自然语言处理国际前沿动态综述报告会
CCL 2016

知 识 图 谱
Knowledge Graphs

刘知远
清华大学
liuzy@tsinghua.edu.cn


NOBEL PRIZE

结构化知识

write

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

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

典型知识图谱

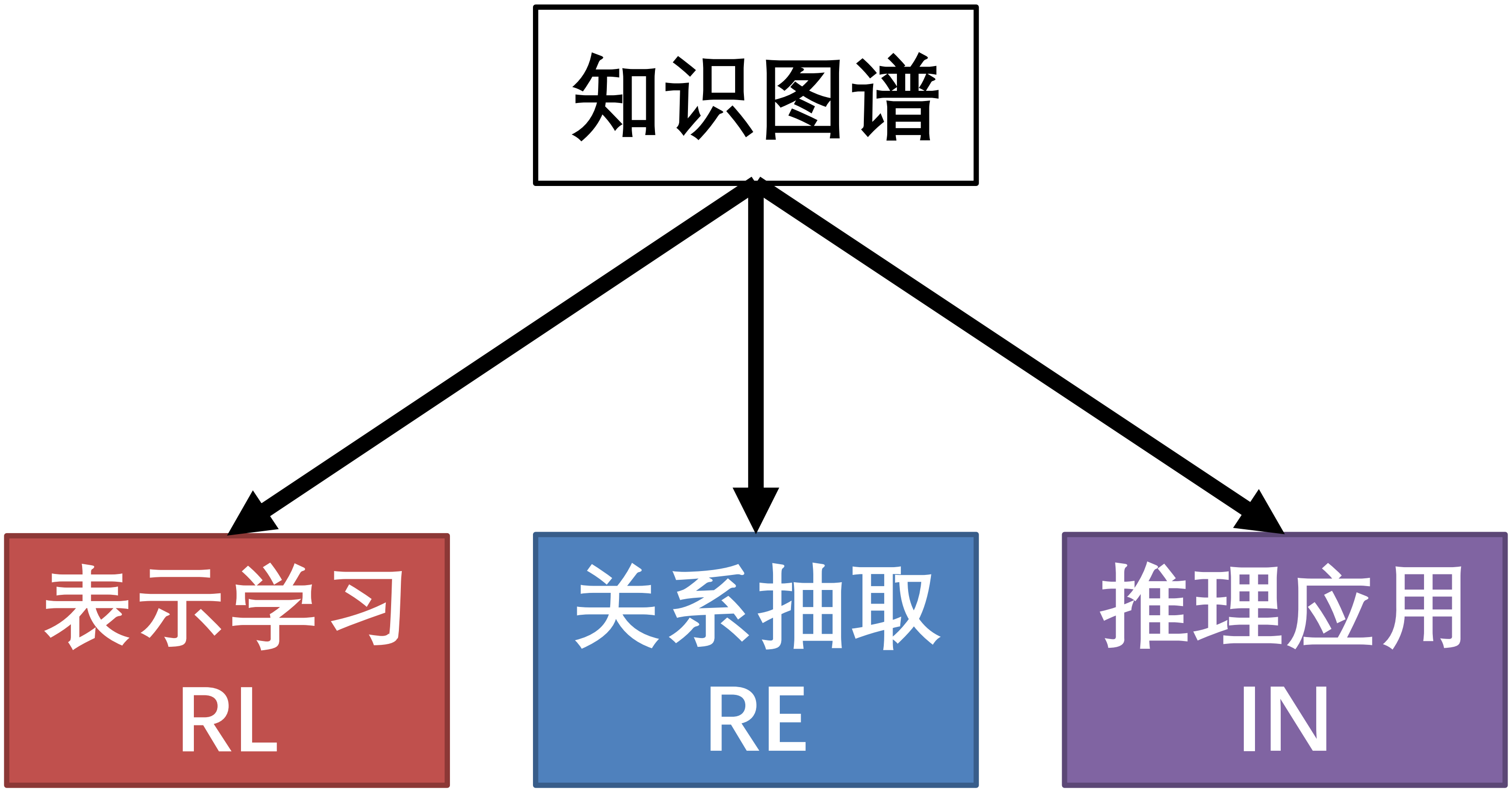
Freebase™

DBpedia

yago
select knowledge

WordNet
A lexical database for English

知识图谱在2016年



AAAI 2016

- **(RL)** Knowledge Graph Completion with Adaptive Sparse Transfer Matrix
- **(RL)** Locally Adaptive Translation for Knowledge Graph Embedding
- **(RL)** Holographic Embeddings of Knowledge Graphs
- **(RL)** Representation Learning of Knowledge Graphs with Entity Descriptions
- **(RE)** PEAK: Pyramid Evaluation via Automated Knowledge Extraction
- **(RE)** Numerical Relation Extraction with Minimal Supervision
- **(RE)** Global Distant Supervision for Relation Extraction
- **(RE)** Aggregating Inter-Sentence Information to Enhance Relation Extraction
- Column-Oriented Datalog Materialization for Large Knowledge Graphs

IJCAI 2016

- **(RL)** Multi-Modal Bayesian Embeddings for Learning Social Knowledge Graphs
- **(RL)** From One Point to A Manifold: Knowledge Graph Embedding For Precise Link Prediction
- **(RL)** Knowledge Representation Learning with Entities, Attributes and Relations
- **(RL)** Text-enhanced Representation Learning for Knowledge Graph
- **(RL)** Representation Learning of Knowledge Graphs with Hierarchical Types
- **(RL)** On the Representation and Embedding of Knowledge Bases Beyond Binary Relations
- **(RL)** KOGNAC: Efficient Encoding of Large Knowledge Graphs
- **(RE)** Building Joint Spaces for Relation Extraction

NAACL 2016

- **(RL)** STransE: a novel embedding model of entities and relationships in knowledge bases
- **(RL)** A Translation-Based Knowledge Graph Embedding Preserving Logical Property of Relations
- **(RE)** Combining Recurrent and Convolutional Neural Networks for Relation Classification
- **(RE)** Multilingual Relation Extraction using Compositional Universal Schema
- **(RE)** Effective Crowd Annotation for Relation Extraction

ACL 2016

- **(RL)** Leveraging Lexical Resources for Learning Entity Embeddings in Multi-Relational Data
- **(RL)** Compositional Learning of Embeddings for Relation Paths in Knowledge Base and Text
- **(RL)** TransG : A Generative Model for Knowledge Graph Embedding
- **(RE)** Commonsense Knowledge Base Completion
- **(RE)** Knowledge Base Completion via Coupled Path Ranking
- **(RE)** Bidirectional Recurrent Convolutional Neural Network for Relation Classification
- **(RE)** Investigating LSTMs for Joint Extraction of Opinion Entities and Relations

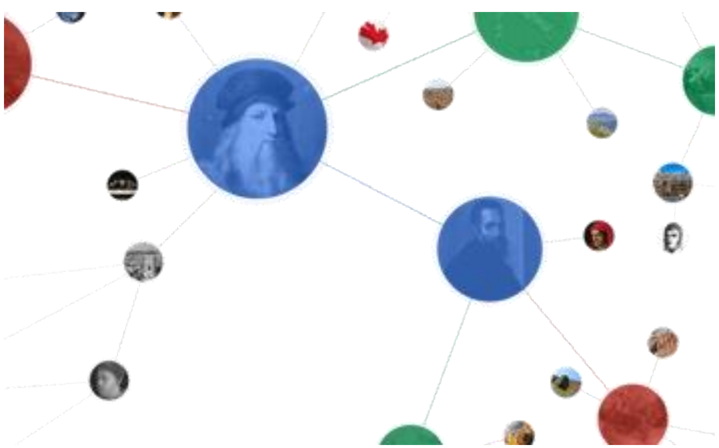
ACL 2016

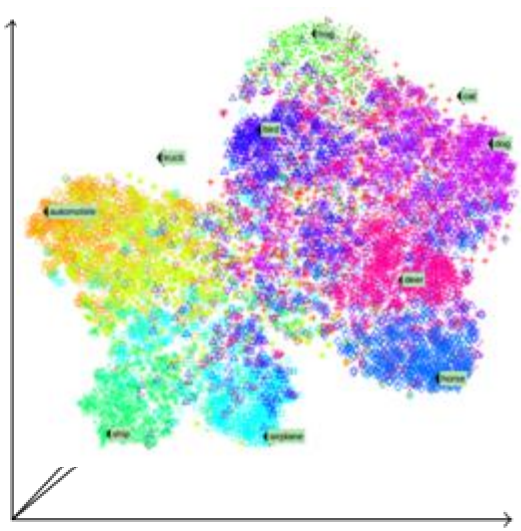
- **(RE)** End-to-End Relation Extraction using LSTMs on Sequences and Tree Structures
- **(RE)** Relation Classification via Multi-Level Attention CNNs
- **(RE)** Neural Relation Extraction with Selective Attention over Instances
- **(RE)** Question Answering on Freebase via Relation Extraction and Textual Evidence
- **(RE)** Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification
- **(RE)** JEDI: Joint Entity and Relation Detection using Type Inference
- Visualizing and Curating Knowledge Graphs over Time and Space

表示学习

知识表示与表示学习

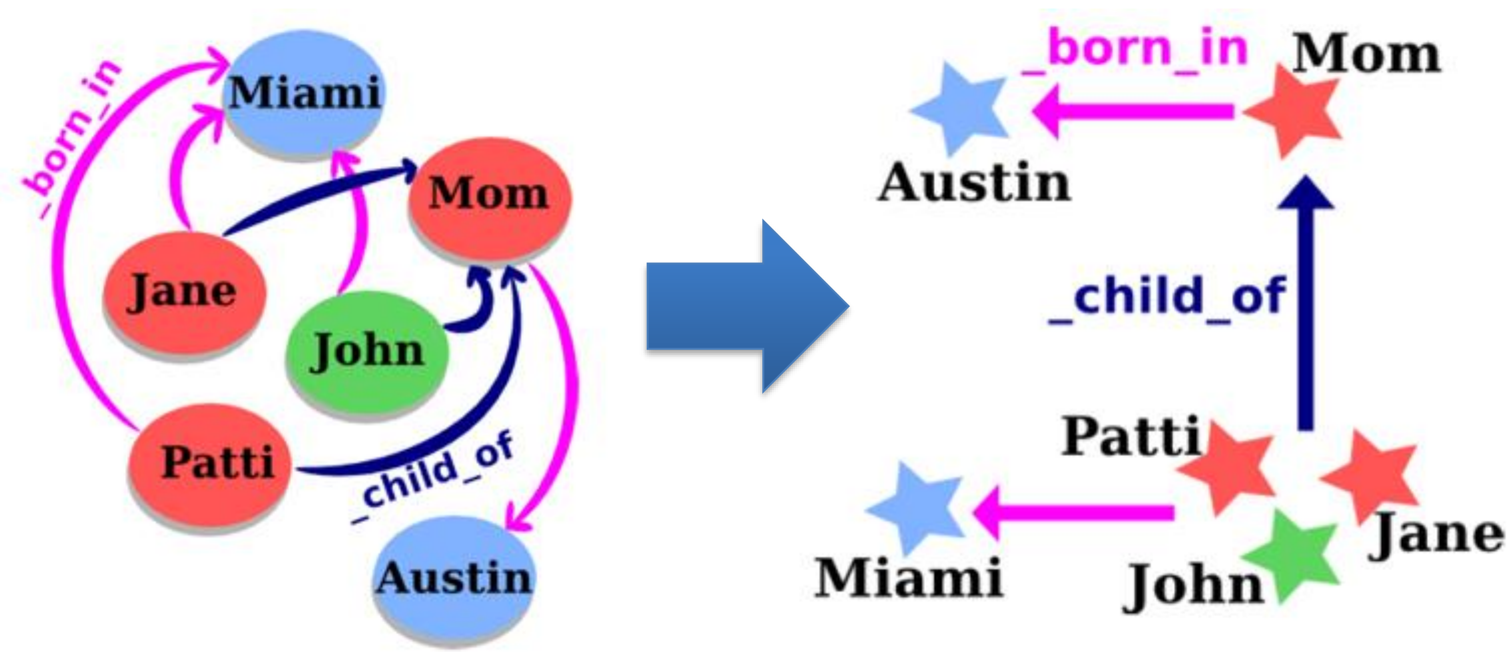
- 知识图谱的典型表示方案
  - 基于符号表示的三元组（RDF）
  - 无法有效计算实体间的语义关系
- 解决方案：将知识映射到低维向量空间





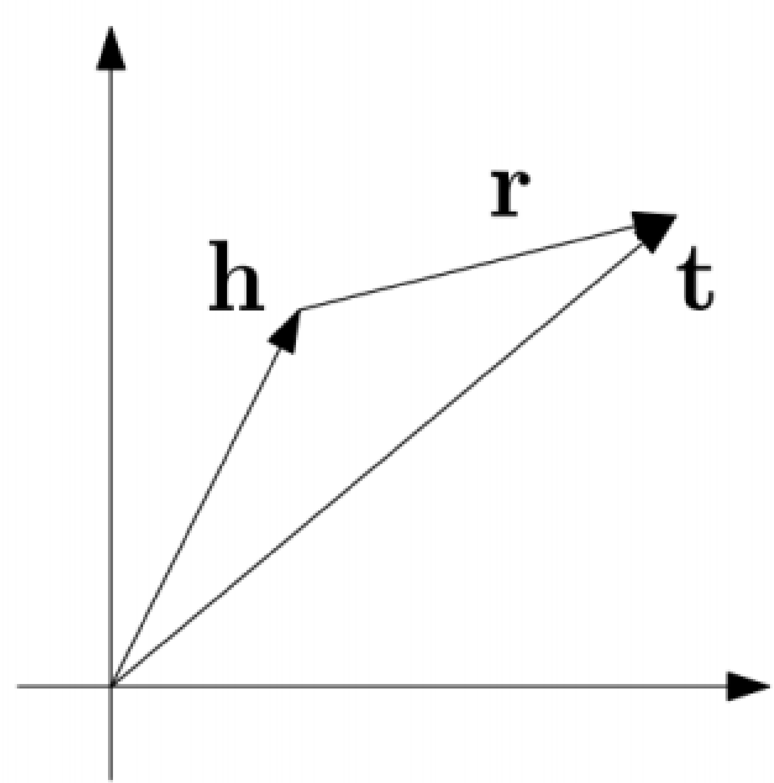
TransE: 将关系表示为翻译

- 对每个事实 (head, relation, tail)，将其中的relation作为从head到tail的翻译操作



TransE: 将关系表示为翻译

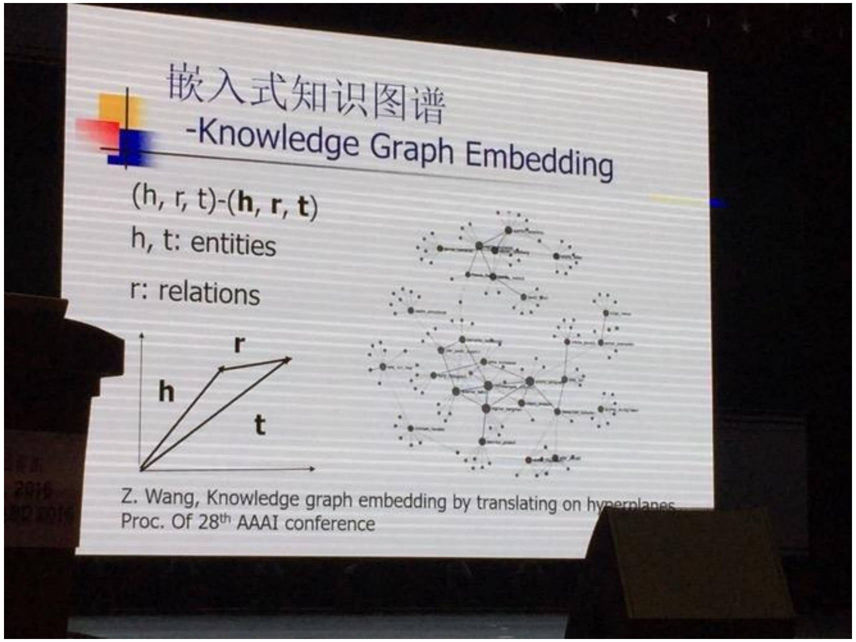
- 对每个事实 (head, relation, tail)，将relation作为从head到tail的翻译操作



优化目标: **h + r = t**

知识表示与表示学习

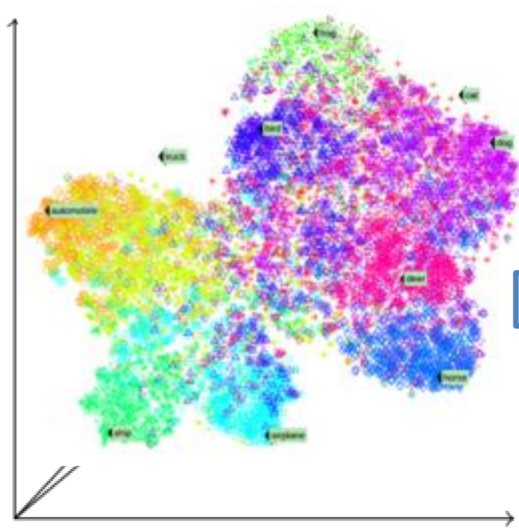

后深度学习时代的计算语言学
Computational Linguistics under the
Post-Deep Learning Age
清华大学计算机系 张钹


嵌入式知识图谱
-Knowledge Graph Embedding
(h, r, t)-(**h**, **r**, **t**)
h, t: entities
r: relations
r
h
t
Z. Wang, Knowledge graph embedding by translating on hyperplanes
Proc. Of 28th AAAI conference

知识表示学习的意义

- 融合知识与文本，建立统一语义空间
- 实现知识驱动的自然语言处理

知识
文档
句子
词汇

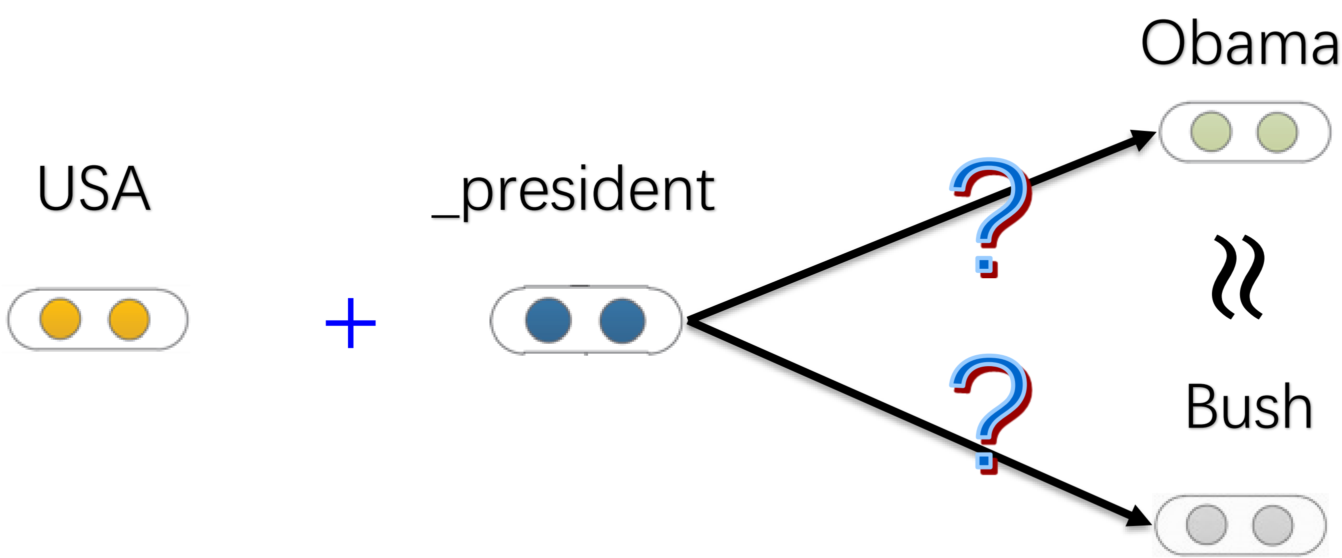


统一语义空间

语义分析
句法分析
词法分析

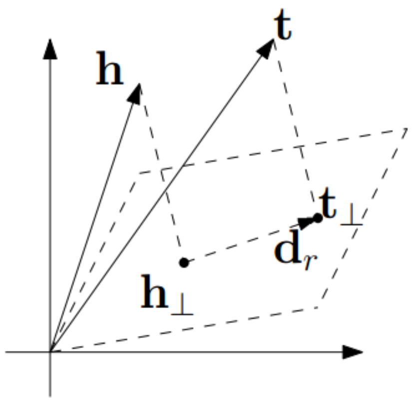
知识表示学习热点一

- 面向复杂关系的建模：1-to-N, N-to-1, N-to-N
  - (USA, _president, Obama)
  - (USA, _president, Bush)

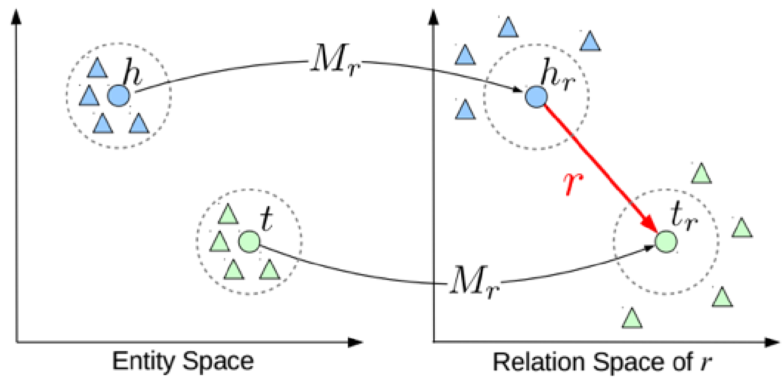


知识表示学习热点一

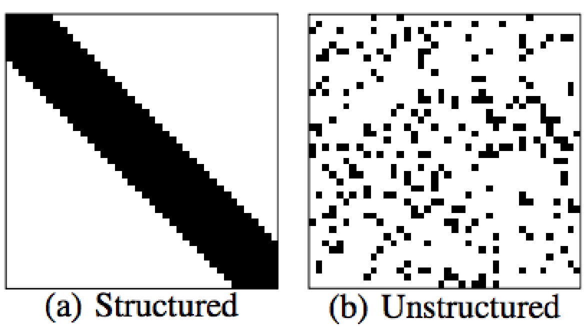
- 建立与特定关系有关的实体表示
  - TransA， TransD， TransE， TransG， TransH， TransR, KG2E, TranSparse, Hole



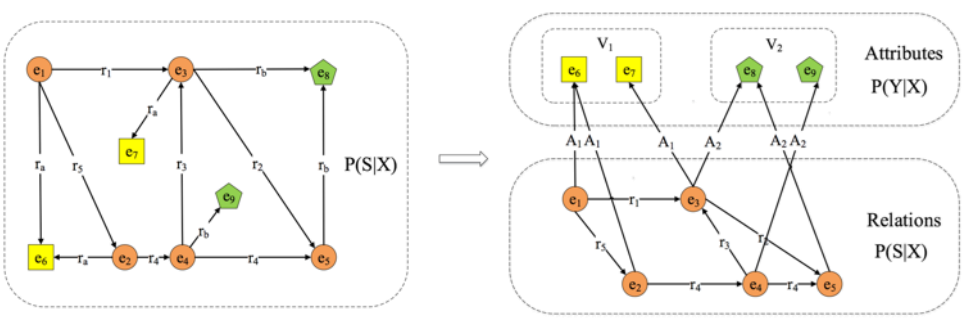
TransH (AAAI 2014)



TransR (AAAI 2015)



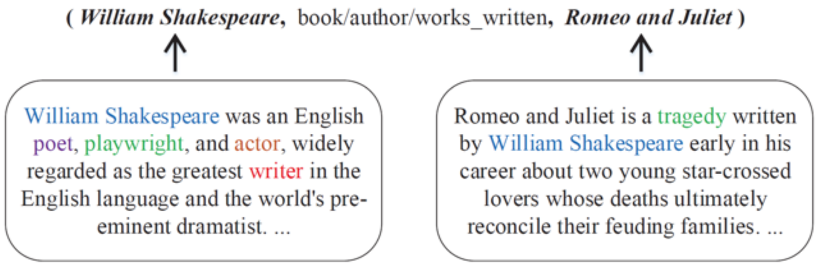
TransSparse (AAAI 2016)

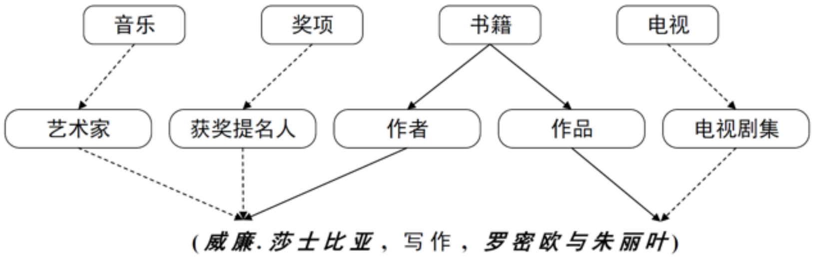


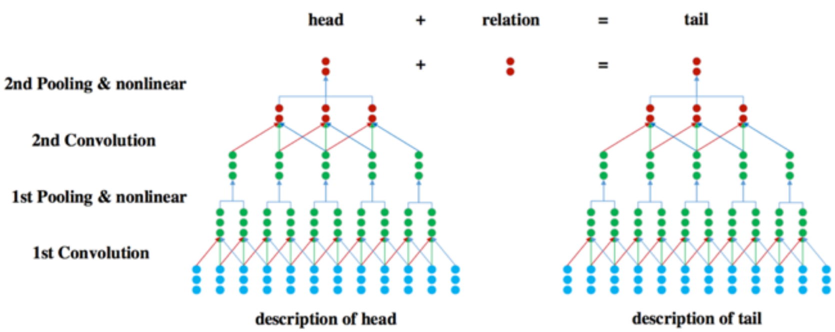
KR-EAR (IJCAI 2016)

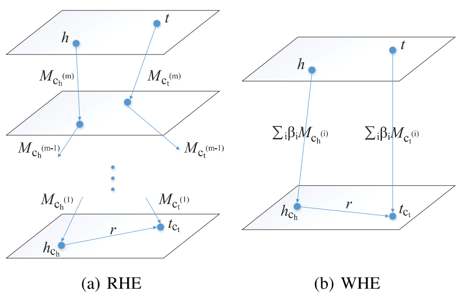
知识表示学习热点二

- 利用实体的描述、类别信息提供丰富语义信息









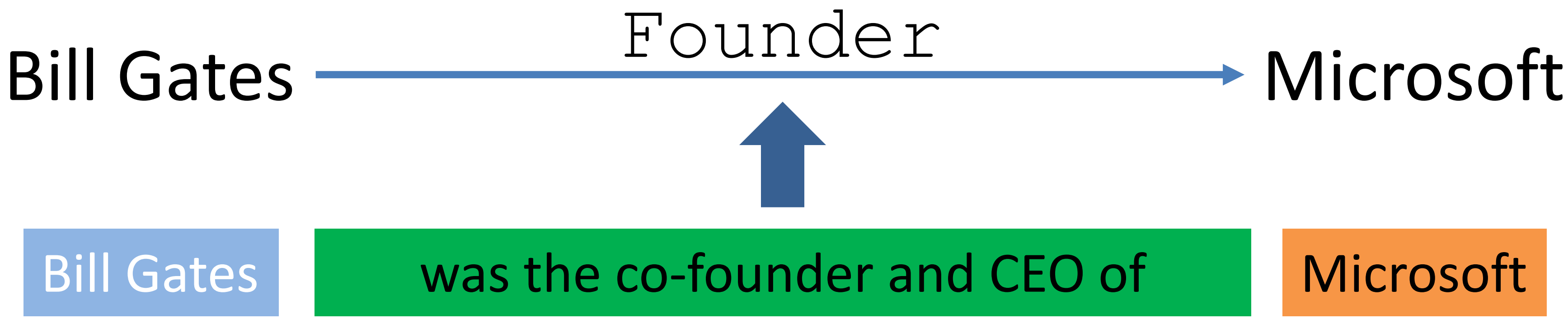
DKRL (AAAI 2016)

TKRL (IJCAI 2016)

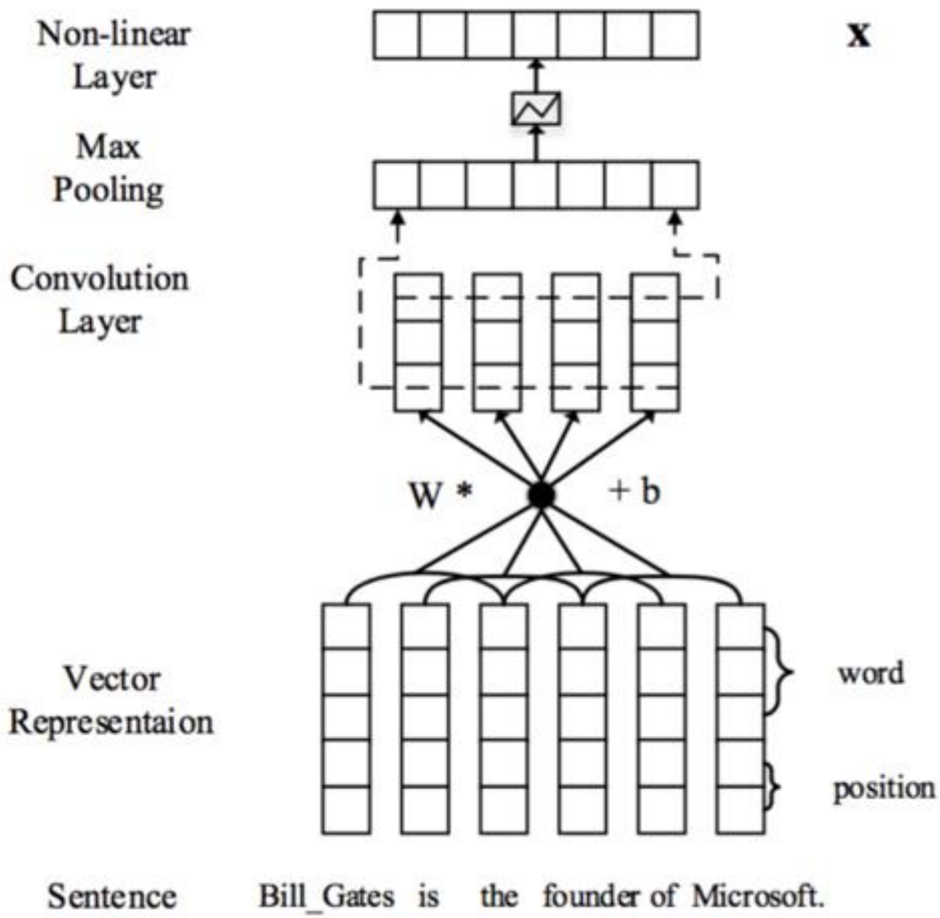
关系抽取

关系分类

- 基于文本信息的关系分类



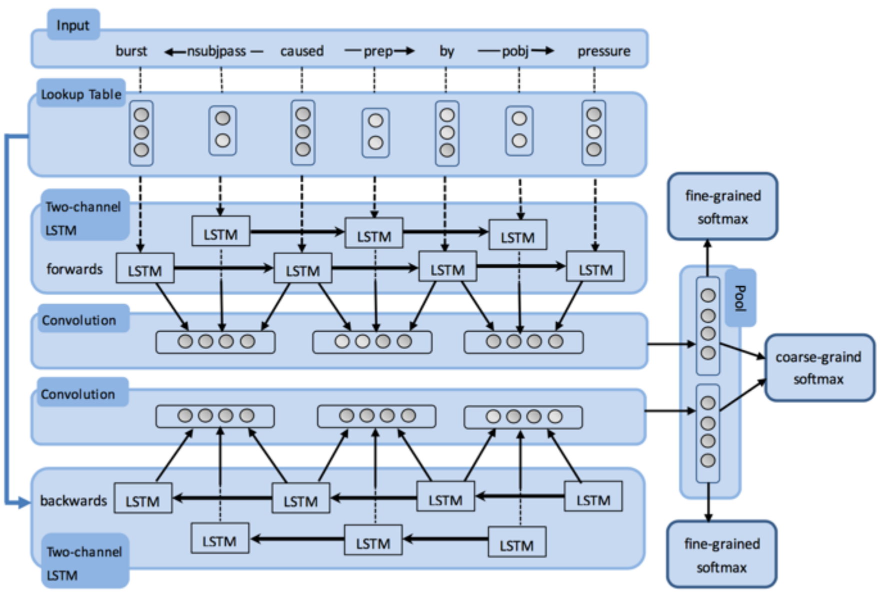
神经网络关系分类模型



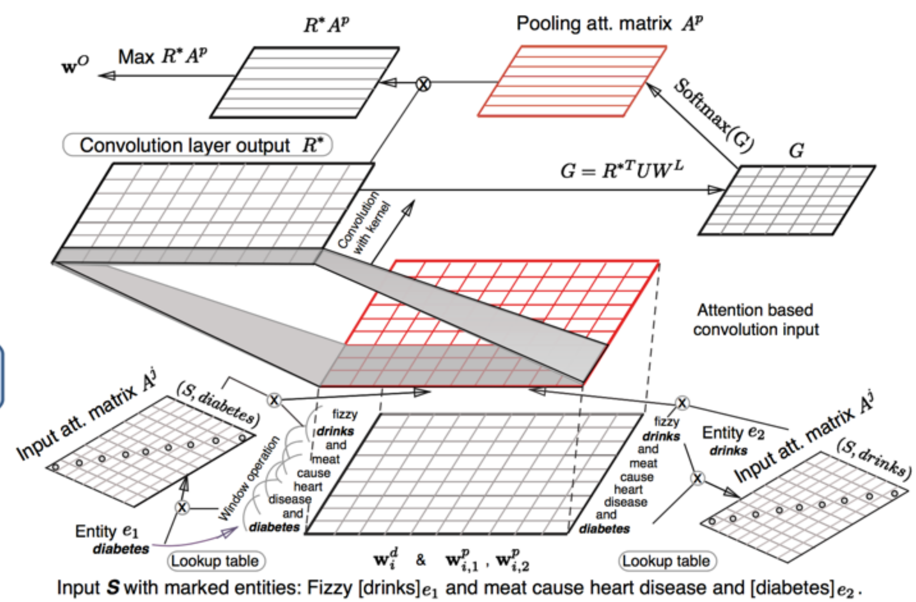
Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. Relation classification via convolutional deep neural network. In Proceedings of COLING 2014.
Daojian Zeng,Kang Liu,Yubo Chen,and Jun Zhao. Distant supervision for relation extraction via piecewise convolutional neural networks. In Proceedings of EMNLP 2015.

关系分类热点

- 提出各种新颖的神经网络模型
  - CNN, RNN, LSTM, …
  - Bi-directional, tree-structure, attention, …



BRCNN (ACL 2016)

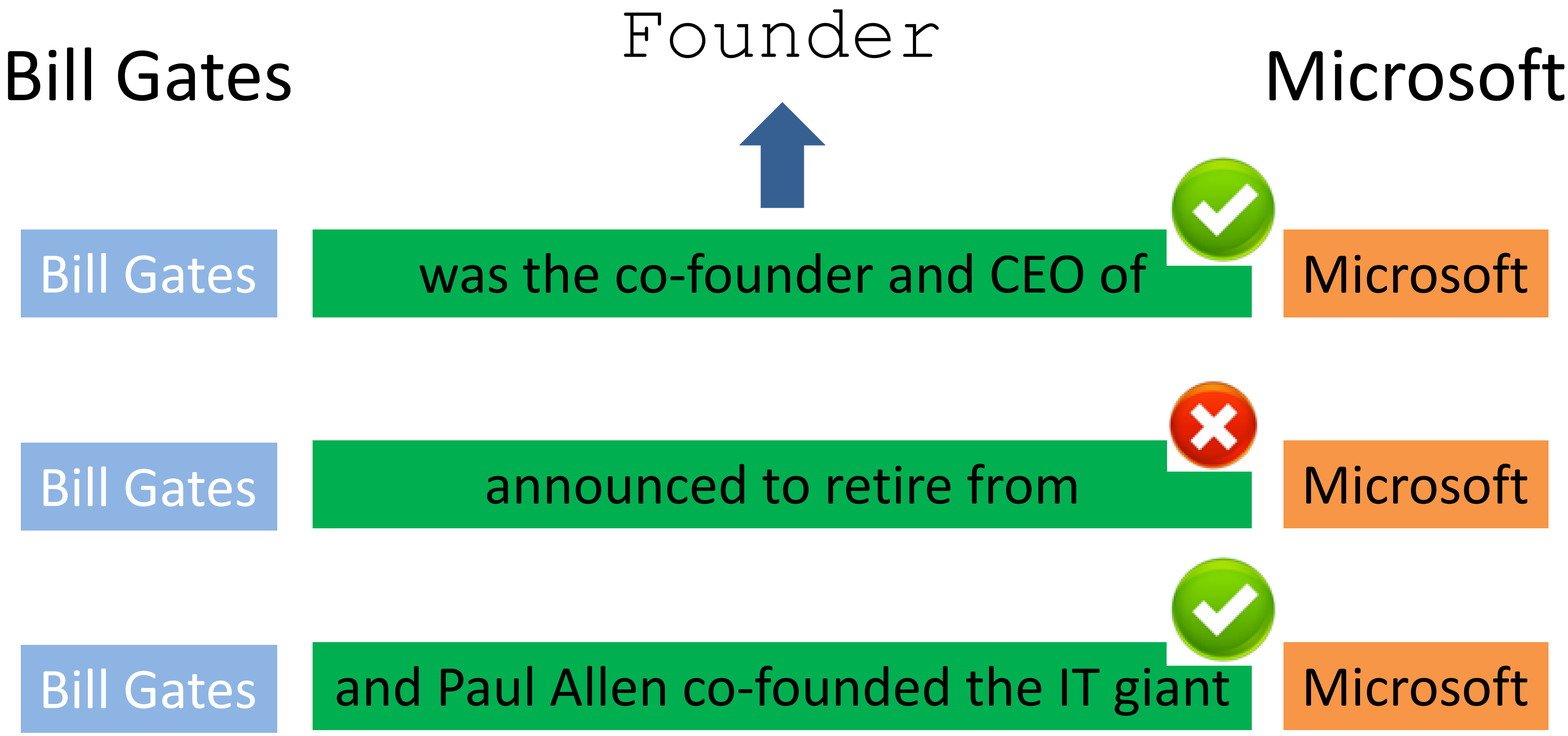


BRCNN (ACL 2016)

关系分类热点

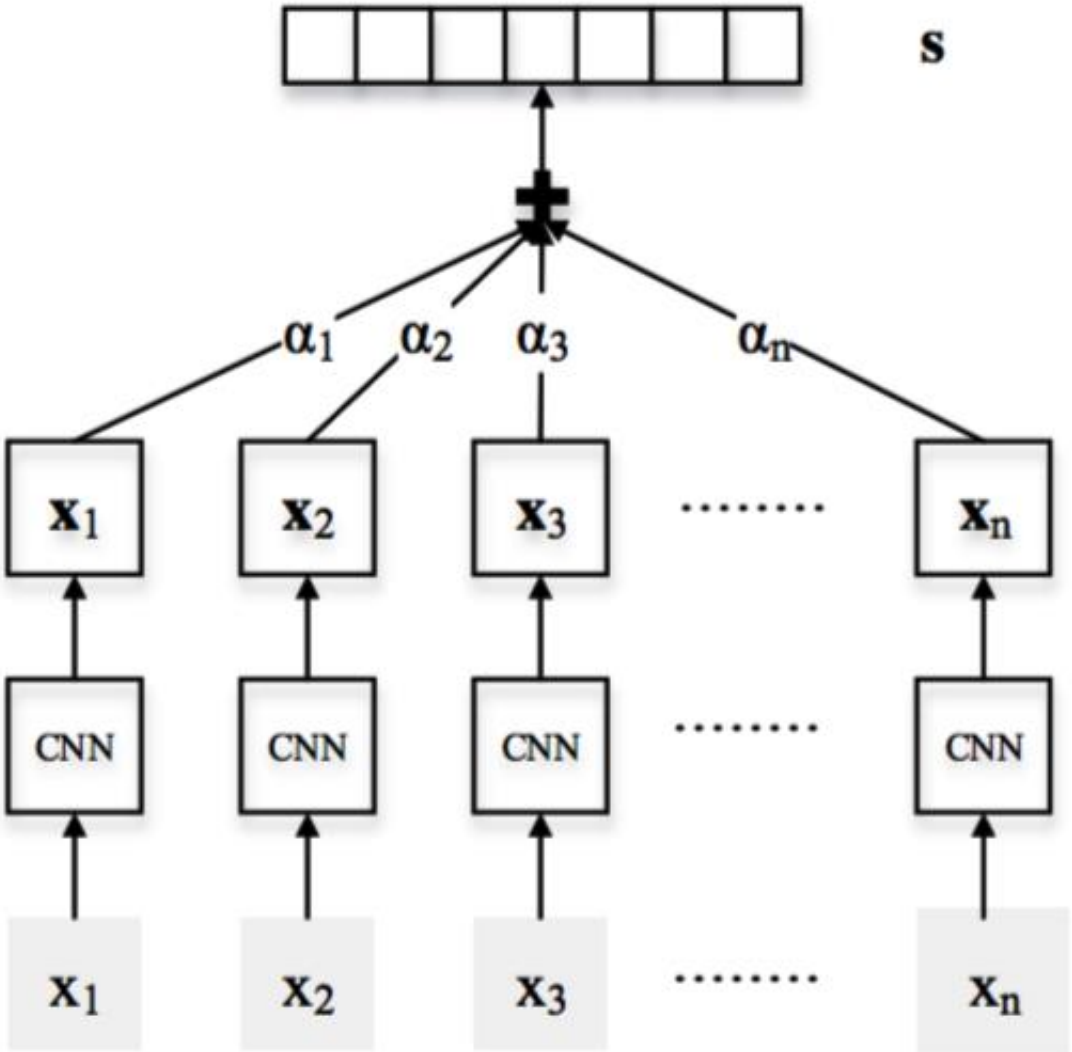
Classifier	**F1**
Manually Engineered Methods	
SVM (Rink and Harabagiu, 2010)	82.2
Dependency Methods	
RNN (Socher et al., 2012)	77.6
MVRNN (Socher et al., 2012)	82.4
FCM (Yu et al., 2014)	83.0
Hybrid FCM (Yu et al., 2014)	83.4
SDP-LSTM (Xu et al., 2015b)	83.7
DRNNs (Xu et al., 2016)	85.8
SPTree (Miwa and Bansal, 2016)	84.5
End-To-End Methods	
CNN+ Softmax (Zeng et al., 2014)	82.7
CR-CNN (dos Santos et al., 2015)	84.1
DepNN (Liu et al., 2015)	83.6
depLCNN+NS (Xu et al., 2015a)	85.6
STACK-FORWARD∗	83.4
VOTE-BIDIRECT∗	84.1
VOTE-BACKWARD∗	84.1
Our Architectures	
Att-Input-CNN	87.5
Att-Pooling-CNN	**88.0**

远程监督与多实例问题



Bill Gates
Founder
Microsoft
Bill Gates
was the co-founder and CEO of
Microsoft
Bill Gates
announced to retire from
Microsoft
Bill Gates
and Paul Allen co-founded the IT giant
Microsoft

Sentence-Level Attention

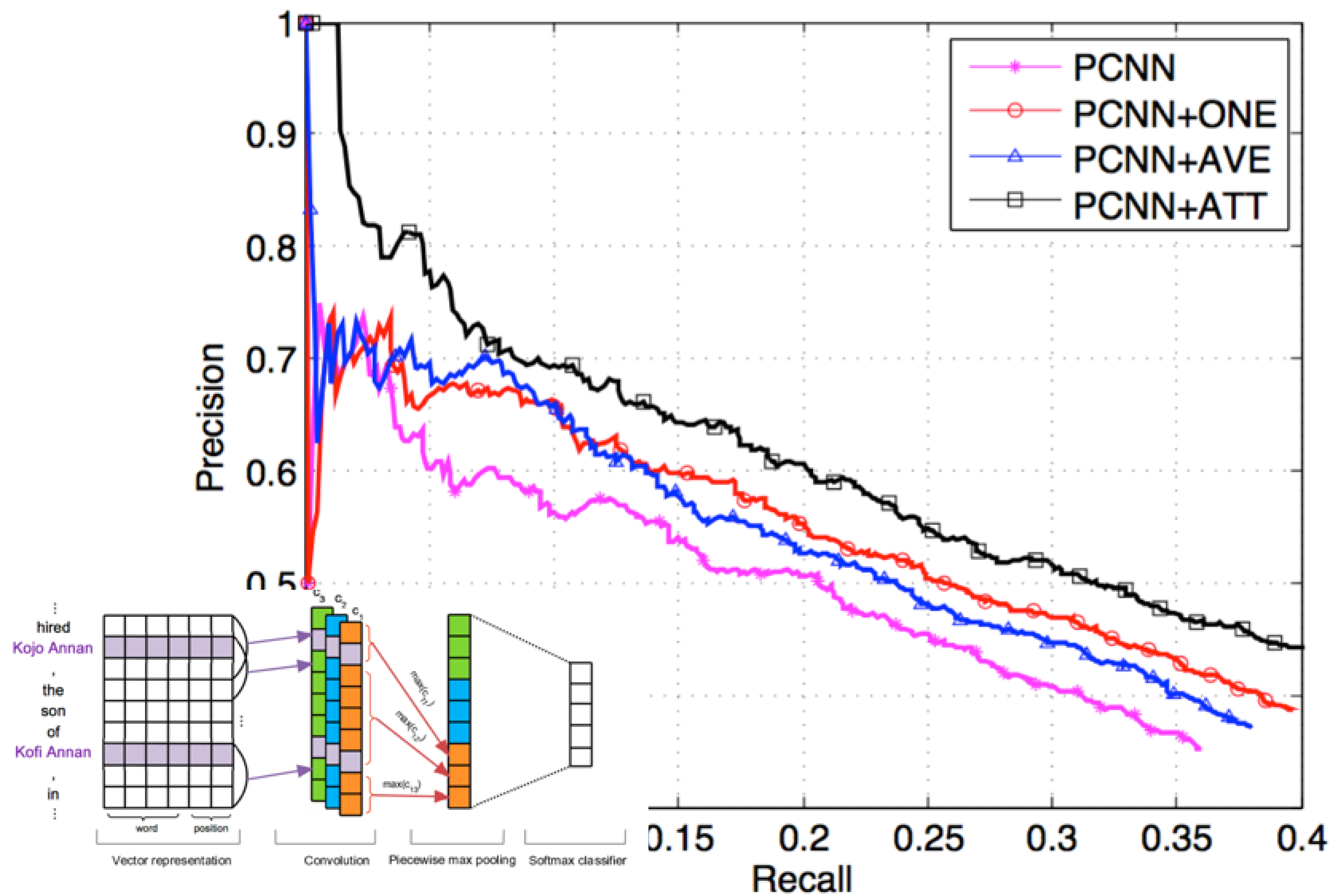


$$\alpha_i = \frac{\exp(e_i)}{\sum_k \exp(e_k)}$$

$$e_i = \mathbf{x}_i \mathbf{A} \mathbf{r}$$

Lin, et al. (2016). Neural Relation Extraction with Selective Attention over Instances. ACL.

关系抽取评测结果



Zeng, et al. (2015). Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks. EMNLP.

小结

- 知识表示学习、关系抽取是近期的研究热点

- 表示学习研究趋势预测
  - 考虑更丰富外部信息（脑电信号、跨语言、…）
  - 融合先验模式（逻辑规则）的表示学习
  - 大规模KG的高效表示学习

- 关系抽取研究趋势预测
  - 应用更多深度学习最新技术（MN、NTM、VAE、…）
  - 通过表示学习融合KG信息
  - 利用更丰富的无监督数据
  - 融合不同关系抽取方案（OpenIE、…）

Thanks!

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

liuzy@tsinghua.edu.cn