

语义知识获取及推理规则发现

韩先培 (hanxianpei@qq.com) 孙乐 (sunle@iscas.ac.cn)

Xianpei Han and Le Sun. *Distant Supervision via Prototype-based Global Representation Learning*. In: the Thirty-First AAAI Conference (AAAI 2017).Xianpei Han and Le Sun. *Context-Sensitive Inference Rule Discovery: A Graph-based Method*. In: the 26th International Conference on Computational Linguistics. (COLING 2016)

海量文本

唐朝中央的三省中书、
门下和尚书，分别负责
决策、审议和执行。
三省的长官都是宰相，
相权分散。

语义知识
获取

ISA(中书省, 机构)
ISA(门下省, 机构)
ISA(尚书省, 机构)
...

属性

中书省
长官: 宰相
职能: 决策
时代: 隋朝到元朝

实体集

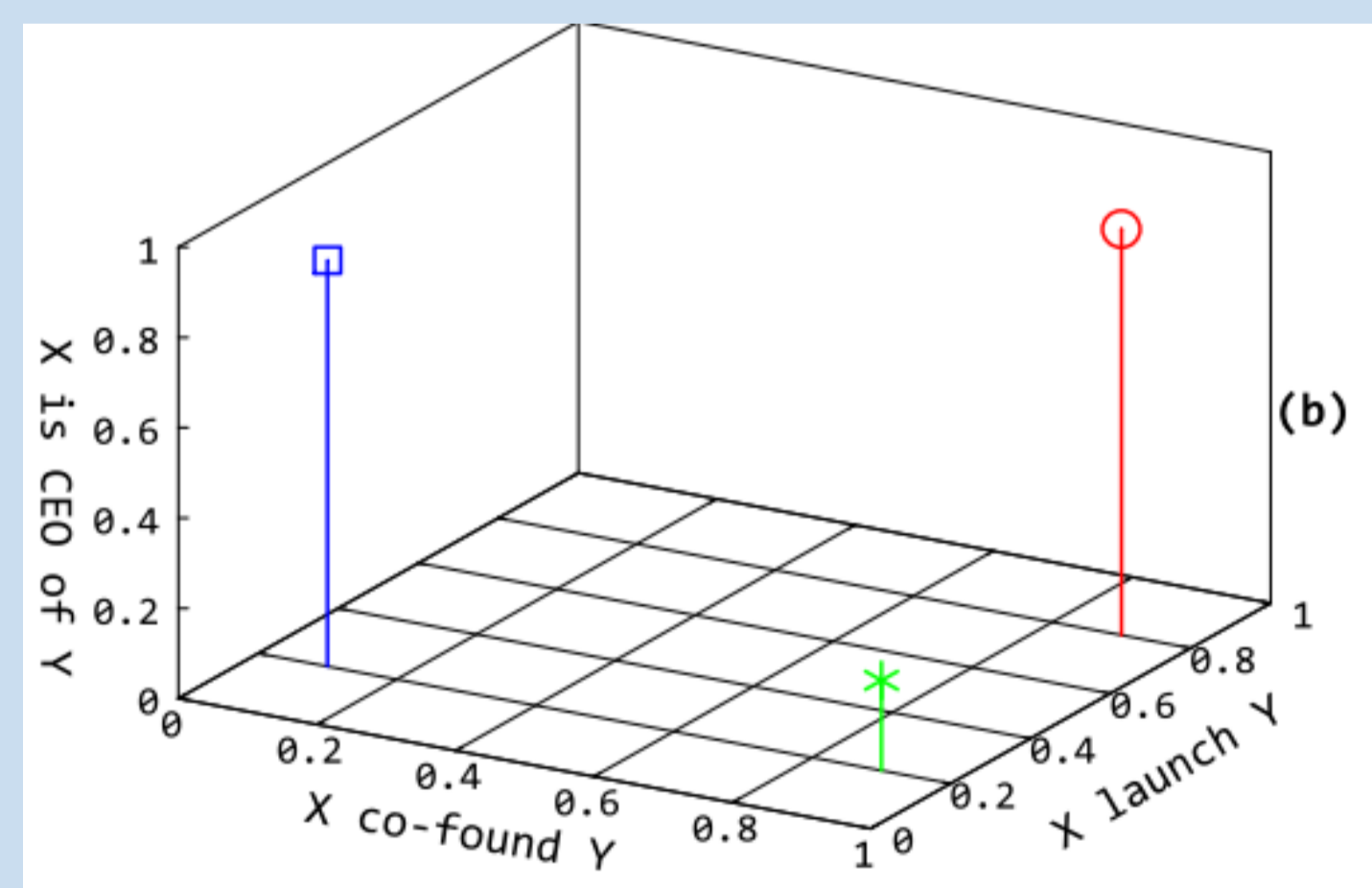
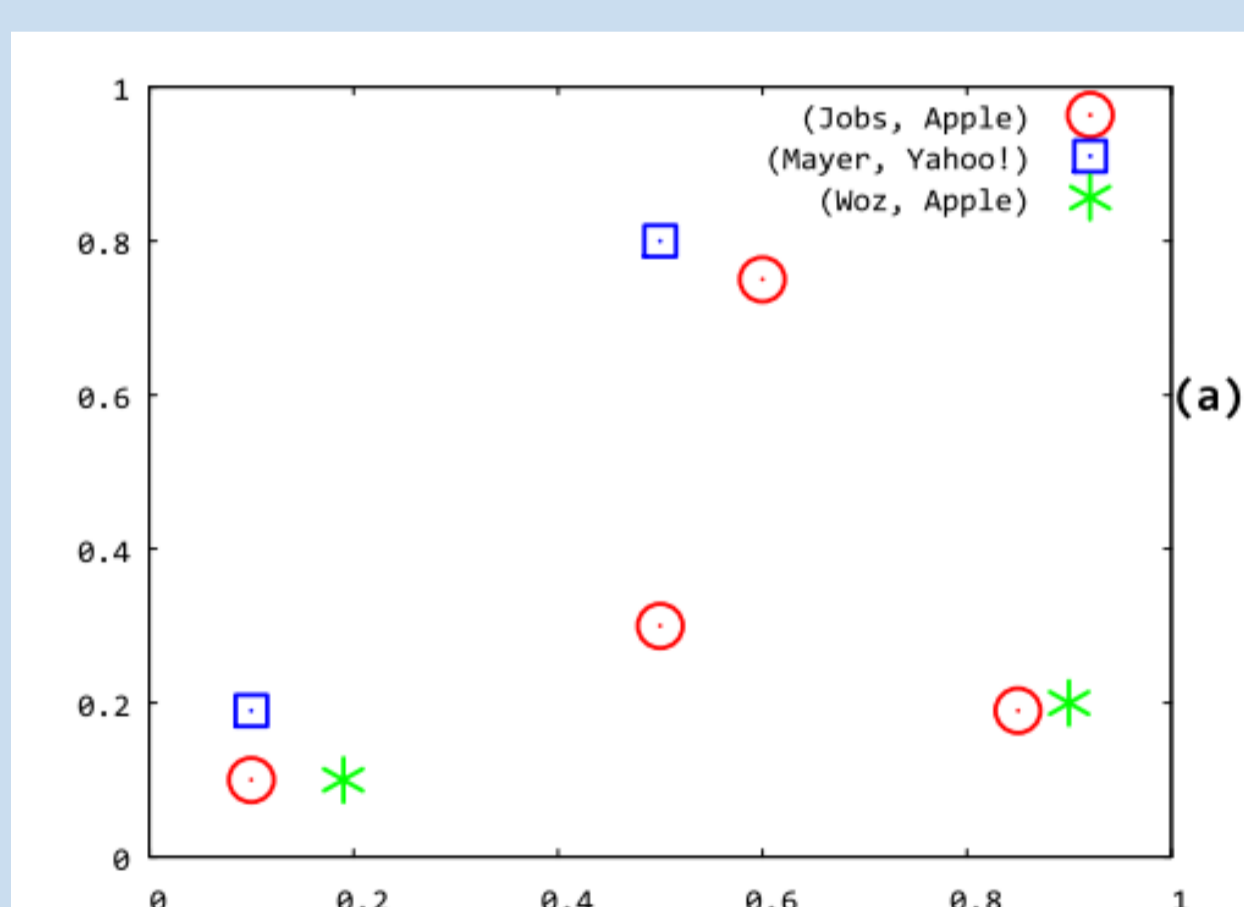
朝代{唐, 宋, 元, 明, 清...}
政治制度{门下省分封制, 宗法制, 郡县制...}

基于深度学习的全局表示学习

创新点

基于深度神经网络，将包含多个句子的实体对表示嵌入到一个原型空间中

- Prototype Learning: 对每一个关系，学习其原型表示；
- Prototype Embedding: 将实体对嵌入到原型空间中，每一维对应一个相应的原型



原始空间

原型空间

实验

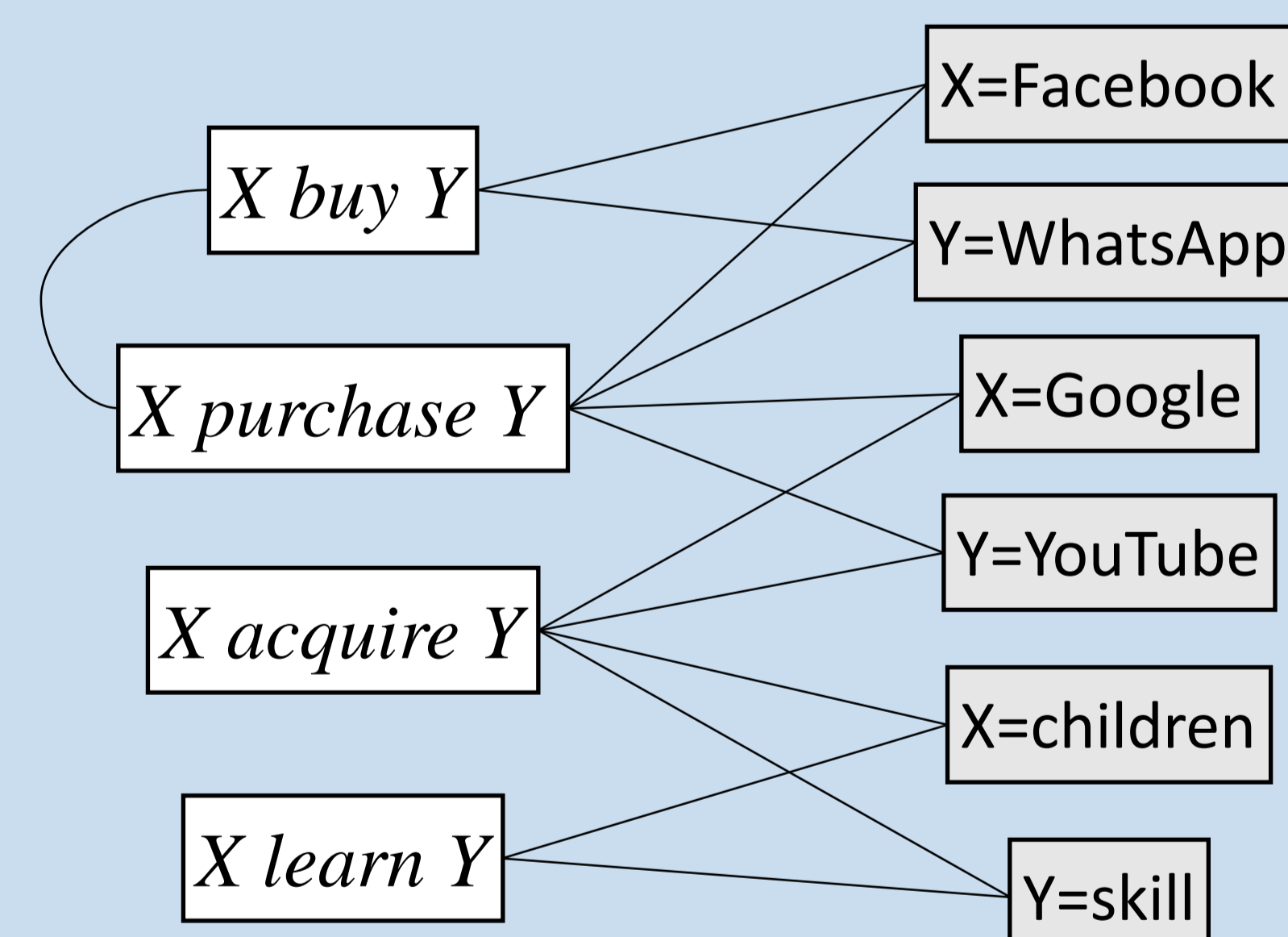
KBP			
System	P	R	F1
Mintz++	0.260	0.250	0.255
Hoffmann	0.306	0.198	0.241
Surdeanu	0.249	0.314	0.278
DSCNN	0.286	0.214	0.244
DSProto	0.459	0.231	0.307

文本推理规则发现

创新点

针对推理规则发现中的模板歧义性问题和上下文表示稀疏性问题，提出了一种基于图的上下文敏感推理规则发现方法

模型



$$P(i \rightarrow j|c) = \begin{cases} (1 - \lambda_{c,i}) \frac{w_{ij}}{\sum_k w_{ik}} & \text{transmit to neighborhood } j \\ \lambda_{c,i} & \text{restart to start node } s \end{cases}$$

Context-Aware Random Walking

实验

Context Feature	No Context	X=Microsoft Y=Nokia	X=people Y=language
X=Facebook	0.045	0.055	0.003
Y=WhatsApp	0.045	0.055	0.003
X=Google	0.064	0.092	0.073
Y=YouTube	0.064	0.092	0.073
X=children	0.119	0.080	0.163
Y=skill	0.119	0.080	0.163

Table 1. The representations of 'X acquire Y' in different contexts ($\lambda=0.1$ and semantic-dependent edge weight 0.5)