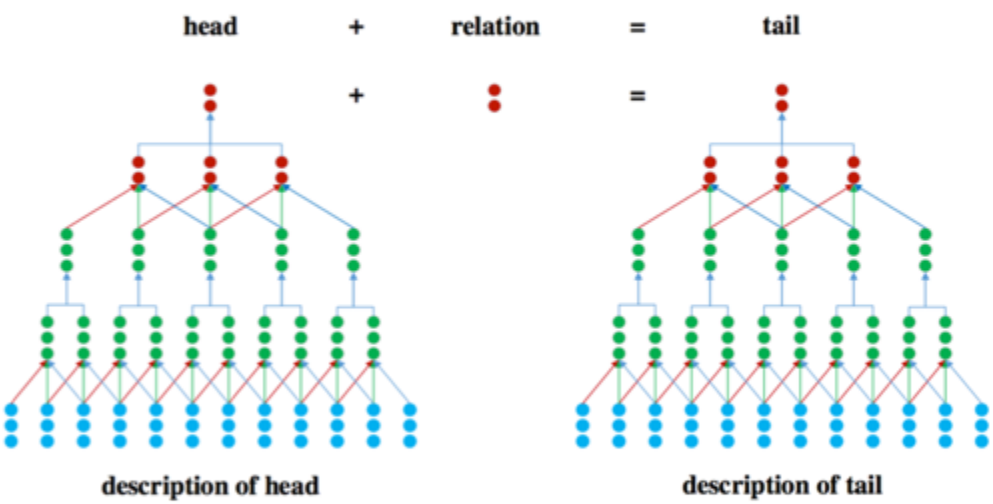
Rank	Sentence
1	***Economics*** is the social science that describes the factors that determine the production, distribution and consumption of goods and services.
2	***Economics*** focuses on the behavior and interactions of economic agents and how economies work.
10	The ultimate goal of ***economics*** is to improve the living conditions of people in their everyday life.
44	There are a variety of modern definitions of ***economics***.

Entity	Rank No.1 Sentence
Productivity	***Productivity*** is the ratio of output to inputs in production.
February	***February*** is the second month of the year in the Julian and Gregorian calendars.
Food	***Food*** is any substance consumed to provide nutritional support for the body.
Travis County	Austin is the capital of Texas and the seat of ***Travis County***.

小结

- 文本能够有效辅助知识表示学习
- 纯文本，实体描述等



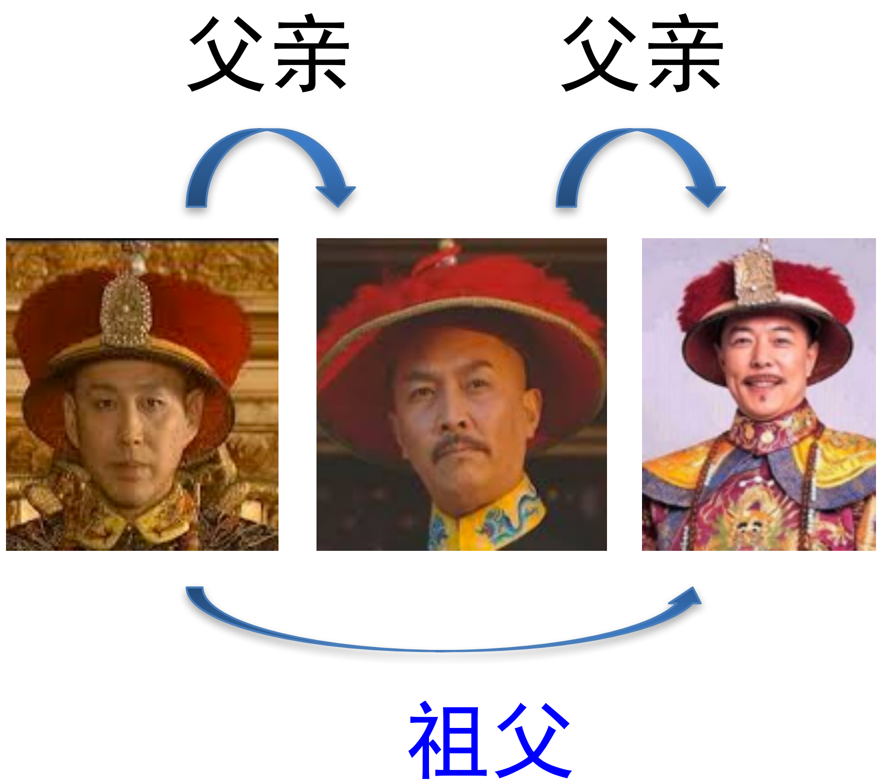


对关系路径建模

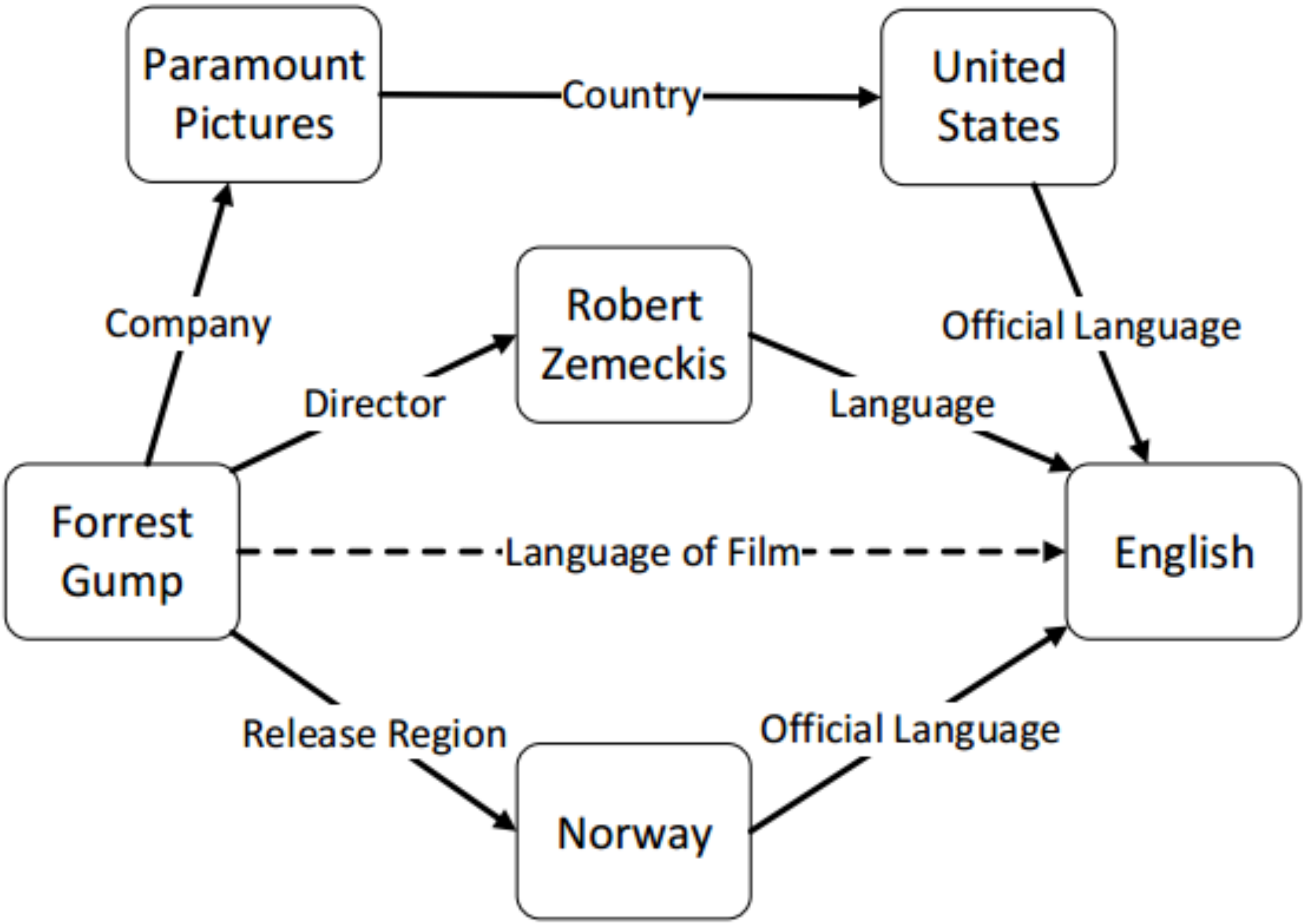
Y Lin, Z Liu, M Sun. Modeling Relation Paths for Representation Learning of Knowledge Bases, EMNLP 2015.

知识推理

- 目前模型孤立地学习每个事实三元组
- 实际上关系之间存在复杂的关系，涉及关系推理



知识图谱关系之间存在复杂推理关系



关系路径用于关系抽取

- Path Ranking Algorithm

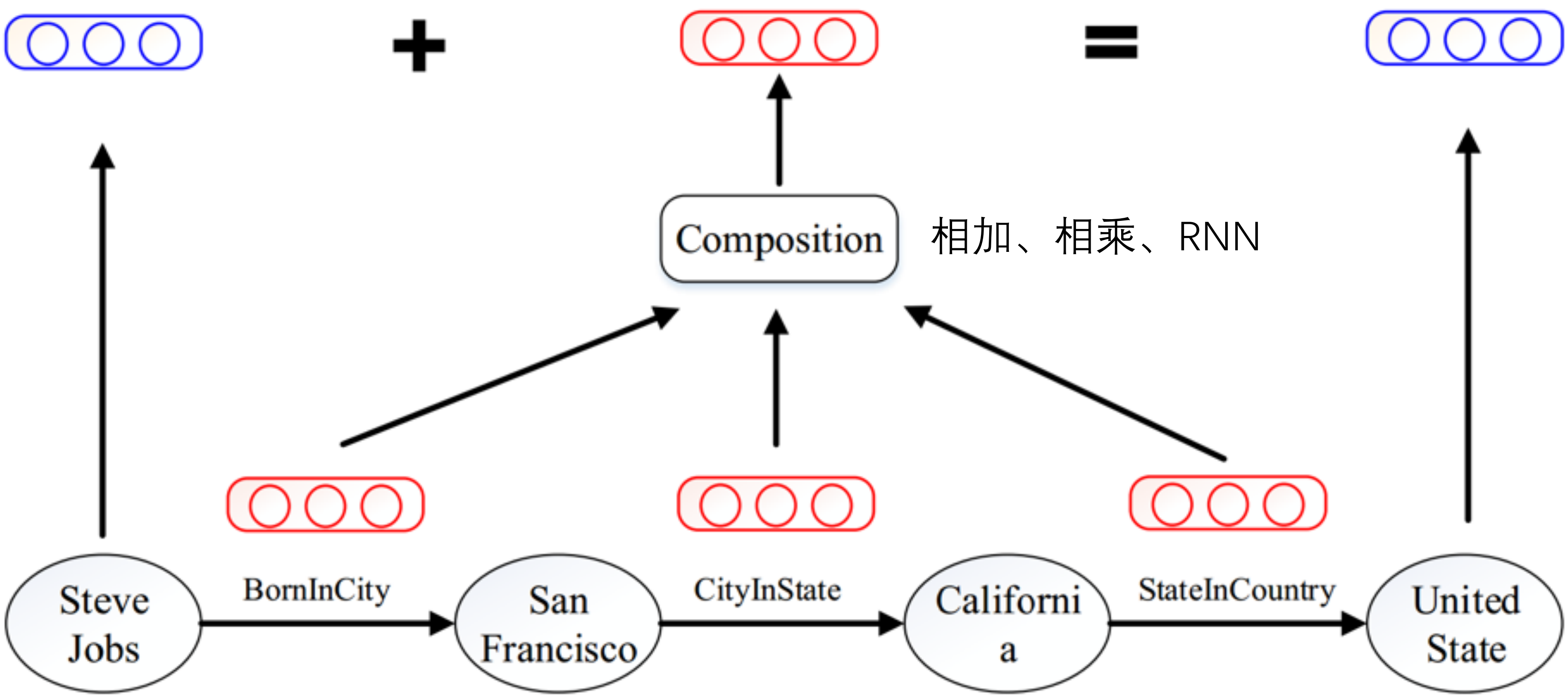
ID	PRA Path (Comment)
	athletePlaysForTeam
1	$c \xrightarrow{\text{athletePlaysInLeague}} c \xrightarrow{\text{leaguePlayers}} c \xrightarrow{\text{athletePlaysForTeam}} c$ (teams with many players in the athlete's league)
2	$c \xrightarrow{\text{athletePlaysInLeague}} c \xrightarrow{\text{leagueTeams}} c \xrightarrow{\text{teamAgainstTeam}} c$ (teams that play against many teams in the athlete's league)
	athletePlaysInLeague
3	$c \xrightarrow{\text{athletePlaysSport}} c \xrightarrow{\text{players}} c \xrightarrow{\text{athletePlaysInLeague}} c$ (the league that players of a certain sport belong to)
4	$c \xrightarrow{\text{isa}} c \xrightarrow{\text{isa}^{-1}} c \xrightarrow{\text{athletePlaysInLeague}} c$ (popular leagues with many players)
	athletePlaysSport
5	$c \xrightarrow{\text{isa}} c \xrightarrow{\text{isa}^{-1}} c \xrightarrow{\text{athletePlaysSport}} c$ (popular sports of all the athletes)
6	$c \xrightarrow{\text{athletePlaysInLeague}} c \xrightarrow{\text{superpartOfOrganization}} c \xrightarrow{\text{teamPlaysSport}} c$ (popular sports of a certain league)
	stadiumLocatedInCity
7	$c \xrightarrow{\text{stadiumHomeTeam}} c \xrightarrow{\text{teamHomeStadium}} c \xrightarrow{\text{stadiumLocatedInCity}} c$ (city of the stadium with the same team)
8	$c \xrightarrow{\text{latitudeLongitude}} c \xrightarrow{\text{latitudeLongitudeOf}} c \xrightarrow{\text{stadiumLocatedInCity}} c$ (city of the stadium with the same location)
	teamHomeStadium
9	$c \xrightarrow{\text{teamPlaysInCity}} c \xrightarrow{\text{cityStadiums}} c$ (stadiums located in the same city with the query team)
10	$c \xrightarrow{\text{teamMember}} c \xrightarrow{\text{athletePlaysForTeam}} c \xrightarrow{\text{teamHomeStadium}} c$ (home stadium of teams which share players with the query)
	teamPlaysInCity
11	$c \xrightarrow{\text{teamHomeStadium}} c \xrightarrow{\text{stadiumLocatedInCity}} c$ (city of the team's home stadium)
12	$c \xrightarrow{\text{teamHomeStadium}} c \xrightarrow{\text{stadiumHomeTeam}} c \xrightarrow{\text{teamPlaysInCity}} c$ (city of teams with the same home stadium as the query)
	teamPlaysInLeague
13	$c \xrightarrow{\text{teamPlaysSport}} c \xrightarrow{\text{players}} c \xrightarrow{\text{athletePlaysInLeague}} c$ (the league that the query team's members belong to)
14	$c \xrightarrow{\text{teamPlaysAgainstTeam}} c \xrightarrow{\text{teamPlaysInLeague}} c$ (the league that the query team's competing team belongs to)
	teamPlaysSport
15	$c \xrightarrow{\text{isa}} c \xrightarrow{\text{isa}^{-1}} c \xrightarrow{\text{teamPlaysSport}} c$ (sports played by many teams)
16	$c \xrightarrow{\text{teamPlaysInLeague}} c \xrightarrow{\text{leagueTeams}} c \xrightarrow{\text{teamPlaysSport}} c$ (the sport played by other teams in the league)

Lao, et al. (2011). Random walk inference and learning in a large scale knowledge base. EMNLP.

PTransE：考虑关系路径的TransE

	TransE	PTransE
KB	$h \xrightarrow{r} t$	$h \xrightarrow{r_1} e_1 \xrightarrow{r_2} t$
Triples	(h, r, t)	(h, r_1, e_1) (e_1, r_2, t) $(h, r_1 \circ r_2, t)$
Objectives	$\mathbf{h} + \mathbf{r} = \mathbf{t}$	$\mathbf{h} + \mathbf{r}_1 = \mathbf{e}_1$ $\mathbf{e}_1 + \mathbf{r}_2 = \mathbf{t}$ $\mathbf{h} + (\mathbf{r}_1 \circ \mathbf{r}_2) = \mathbf{t}$

PTransE：考虑关系路径的TransE



实体预测结果

Metric	Mean Rank		Hits@10 (%)	
	Raw	Filter	Raw	Filter
RESCAL	828	683	28.4	44.1
SE	273	162	28.8	39.8
SME (linear)	274	154	30.7	40.8
SME (bilinear)	284	158	31.3	41.3
LFM	283	164	26.0	33.1
TransE	243	125	34.9	47.1
TransH	212	87	45.7	64.4
TransR	**198**	77	48.2	68.7
TransE (Our)	205	63	47.9	70.2
PTransE (ADD, 2-step)	200	**54**	**51.8**	83.4
PTransE (MUL, 2-step)	216	67	47.4	77.7
PTransE (RNN, 2-step)	242	92	50.6	82.2
PTransE (ADD, 3-step)	207	58	51.4	**84.6**

+35%

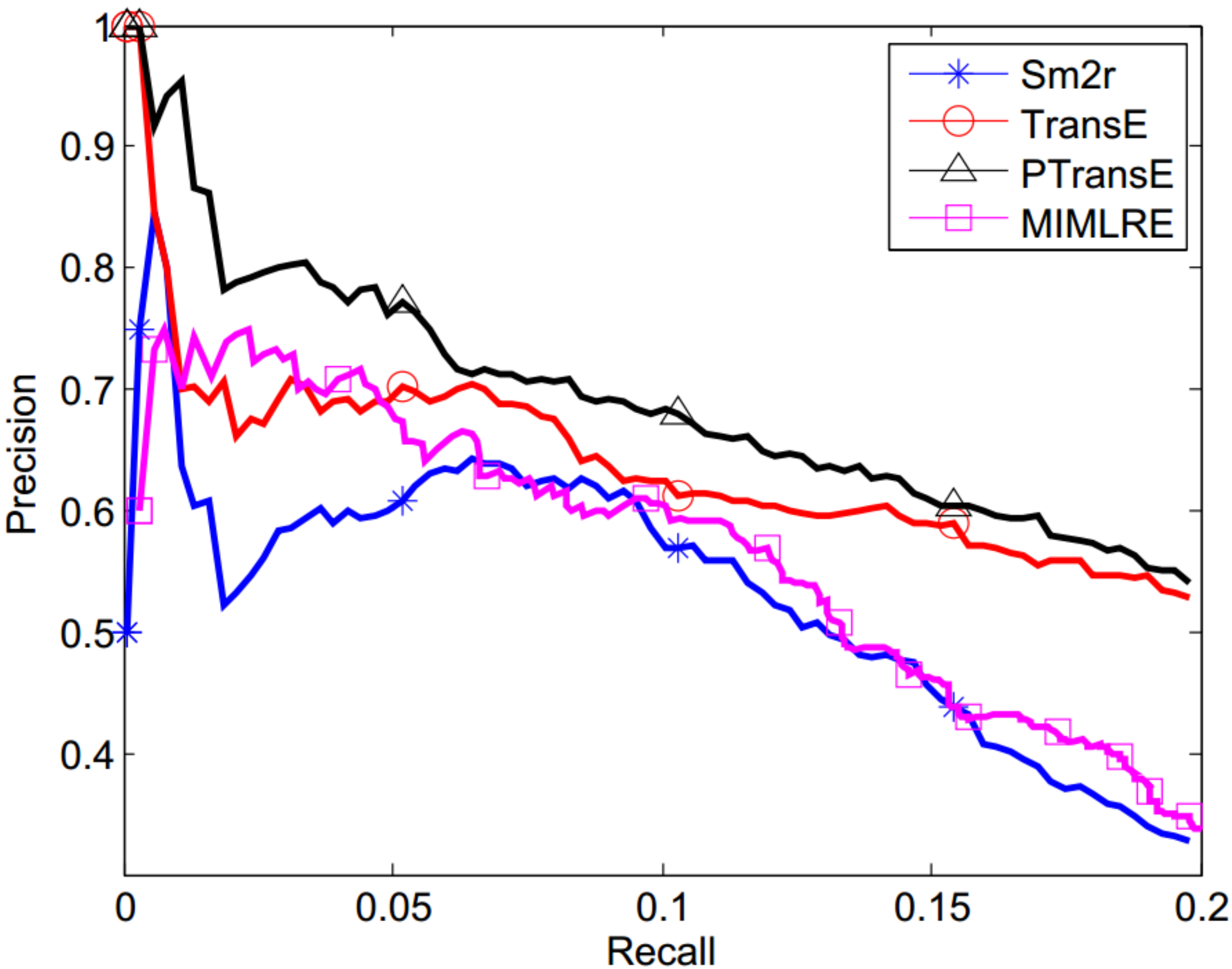
Lin, et al. (2015). Modeling Relation Paths for Representation Learning of Knowledge Bases. EMNLP.

关系预测结果

Metric	Mean Rank		Hits@1 (%)	
	Raw	Filter	Raw	Filter
TransE	2.8	2.5	65.1	$\boxed{84.3}$
+Rev	2.6	2.3	67.1	86.7
+Rev+Path	2.4	1.9	65.2	89.0
PTransE (ADD, 2-step)	**1.7**	**1.2**	69.5	93.6
-TransE	135.8	135.3	51.4	78.0
-Path	2.0	1.6	**69.7**	89.0
PTransE (MUL, 2-step)	2.5	2.0	66.3	89.0
PTransE (RNN, 2-step)	1.9	1.4	68.3	93.2
PTransE (ADD, 3-step)	1.8	1.4	68.5	$\boxed{\mathbf{94.0}}$

+10%

对关系抽取的帮助



Precision
Recall
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.05
0.1
0.15
0.2
Sm2r
TransE
PTransE
MIMLRE

样例

Head	Barack_Obama	
Relation	**/education/education/institution**	
Model	TransE	PTransE
1	Harvard_College	Columbia_University
2	Massachusetts_Institute_of_Technology	Occidental_College
3	American_University	Punahou_School
4	University_of_Michigan	University_of_Chicago
5	Columbia_University	Stanford_University
6	Princeton_University	Princeton_University
7	Emory_University	University_of_Pennsylvania
8	Vanderbilt_University	University_of_Virginia
9	University_of_Notre_Dame	University_of_Michigan
10	Texas_A&M_University	Yale_University

样例

Head	**Stanford_University**	
Relation	**/education/educational_institution/students_graduates**	
Model	TransE	PTransE
1	Steven_Spielberg	Raymond_Burr
2	Ron_Howard	Ted_Danson
3	Stan_Lee	Delmer_Daves
4	Barack_Obama	D.W._Moffett
5	Milton_Friedman	Gale_Anne_Hurd
6	Walter_F._Parkes	Jack_Palance
7	Michael_Cimino	Kal_Penn
8	Gale_Anne_Hurd	Kurtwood_Smith
9	Bryan_Singer	Alexander_Payne
10	Aaron_Sorkin	Richard_D._Zanuck

关系路径样例

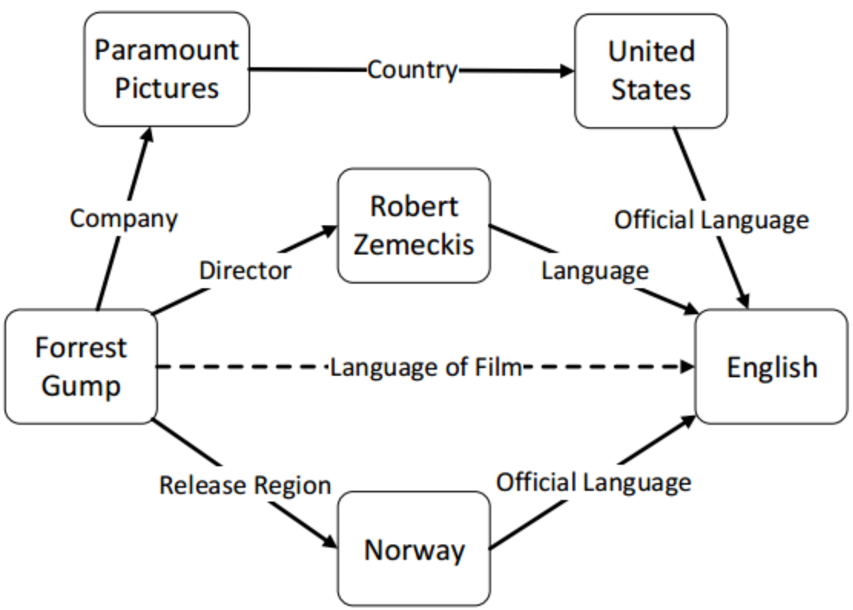
Relation1	**/people/person/place_of_birth**
Relation2	**/location/administrative_division/country**
1	/people/person/nationality
2	/people/person/places_lived./people/place_lived/location
3	/people/person/place_of_birth
4	/music/artist/origin
5	/olympics/olympic_athlete_affiliation/country
6	/government/politician/government_positions_held
7	/base/popstra/vacation_choice/location
8	/people/deceased_person/place_of_death
9	/government/political_appointer/appointees
10	/location/administrative_division/country

关系路径样例

Relation1	**/location/location/contains**
Relation2	**/location/location/contains**
1	/location/location/contains
2	/location/country/second_level_divisions
3	/location/country/administrative_divisions
4	/location/administrative_division/capital
5	/base/locations/continents/countries_within
6	/base/aareas/schema/administrative_area/administrative_children
7	/location/us_county/hud_county_place
8	/location/country/capital
9	/location/country/first_level_divisions
10	/travel/travel_destination/tourist_attractions

小结

- 关系路径蕴藏重要推理信息
- 未来需要考虑更复杂的推理信息



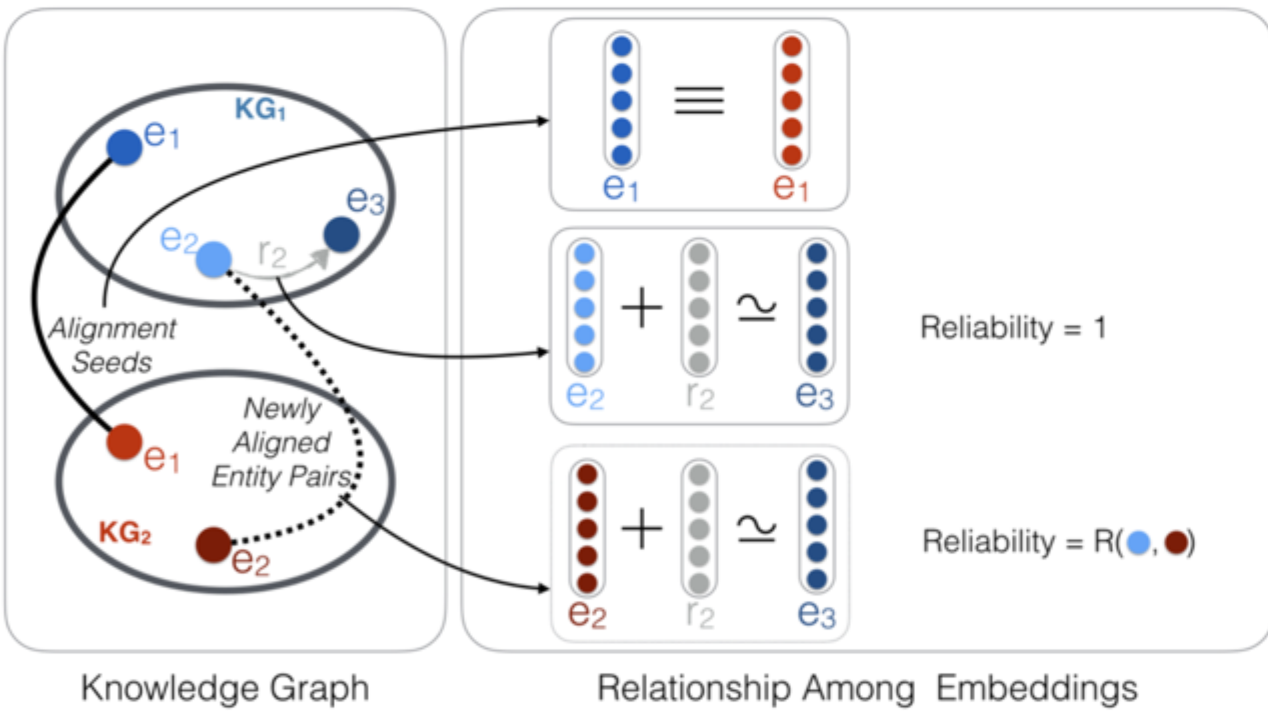
(Obama, _president, USA)

(Obama, _is, American)

知识表示示范应用

实体对齐

- 在两个异质KG之间，根据少量seed对齐实体，可实现大量实体对齐
- 分别学习两KG知识表示，建立两者映射关系



Hao Zhu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. Iterative Entity Alignment via Joint Knowledge Embeddings. International Joint Conference on Artificial Intelligence (IJCAI 2017).

实体对齐

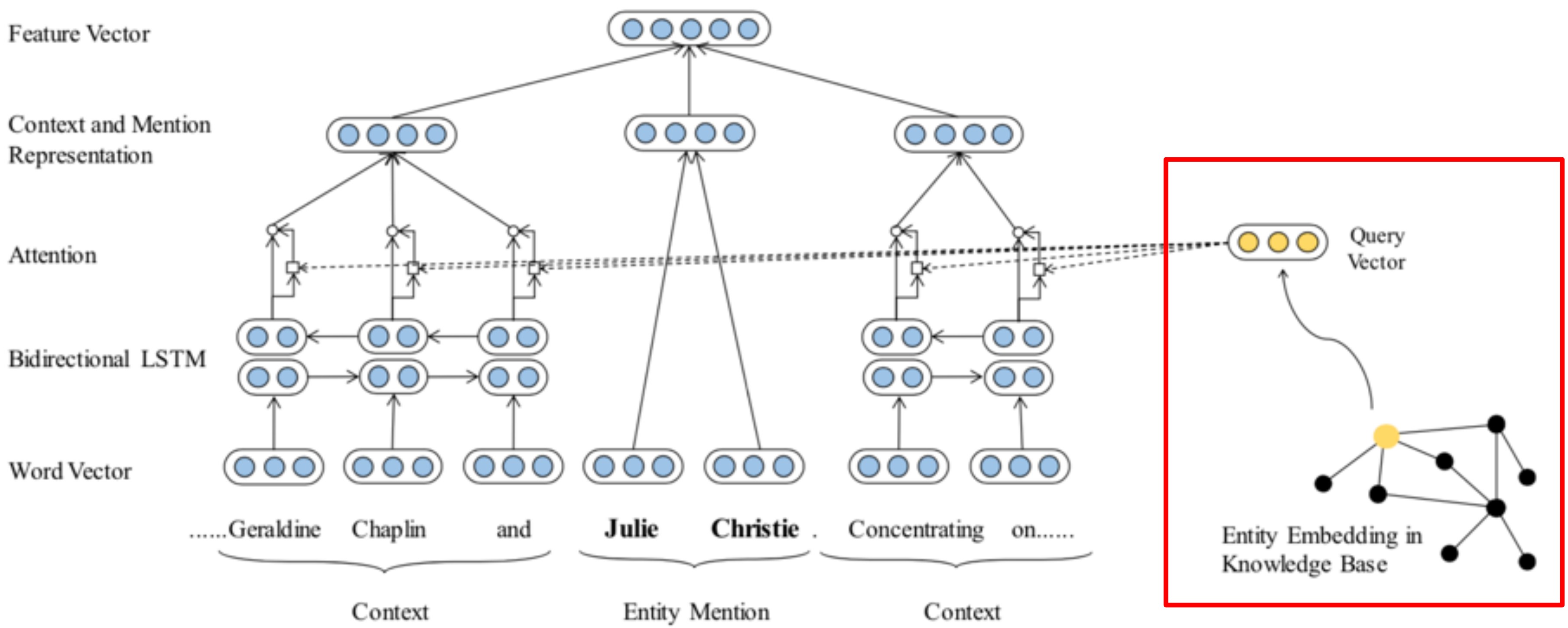
- 在两个异质KG之间，根据少量seed对齐实体，可实现大量实体对齐
- 分别学习两KG知识表示，建立两者映射关系
- 实验证明该方案能够有效提升实体对齐效果

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**

Hao Zhu, Ruobing Xie, Zhiyuan Liu, Maosong Sun. Iterative Entity Alignment via Joint Knowledge Embeddings. International Joint Conference on Artificial Intelligence (IJCAI 2017).

实体分类

- 对文本实体进行细粒度分类，助力深度分析
- 充分利用KG实体表示，提出知识注意力机制，建立对上下文的高效建模



Ji Xin, Yankai Lin, Zhiyuan Liu, Maosong Sun. Improving Neural Fine-Grained Entity Typing with Knowledge Attention. The 32th AAAI Conference on Artificial Intelligence (AAAI 2018).

实体分类

- 对文本实体进行细粒度分类，助力深度分析
- 充分利用KG实体表示，提出知识注意力机制，建立对上下文的高效建模
- 显著提升实体分类性能

Dateset	WIKI-AUTO							WIKI-MAN						
Metrics	Strict	Macro			Micro			Strict	Macro			Micro		
	Acc	Pre	Rec	F1	Pre	Rec	F1	Acc	Pre	Rec	F1	Pre	Rec	F1
AFET	20.32	67.00	45.82	54.75	69.29	42.40	52.61	18.00	64.50	50.00	56.33	64.29	50.43	56.52
KB-ONLY	35.12	69.65	71.35	70.49	54.85	74.99	63.36	17.00	55.50	72.83	63.00	27.81	74.57	40.52
HNM	34.88	68.09	61.03	64.37	72.80	64,48	68.39	15.00	61.80	68.00	64.75	62.35	68.53	65.30
SA	42.77	75.33	69.69	72.40	77.35	72.63	74.91	18.00	66.67	73.67	69.44	65.54	75.43	70.14
MA	41.58	73.64	71.71	72.66	75.94	75.52	75.72	26.00	65.13	78.50	71.19	64.09	82.33	72.08
KA	45.49	74.82	72.46	73.62	76.96	75.49	76.22	23.00	64.69	78.92	71.10	63.25	82.68	71.67
KA+D	**47.20**	**75.72**	**74.03**	**74.87**	**77.96**	**77.87**	**77.92**	**34.00**	**68.41**	**82.83**	**74.94**	**66.12**	**87.50**	**75.32**

Ji Xin, Yankai Lin, Zhiyuan Liu, Maosong Sun. Improving Neural Fine-Grained Entity Typing with Knowledge Attention. The 32th AAAI Conference on Artificial Intelligence (AAAI 2018).

知识表示的研究趋势

- 大规模知识图谱的快速表示学习
  - 在线学习、分布式学习
- 融合外部信息的知识表示学习
  - 利用文本、实体和关系的属性等外部信息
  - 建立统一的知识表示空间
- 考虑常识信息的知识表示学习与信息抽取
  - 特定关系的常识信息（如人的结婚年龄、毕业年龄等）
- 知识表示在信息融合、知识推理中的应用
  - 跨语言、跨知识库的知识融合
  - 在低维向量空间中的知识推理

开源工具

- 在中文分词、文本分类、关键词抽取、表示学习等方面开源数十项软件

https://github.com/thunlp

THULAC：	中文词法分析
THUCTC：	中文文本分类
THUTAG：	关键词抽取与社会标签推荐
KB2E：	知识表示学习
NRE：	神经网络关系抽取
NSC：	神经网络情感分类
MMDW：	最大间隔网络表示学习

开源平台：OpenKE

http://openke.thunlp.org/

- **工具包**：统一接口，包括TransE、TransH、TransR、TransD、RESCAL、DistMult、HolE、ComplEx等算法的高效实现
- **表示模型**：面向WikiData和Freebase两大通用KG全量数据的预训练好的知识表示模型下载

开源平台： OpenNE

https://github.com/thunlp/OpenNE

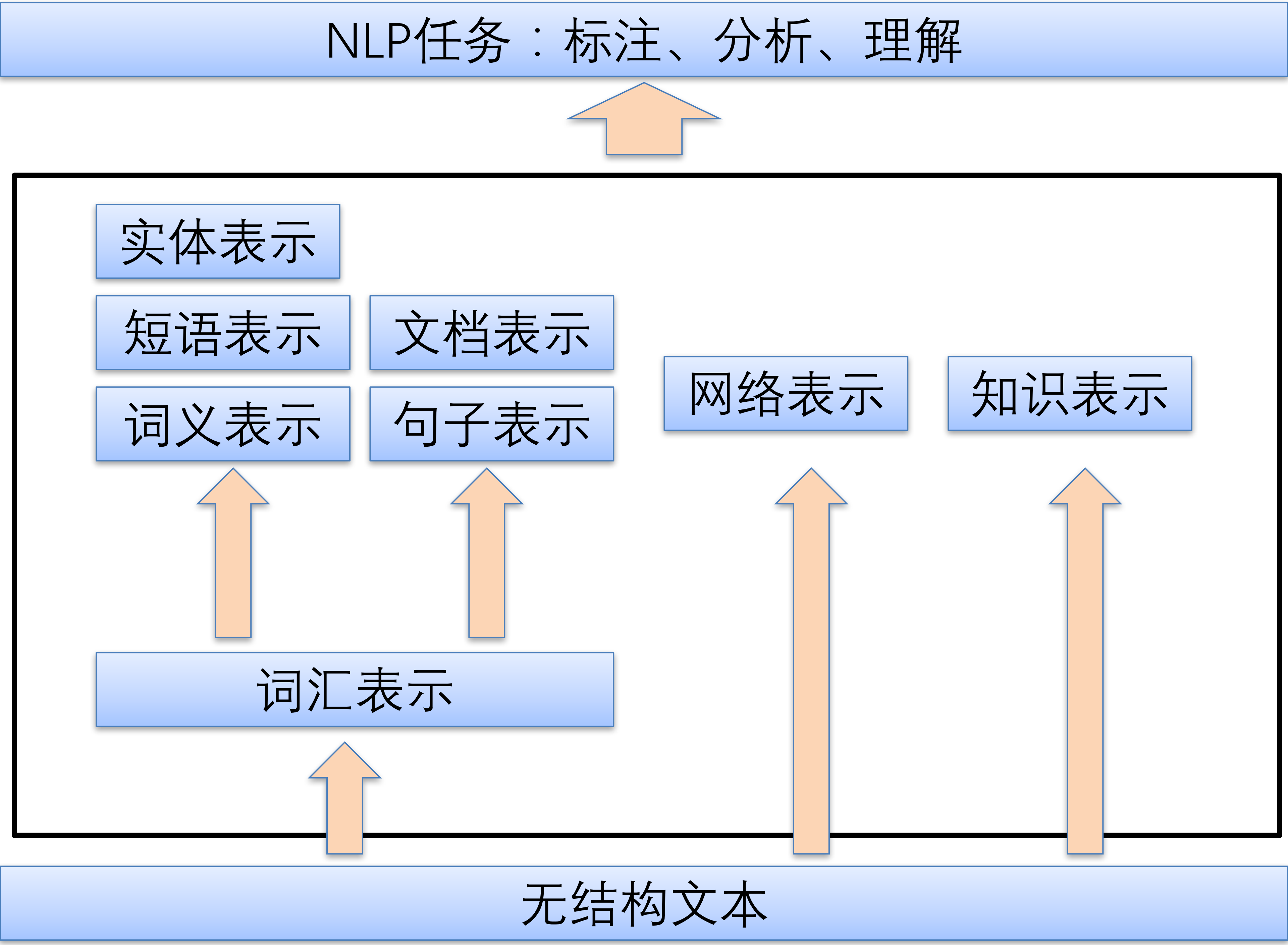
- **工具包**：统一接口，基于TensorFlow实现DeepWalk, LINE, node2vec, GraRep, TADW and GCN等网络表示学习方法
- **论文清单**：及时整理网络表示学习论文清单，https://github.com/thunlp/nrlpapers

开源平台：OpenNRE

https://github.com/thunlp/OpenNRE

- **工具包**：统一接口，基于TensorFlow实现不同的Embedding、Encoder、Selector、Classifier模块，方便用户使用和扩展

- **论文清单**：及时整理网络表示学习论文清单，https://github.com/thunlp/nrepapers

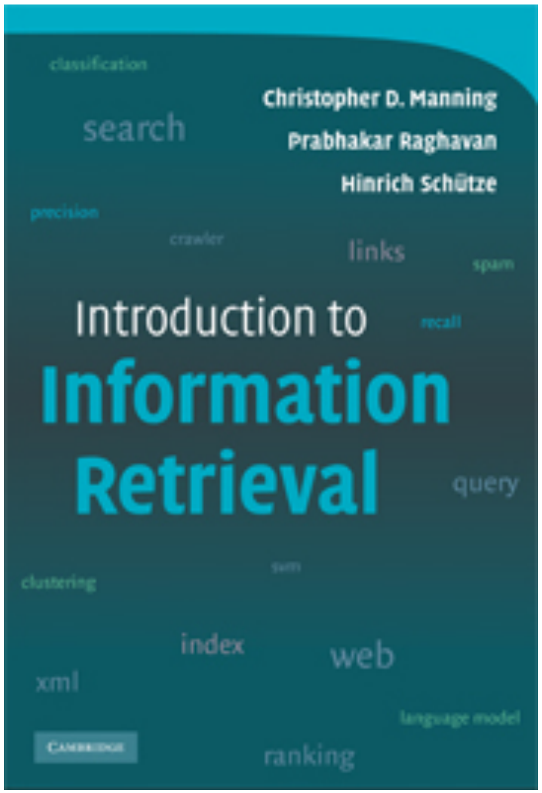

NLP任务：标注、分析、理解
实体表示
短语表示
文档表示
网络表示
知识表示
词义表示
句子表示
词汇表示
无结构文本

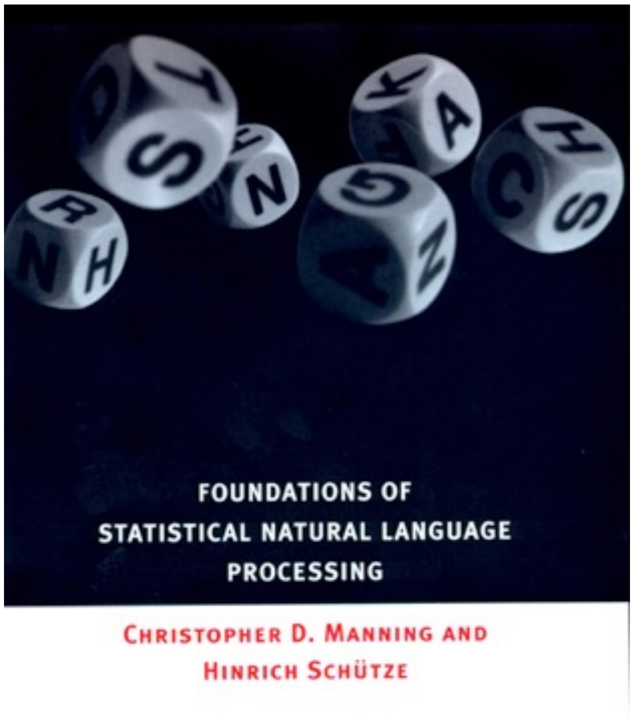
总结

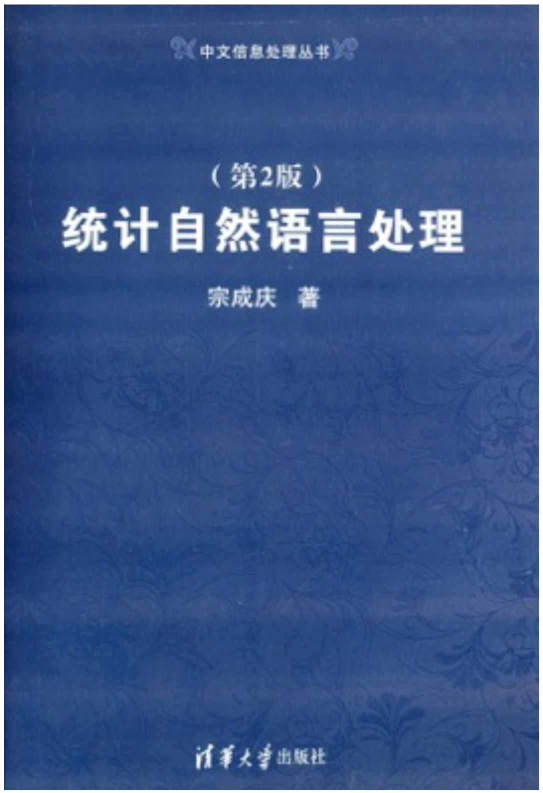
- 分布式表示将研究对象语义信息编码到低维向量空间中
- 分布式表示可扩展性强，可有效解决数据稀疏问题，用于跨领域、跨对象的语义计算和知识迁移
- 分布式表示已被广泛应用于汉字、词汇、词义、实体、短语、句子、文档、网络和知识的表示
- 分布式表示仍有很多开放性问题需要解决
  – 如何在深度学习框架引入人类先验知识（如人工标注的语言和世界知识库）
  – 比向量更具表达能力的表示形式
  – 对复杂结构信息（如层次树、网络等）的表示学习框架

推荐书目

- 自然语言处理


classification
search
Christopher D. Manning
Prabhakar Raghavan
Hinrich Schütze
precision
crawler
links
spam
Introduction to
recall
Information
Retrieval
query
clustering
index
web
xml
language model
ranking


FOUNDATIONS OF
STATISTICAL NATURAL LANGUAGE
PROCESSING
CHRISTOPHER D. MANNING AND
HINRICH SCHÜTZE

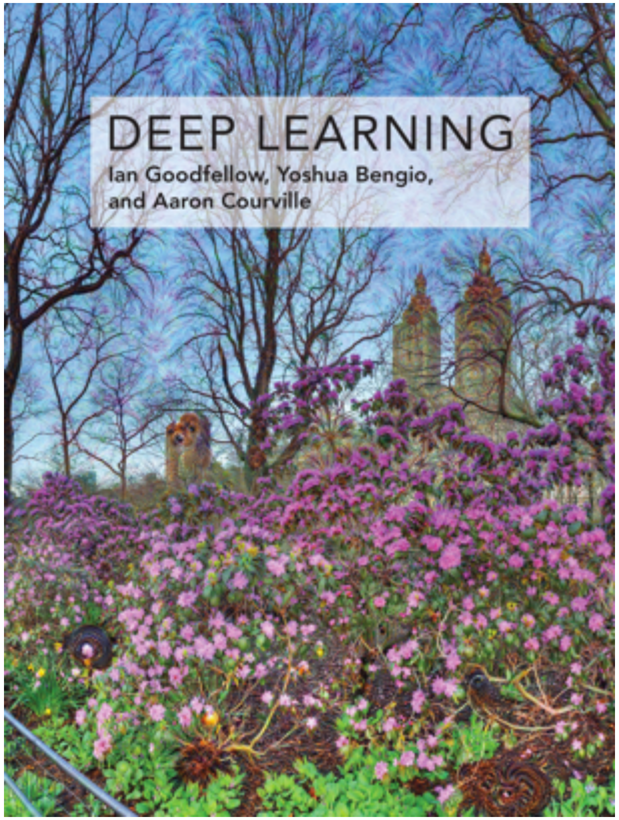

中文信息处理丛书
（第2版）
统计自然语言处理
宗成庆 著
清华大学出版社

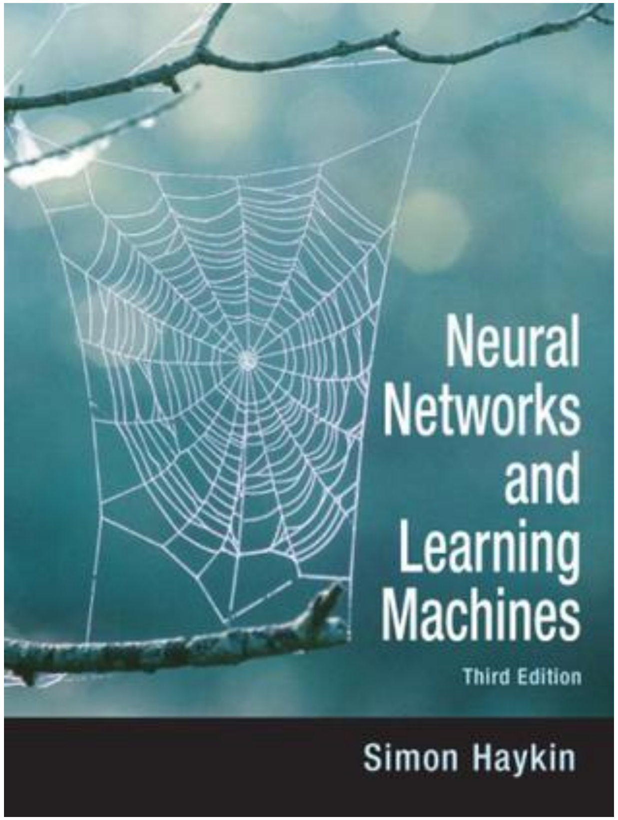
推荐书目

- 机器学习

推荐书目

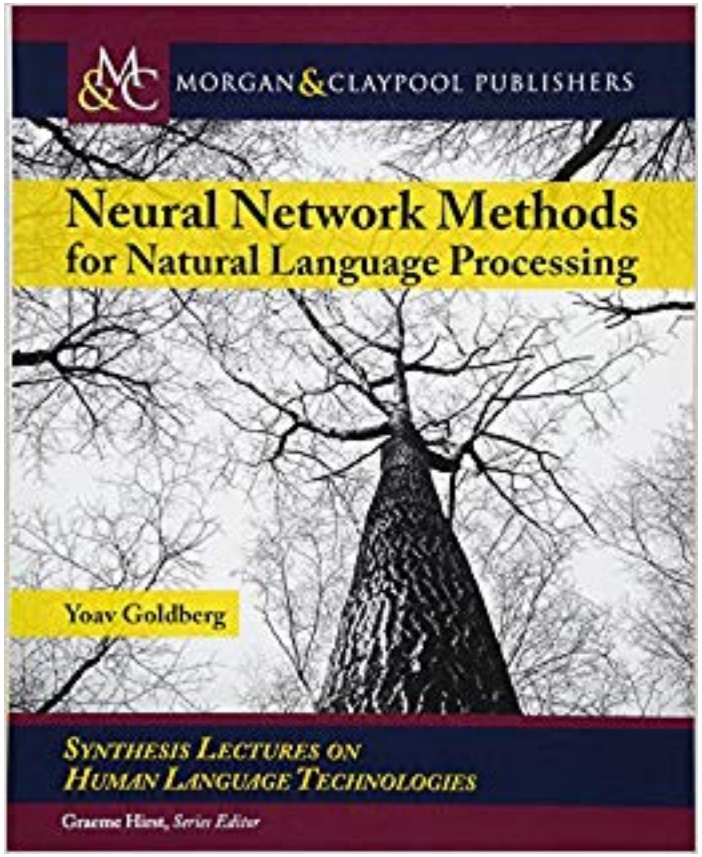
- 深度学习

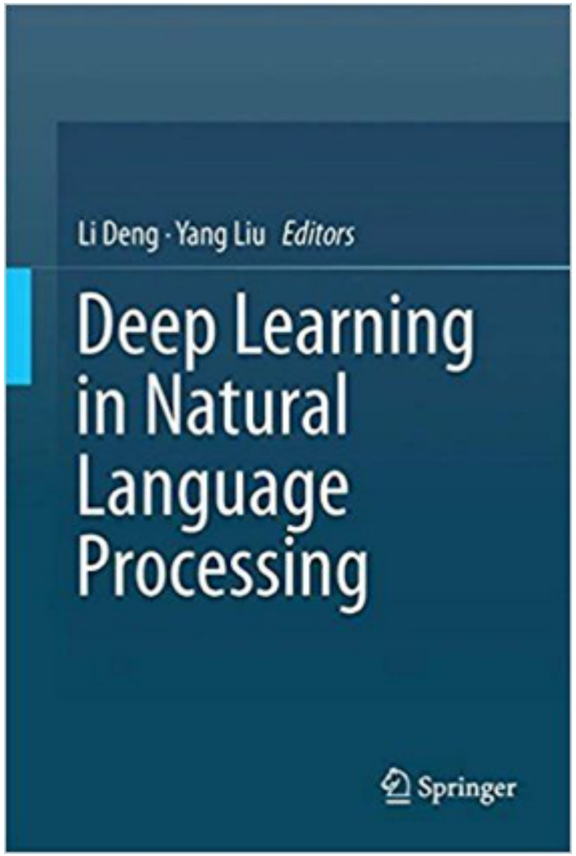

DEEP LEARNING
Ian Goodfellow, Yoshua Bengio,
and Aaron Courville


Neural Networks and Learning Machines
Third Edition
Simon Haykin

推荐书目

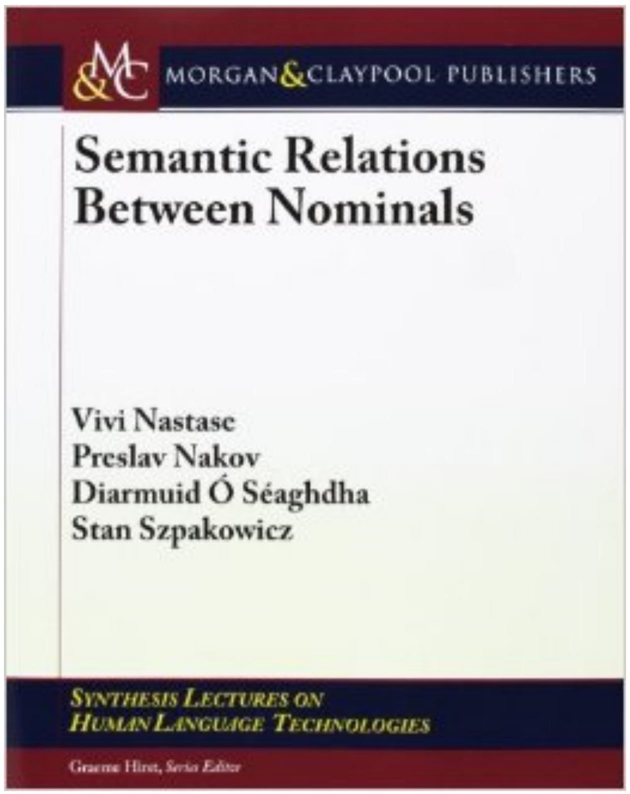
- 深度学习+自然语言处理

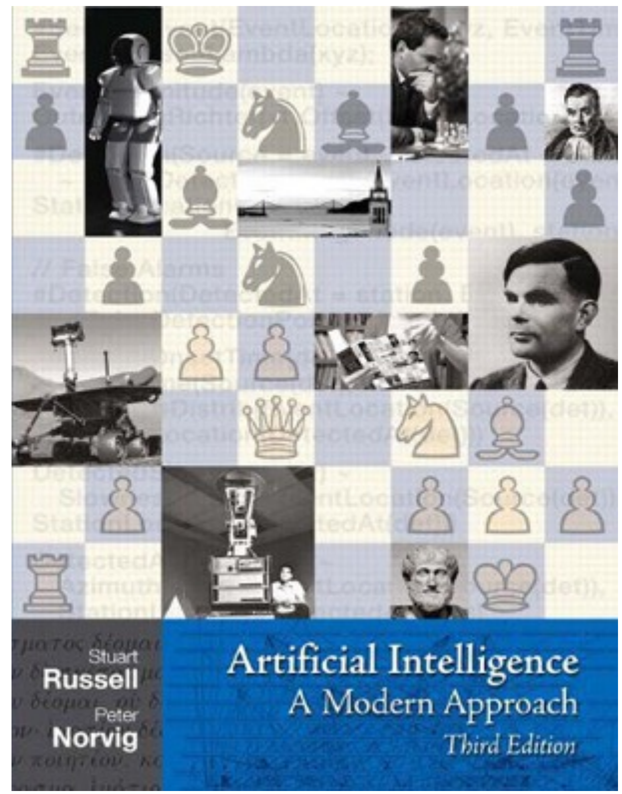




推荐书目

- 知识图谱


MORGAN&CLAYPOOL PUBLISHERS
Semantic Relations
Between Nominals
Vivi Nastase
Preslav Nakov
Diarmuid Ó Séaghdha
Stan Szpakowicz
SYNTHESIS LECTURES ON
HUMAN LANGUAGE TECHNOLOGIES
Graeme Hirst, Series Editor


Stuart
Russell
Peter
Norvig
Artificial Intelligence
A Modern Approach
Third Edition

谢谢大家，欢迎交流

liuzy@tsinghua.edu.cn

http://nlp.csai.tsinghua.edu.cn/~lzy

https://github.com/thunlp