



# 事理图谱的构建及应用

丁 效

哈尔滨工业大学

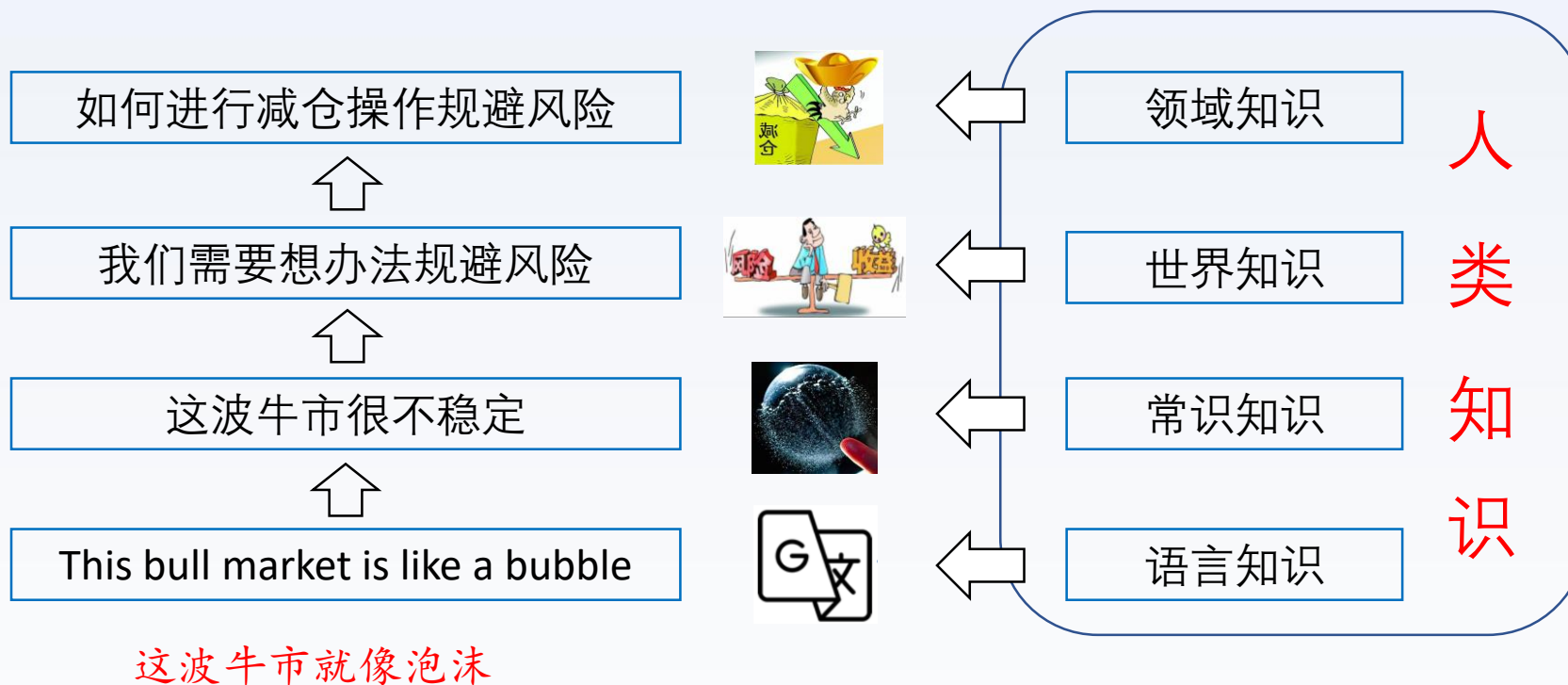
社会计算与信息检索研究中心

2019.12.14 南京



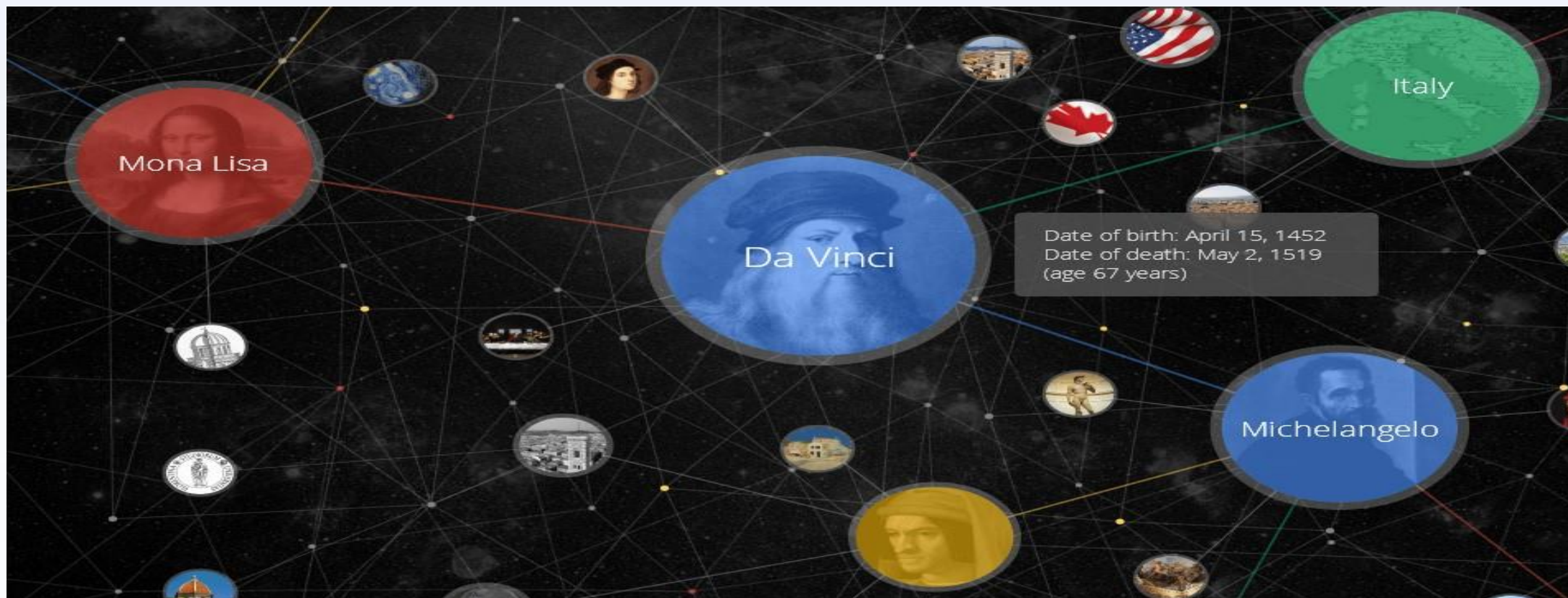
# 面临挑战

- 对自然语言的深度理解需要复杂**知识**的支持



# 技术挑战

- 知识库的构建、表示及应用





# 知识图谱的构建，从人工到自动





# 封闭域知识图谱 vs 开放域知识图谱

## 封闭域

**DBpedia:**  
4 类基本关系  
Part-of/kind-of/  
instance-of/  
attribute-of

关系类型人为设定

- 优点：易于扩充、无冗余
- 缺点：灵活性差、不便应用

## 开放域

**Google Knowledge Graph**

**百度知识图谱**

关系类型自动挖掘

- 优点：关系灵活、覆盖全面
- 缺点：冗余不一致、难以支持逻辑推理





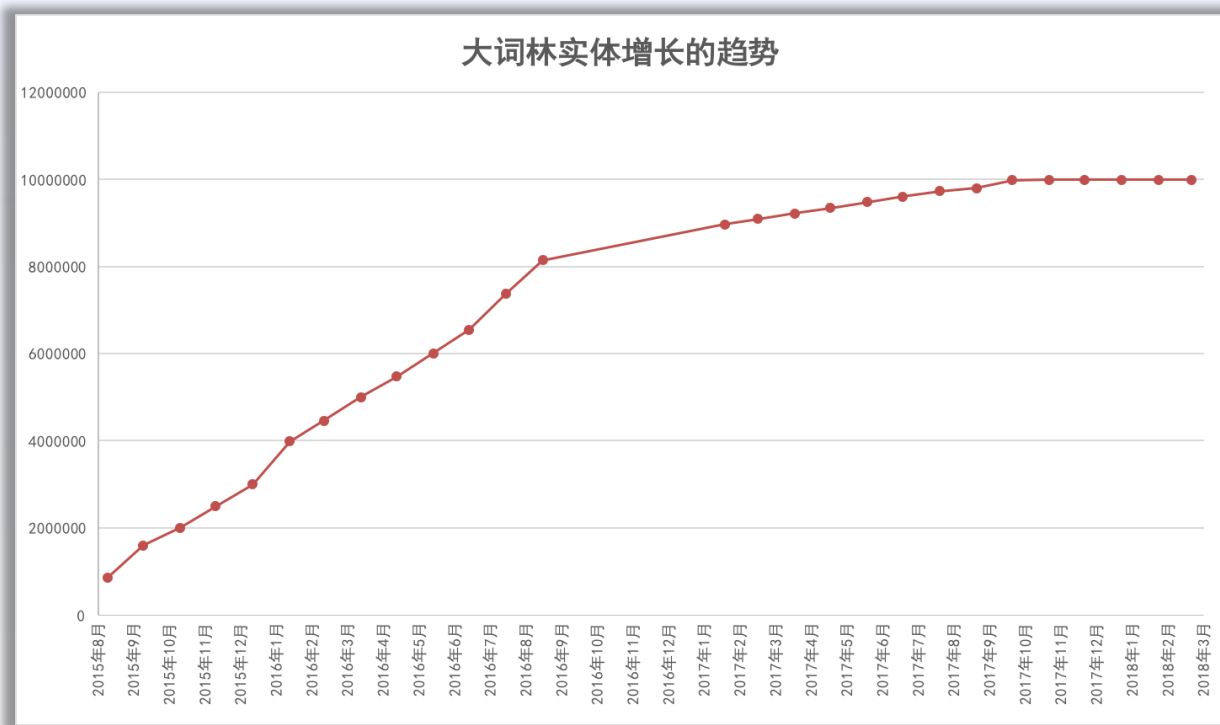
# 2014年末《大词林》正式发布

- 2014年11月27日，哈工大正式发布《大词林》
- 《大词林》是**自动构建**的基于**上下位关系**的大规模**开放域**中文知识图谱





# 大词林 (BigCilin, <http://www.bigcilin.com>)



**实体词数量：** 10,672,790 (1千万)

**概念词数量：** 182,083 (18万)

**实体间三元组对：** 45,854,785 (4千万)

**关系（属性）数：** 161,983 (十六万)





# 事理图谱的研究意义

- 现有的知识库普遍是以“**概念及概念间的关系**”为核心的，缺乏对“**事理逻辑**”知识的挖掘
- 事理逻辑（**事件之间的演化规律与模式**）是一种非常有价值的人类知识，挖掘这种知识对我们**认识人类行为**和**社会发展变化规律**非常有意义







# 什么是事理图谱？

- 事理图谱：Eventic Graph (EG)，事理：Eventics
- 英文翻译，类比：语义Semantics，语义网Semantic Web
- 定义：事理图谱是一个**事理逻辑知识库**，描述了事件之间的演化规律和模式。结构上事理图谱是一个有向有环图，节点代表事件，有向边代表事件之间的顺承、因果、条件和上下位等逻辑关系。





# 事理图谱的应用

- 事理图谱可应用于事件预测、常识推理、消费意图挖掘、对话生成、问答系统、辅助决策等任务中
- 大规模事理图谱将和传统知识图谱一样，具有非常巨大的应用价值。





# 事理图谱与知识图谱的区别与联系

	事理图谱	知识图谱
研究对象	谓词性事件及其关系	体词性实体及其关系
组织形式	有向有环图	有向有环图
主要知识形式	事理逻辑关系，以及概率转移信息	实体属性和关系
知识的确定性	事件间的演化关系多数是不确定的	多数实体关系是确定性的





# 事理图谱中的事件定义

- 前人工作: 事件是特定时间、地点下的一个状态变化
  - 经典的事件抽取和分类任务, ACE 2005
  - 话题检测与跟踪
- 事理图谱中的事件
  - 理论上, 事理图谱中的事件是具有一定抽象程度的**泛化事件**。
  - 表示为**抽象、语义完备**的谓词性词或词组, 也可以表示为可变长度的、结构化的(主体, 事件词, 客体)多元组
  - 其中**必然包含一个事件词**, 标志事件的发生, 例如: “跑步”; 而主体和客体都可以在不同的应用场景下被省略, 例如: “(元首, 出访)”可以省略事件的客体; “(购买, 机票)”可以省略事件的主体

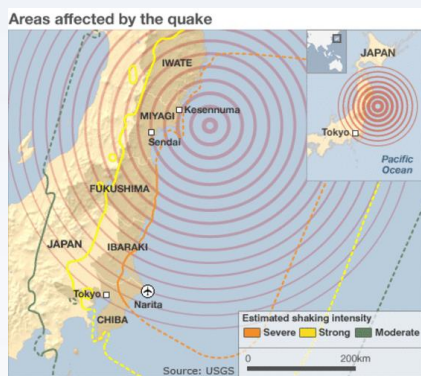




# 事件关系可归纳：从事实到认知

- 例子：A massive 8.9-magnitude *earthquake* hit northeast Japan on Friday, which cause a large amount of *houses collapsed*.

(星期五，日本东北部发生8.9级大地震，造成大量房屋倒塌。)



常识性知识:

地震 → 房屋倒塌



# 事理图谱的研究动机

- 从具体的、细节化的事件因果（**事实**）中归纳总结出事件因果律（**认知**）
- 利用事件因果律（**认知**）指导实践，例如事件预测和股票预测（**事实**）





# 事件间顺承关系

- 顺承关系(Sequential)是指两个事件在时间上相继发生的偏序关系
- 两个前后顺承的事件之间存在一个介于0到1之间的**转移概率**，表示从一个事件按时序顺承关系演化到下一事件的**概率**

吃过午饭后，小明到前台买单，然后离开了餐馆。





# 事件间因果关系

- 因果关系 (Causal) 是指两个事件之间，前一事件 (原因) 导致后一事件 (结果) 的发生，满足发生时间上的偏序约束
- 因果关系是顺承关系的子集；因果事件对之间存在一个介于0到1之间的因果强度值，表示该因果关系的**置信度**

核泄漏引起了严重的海洋污染。



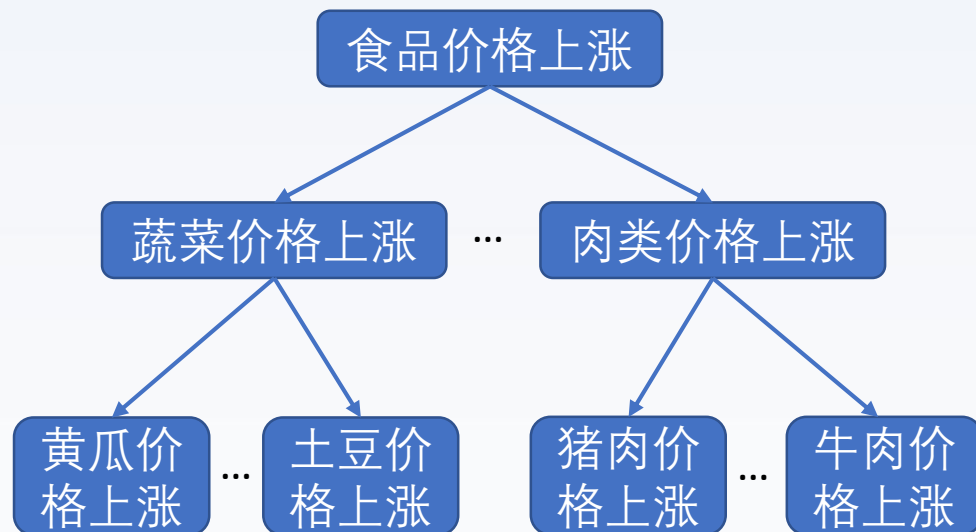




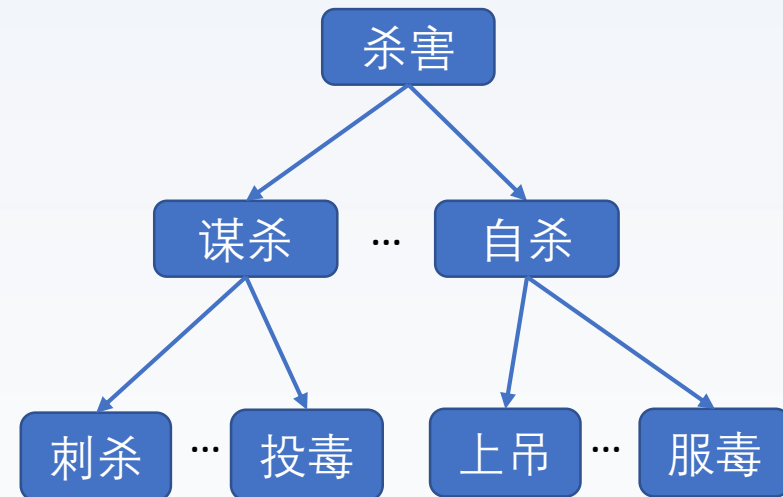
# 事件间上下位关系

- 事件之间的上下位关系有两种：名词性上下位关系和动词性上下位关系

名词性上下位关系



动词性上下位关系





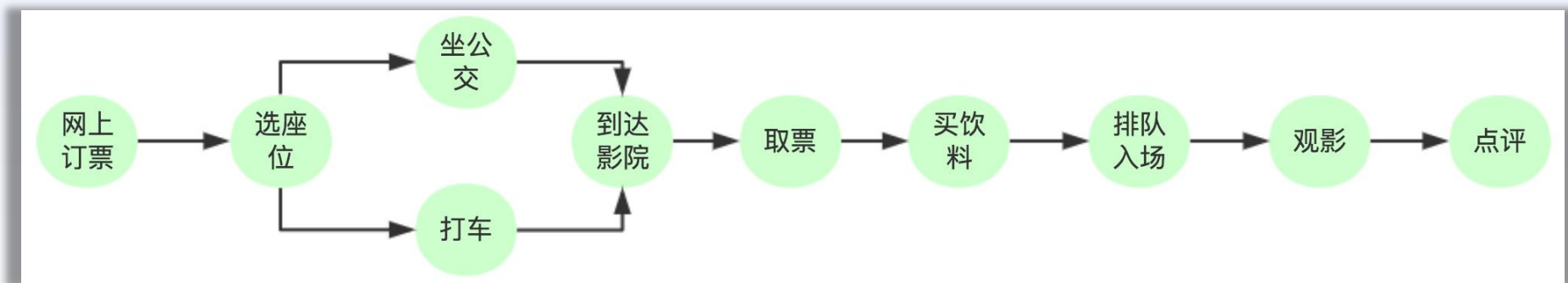
# 事件间条件关系

- 条件关系是指前一个事件的发生是后一个事件发生的条件

只有刻苦学习，才能取得好成绩。

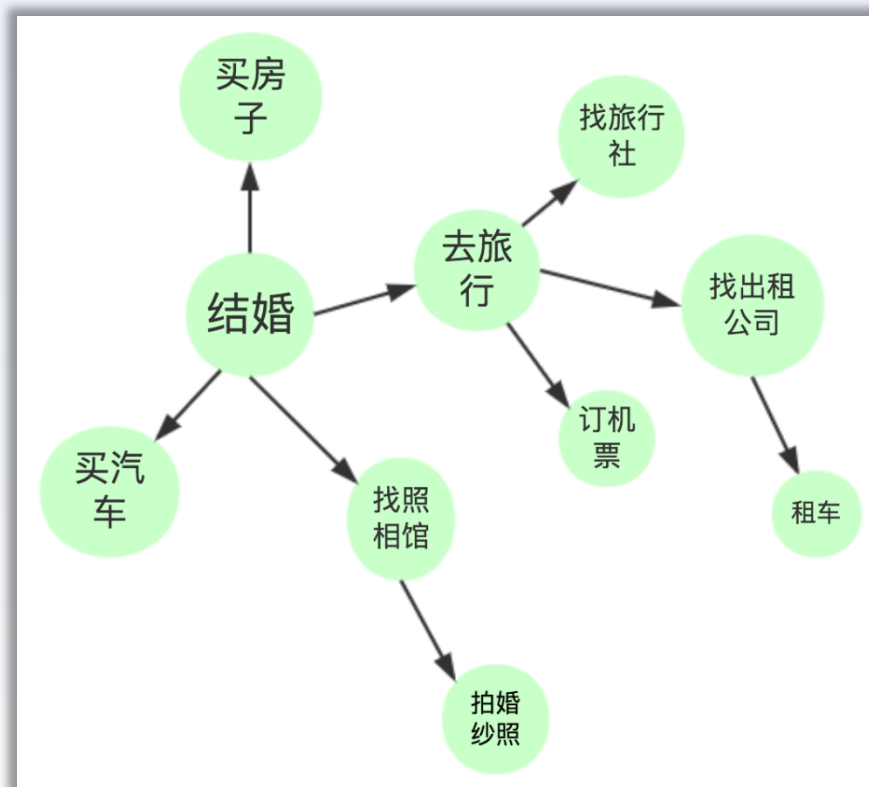


# 事理图谱的三种典型子图结构



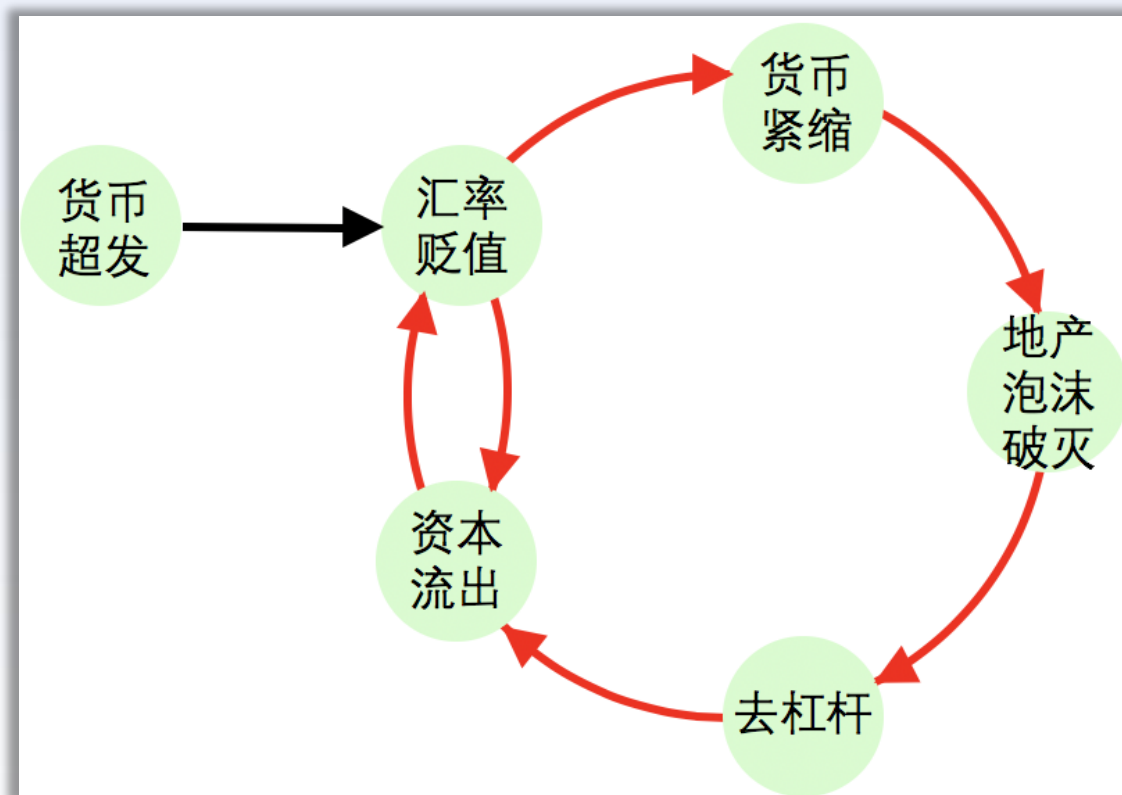
‘看电影’场景下的**链状**事理图谱

# 事理图谱的三种典型子图结构



‘结婚’场景下的**树状**事理图谱

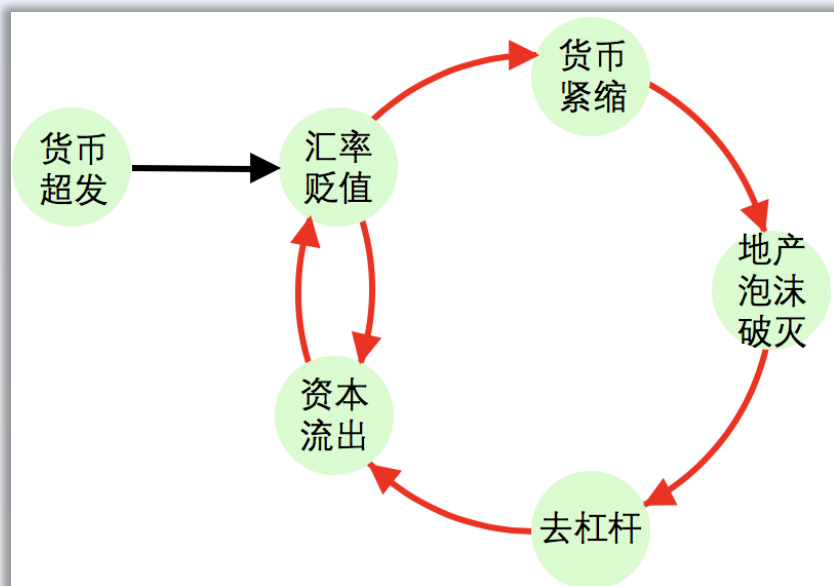
# 事理图谱的三种典型子图结构



金融领域的**环状**事理图谱

# 事理图谱的三种典型子图结构

## 漩涡——汇率机制带来的贬值与资本流出恶性循环

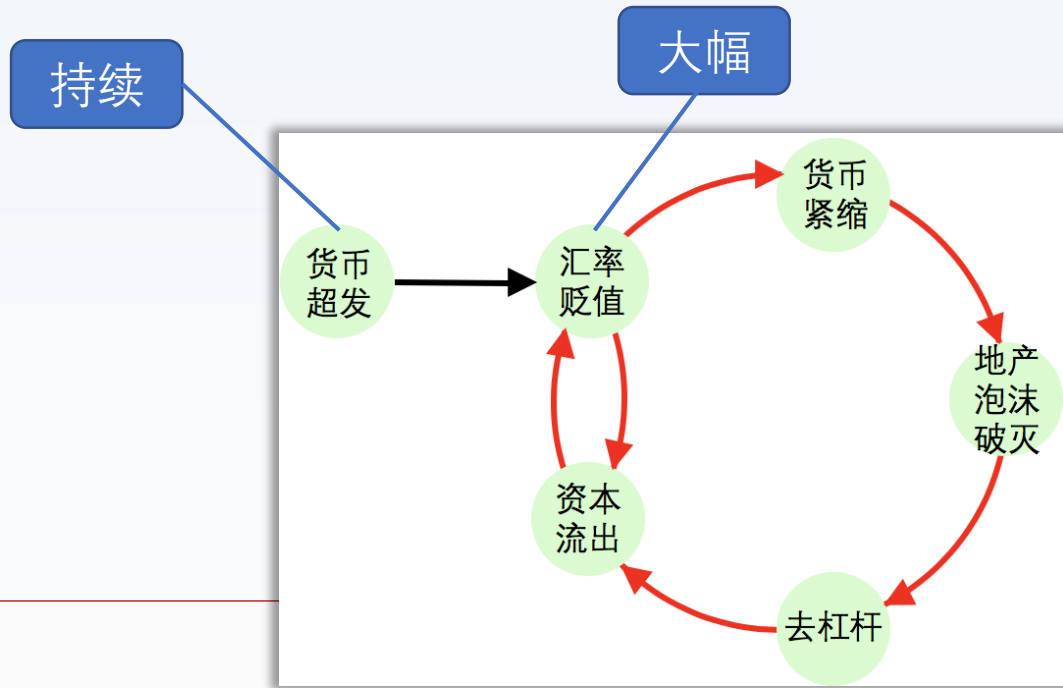


经济和政策基本面冲击的力度大小、外汇市场所处的前定格局都会影响到净短期资本流动规模。汇率形成机制同样会影响净资本流动规模。不同的汇率形成机制下，冲击所带来的净短期资本流动变化大相径庭。浮动汇率下，内部或外部冲击所引发的资本流动压力会导致汇率调整，也因此改变国内外金融资产的相对价格，汇率价格浮动自发地平抑资本流入或者流出。固定汇率下，内部或者外部冲击所引发的资本流动对汇率形成压力，货币当局不得不调整利率，也因此改变国内外金融资产价格相对价格，平抑资本流动并维持固定汇率水平。

金融领域的**环状**事理图谱

# 事件属性

- 事理图谱除了关注事件之间的事理逻辑关系外，还关注事件自身的属性
- 事件属性是指附加到事件上的所有细节，这些细节有助于我们理解事件的上下文
  - 例如，事件属性可以用来描述事件发生的程度、持续时间等





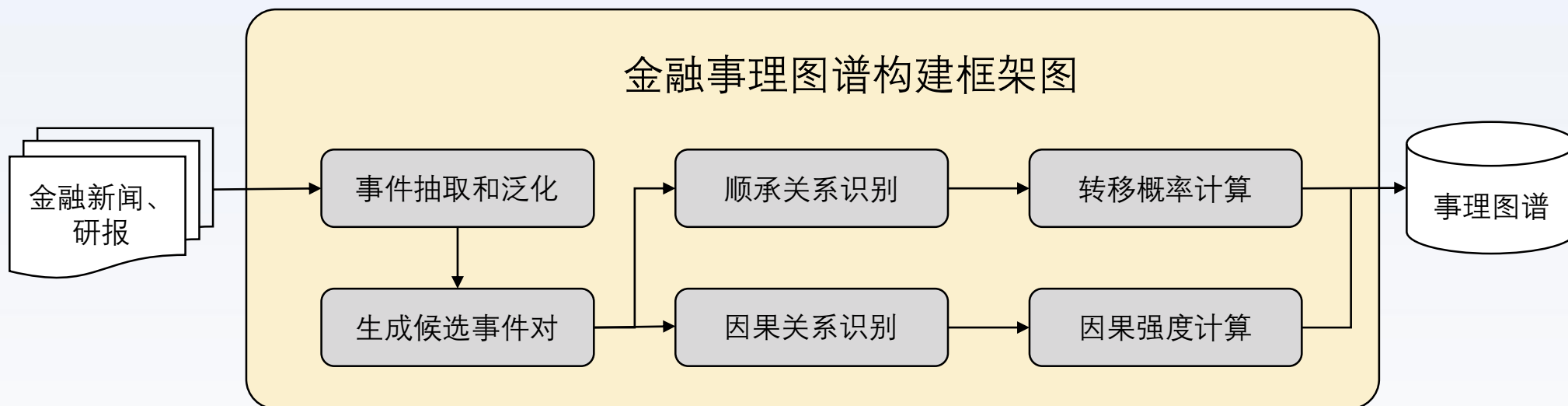
# 事理图谱的构建过程

- **前向演化**：按照时间发生顺序和四种事理关系向前演化（由因导果）
- **反向演化**：按照时间发生顺序和四种事理关系反向演化（执果溯因）
- **向上抽象**：将某些同类事件进行合并，得到一个更加抽象的上层事件
- **向下分裂**：选择某个事件节点作为场景事件，将其进行分裂，得到更加具体、精确的局部子事件





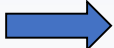
# 金融事理图谱的构建

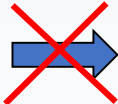




# 无监督“因果事件关系”抽取方法

- 规则模板+词性标注信息抽取具体因果事件
  - 使用规则<Pattern, Constraint, Priority>匹配因果事件因果提及文本
  - Pattern是抽因果的正则表达式，例如“after [sentence1], [sentence2]”
  - Constraint是句法结构约束，例如[sentence1]不能以数字开头
  - Priority表示如果多个Pattern匹配成功时的决策

“after Iraq war, the oil price falls”  (Iraq war, the oil price falls)

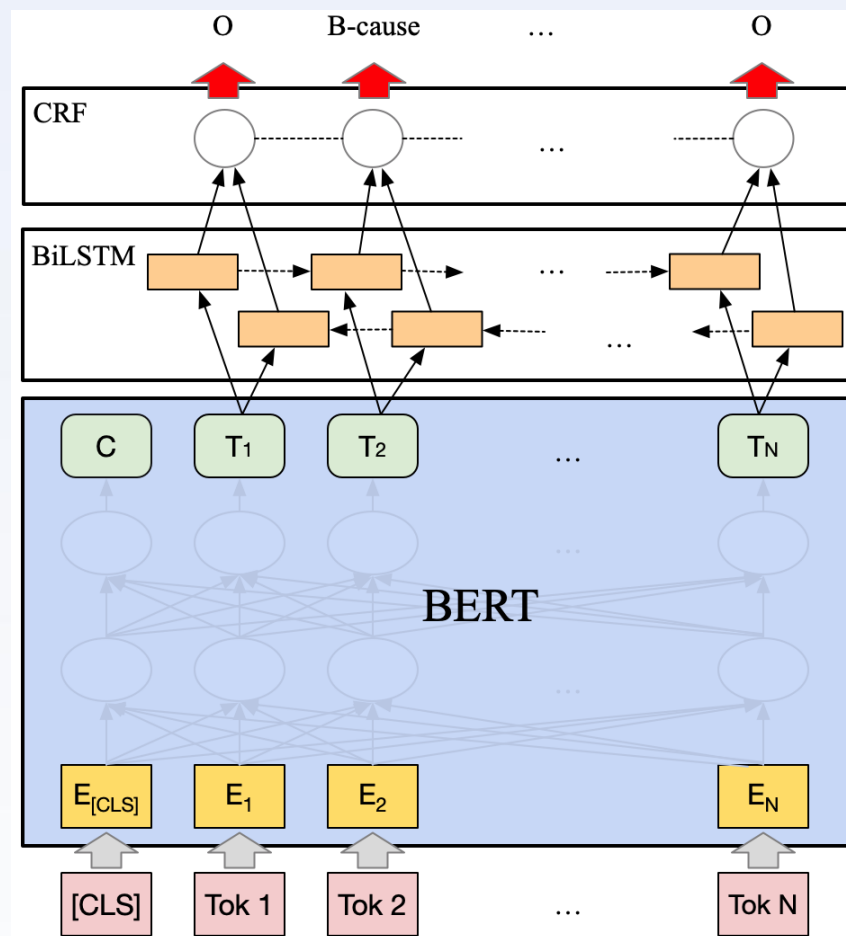
“After 2 Years, US Capitol Restoration Project Is Complete”  (2 Years, US Capitol Restoration Project Is Complete)





# 有监督“因果事件关系”抽取方法

- 基于BERT的因果关系抽取模型
  - 使用BERT进行特征提取
  - 使用BiLSTM+CRF输出标签



根据特定任务标注数据训练序列标注模型

预训练语言模型





# Transferable BERT (可迁移的BERT)

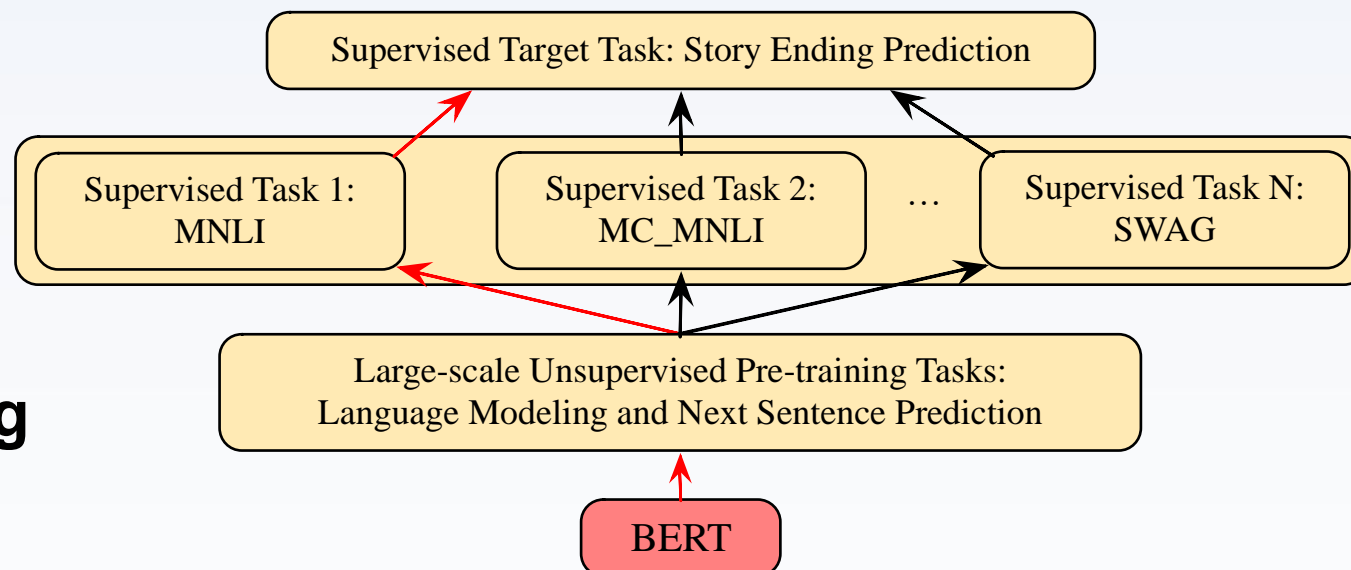
## • 动机

- 传统BERT在大规模无监督数据上进行预训练
- **Transferable BERT**提出三层训练框架，在原有Pre-training和Fine-tuning框架下，增加在一个**相关任务预训练层**，学习相关任务的有监督信息

## 3. Fine-tuning

## 2. Supervised Pre-training

## 1. Unsupervised Pre-training





# 有监督 “因果事件关系” 抽取方法

- 测试集实验结果
  - Token-level

模型	句子数	Token数	评价方法	Token-level accuracy
BERT+BiLSTM+CRF	178	4027	自动评价	<b>0.8423</b>
BERT	178	4027	自动评价	0.8361
全部预测O标签	178	4027	自动评价	0.3241

- Mention-level

模型	数据集大小	评价方法	Precision	Recall	F1 score
BERT+BiLSTM+CRF	178	自动评价	0.8421	0.8623	<b>0.8521</b>
BERT	178	自动评价	0.8218	0.8563	0.8190
规则模板	178	自动评价	0.6364	0.6707	0.6531





# 因果强度计算 (1)

- 基于统计的因果强度计算方法

记 $f(i_c, j_e)$ 为单词 $i$ 出现在原因中且单词 $j$ 出现在结果中的频数  
使用最大似然估计得到 $i$ 作为 $j$ **必要条件**的概率:

$$CS_{nec}(i_c, j_e) = \sqrt{\frac{f(i_c, j_e)}{\sum_w f(w_c, j_e)}}$$

类似地得到 $i$ 作为 $j$ **充分条件**的概率:

$$CS_{suf}(i_c, j_e) = \sqrt{\frac{f(i_c, j_e)}{\sum_w f(i_c, w_e)}}$$





## 因果强度计算 (2)

$$CS_{nec}(i_c, j_e) = \sqrt{\frac{f(i_c, j_e)}{\sum_w f(w_c, j_e)}} \quad CS_{suf}(i_c, j_e) = \sqrt{\frac{f(i_c, j_e)}{\sum_w f(i_c, w_e)}}$$

求两个概率的加权几何平均数，作为*i*与*j*间的因果强度

$$CS(i_c, j_e) = CS_{nec}(i_c, j_e)^\lambda CS_{suf}(i_c, j_e)^{1-\lambda}$$

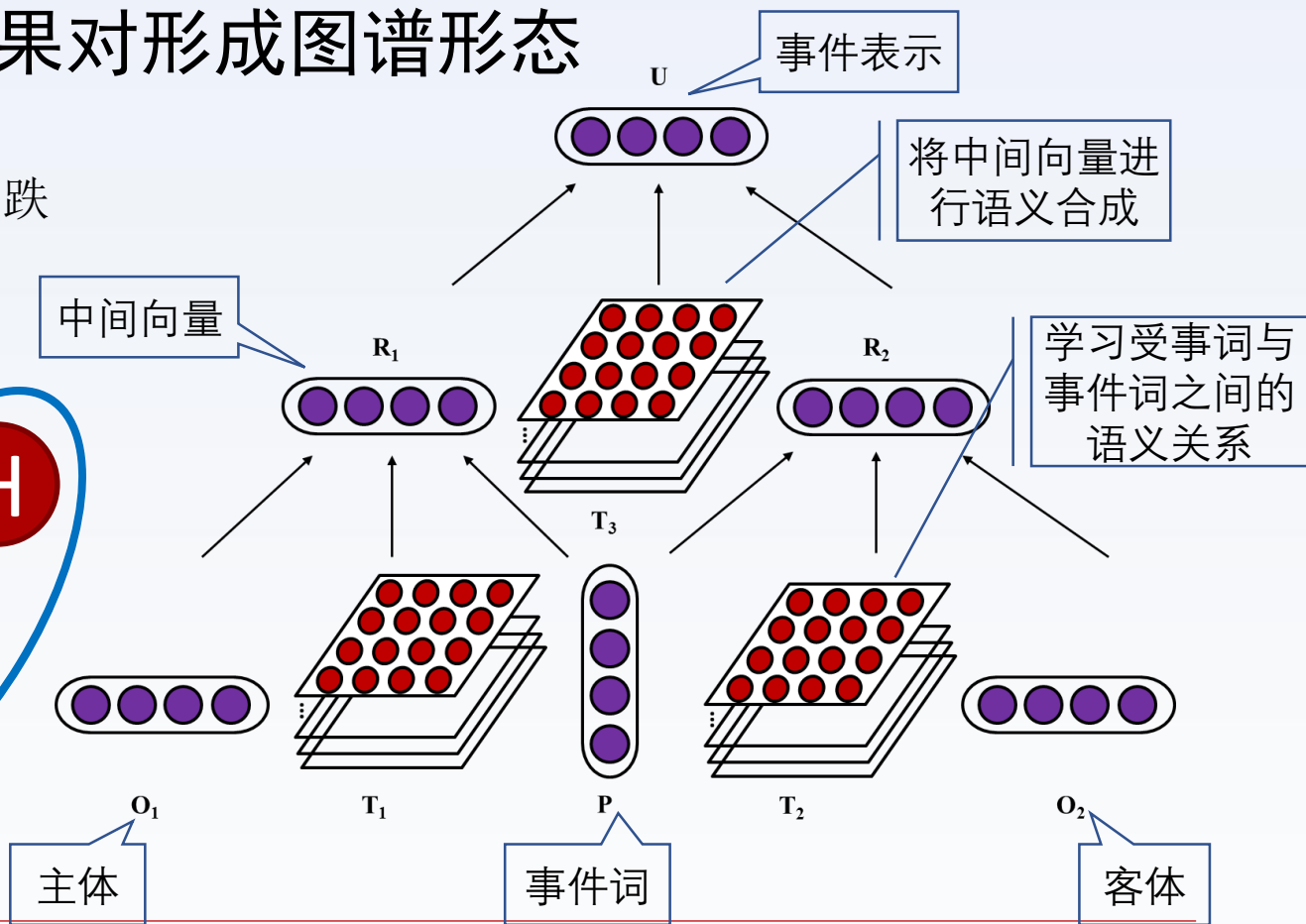
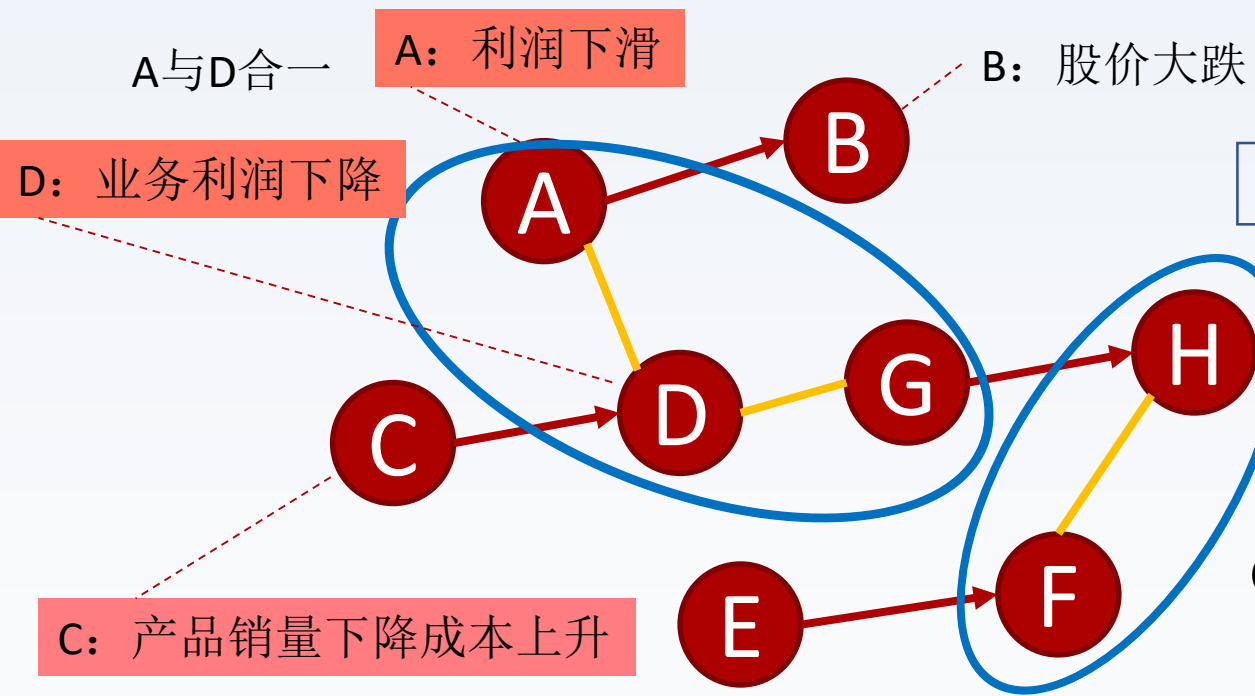
给定事件 $T_1, T_2$ ，将其中**所有单词的因果强度组合作为事件的因果强度**

$$CS_T(T_1, T_2) = \frac{1}{|T_1| + |T_2|} \sum_{i \in T_1} \sum_{j \in T_2} CS(i, j)$$



# 事件合一

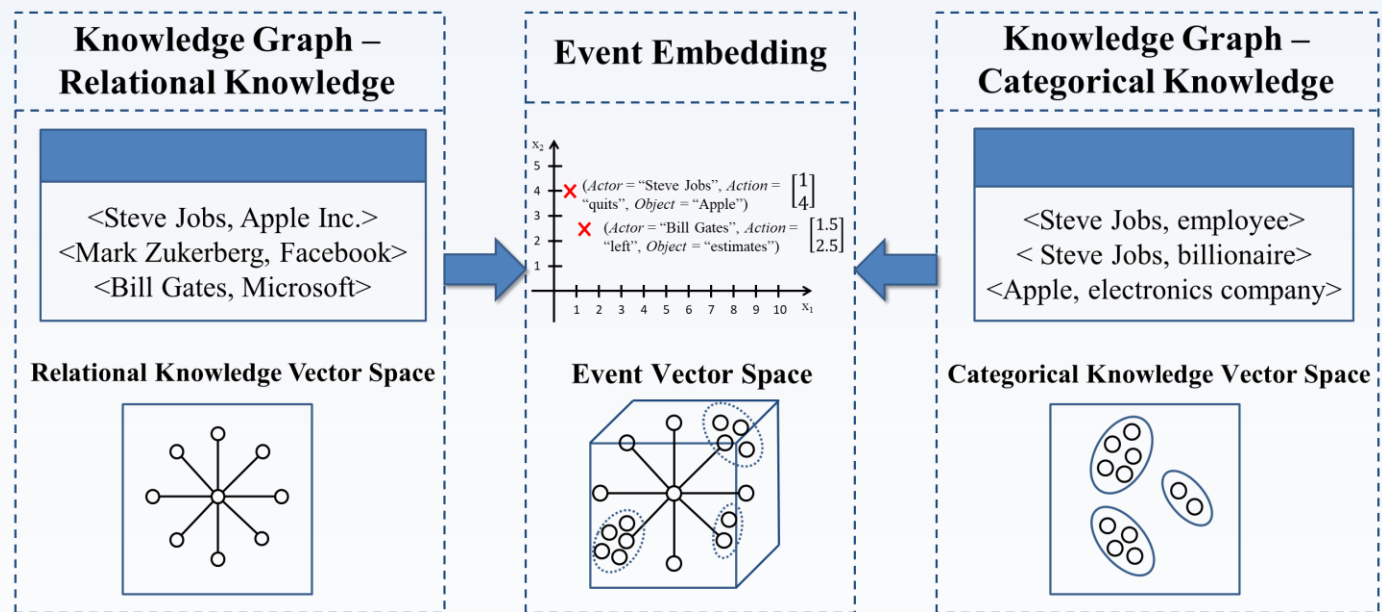
- 通过对某些事件的合一，将因果对形成图谱形态





# 知识驱动的事件表示学习

- 缺乏知识造成的歧义
  - 具有相同或相似的词但是语义不同的事件，学到相似的事件向量
  - 举例：乔布斯离开苹果 vs 小明离开星巴克





# 知识驱动的事件表示学习模型

联合学习模型:

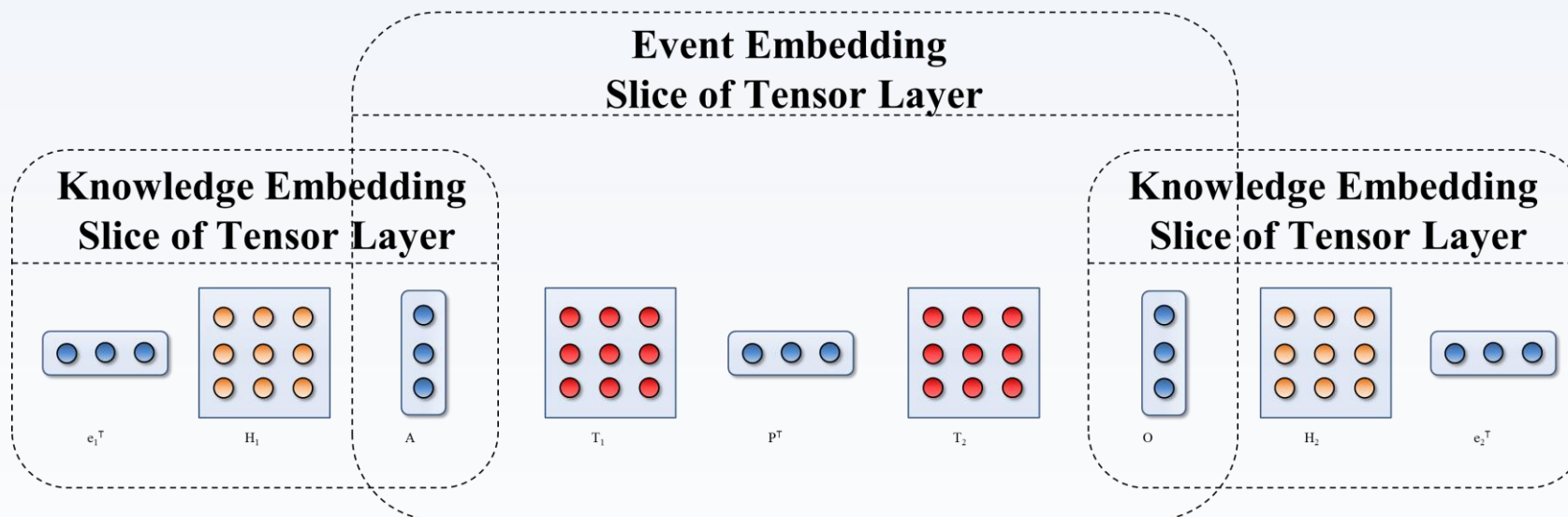
$$L = \alpha L_{\mathcal{E}} + (1 - \alpha) L_{\mathcal{K}}$$

事件表示学习:

$$L_{\mathcal{E}} = \text{loss}(E, E^r) = \max(0, 1 - g(E) + g(E^r)) + \lambda \|\Phi\|_2^2,$$

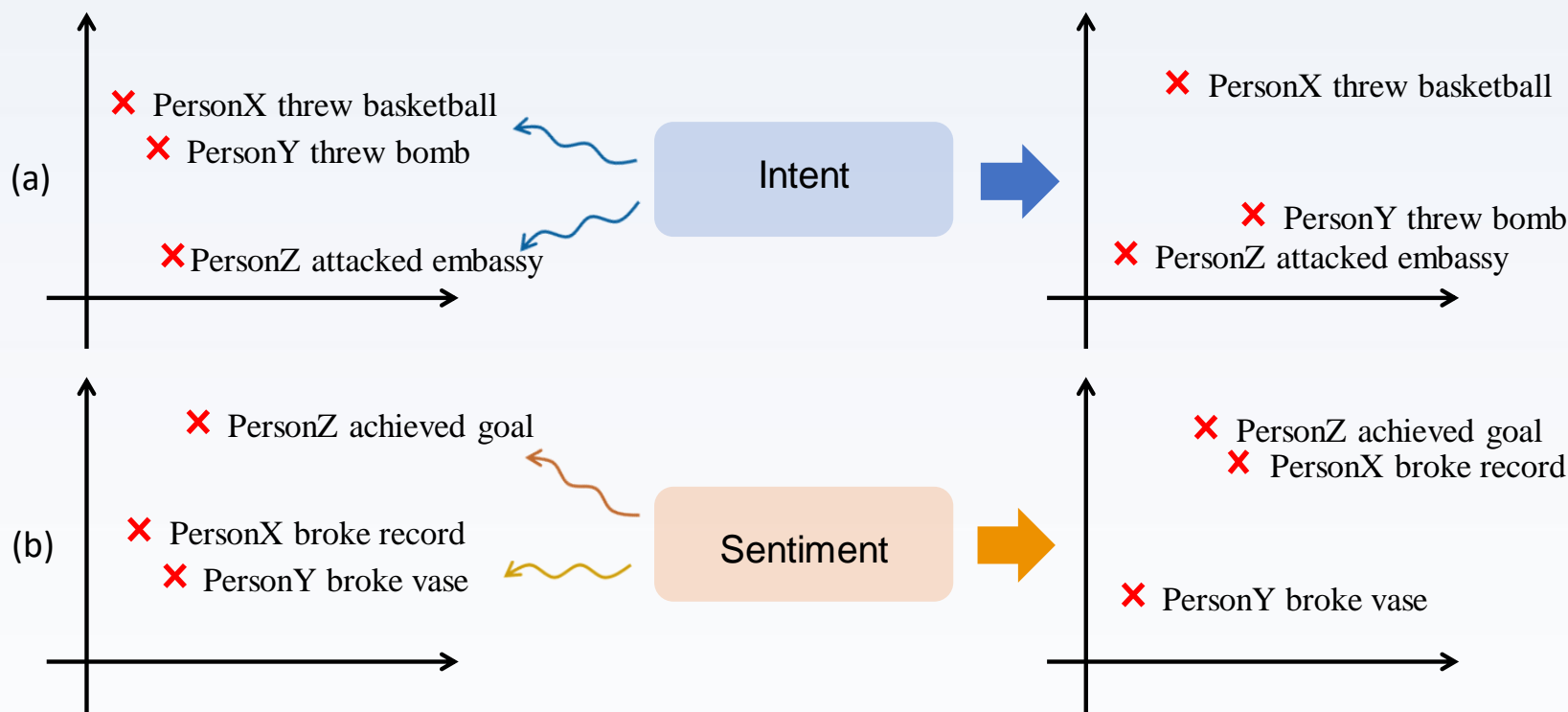
知识图谱表示学习:

$$L_{\mathcal{K}} = \sum_{i=1}^N \sum_{m=1}^M \max(0, 1 - g(T^{(i)}) + g(T_c^{(i)})) + \lambda \|\Omega\|_2^2,$$



# 常识知识（情感、意图）增强的事件表示学习

- 事件是对客观事实的表达，然而客观事件的发生会对人类的主观情感产生影响，不同的事件其背后的意图也有所不同





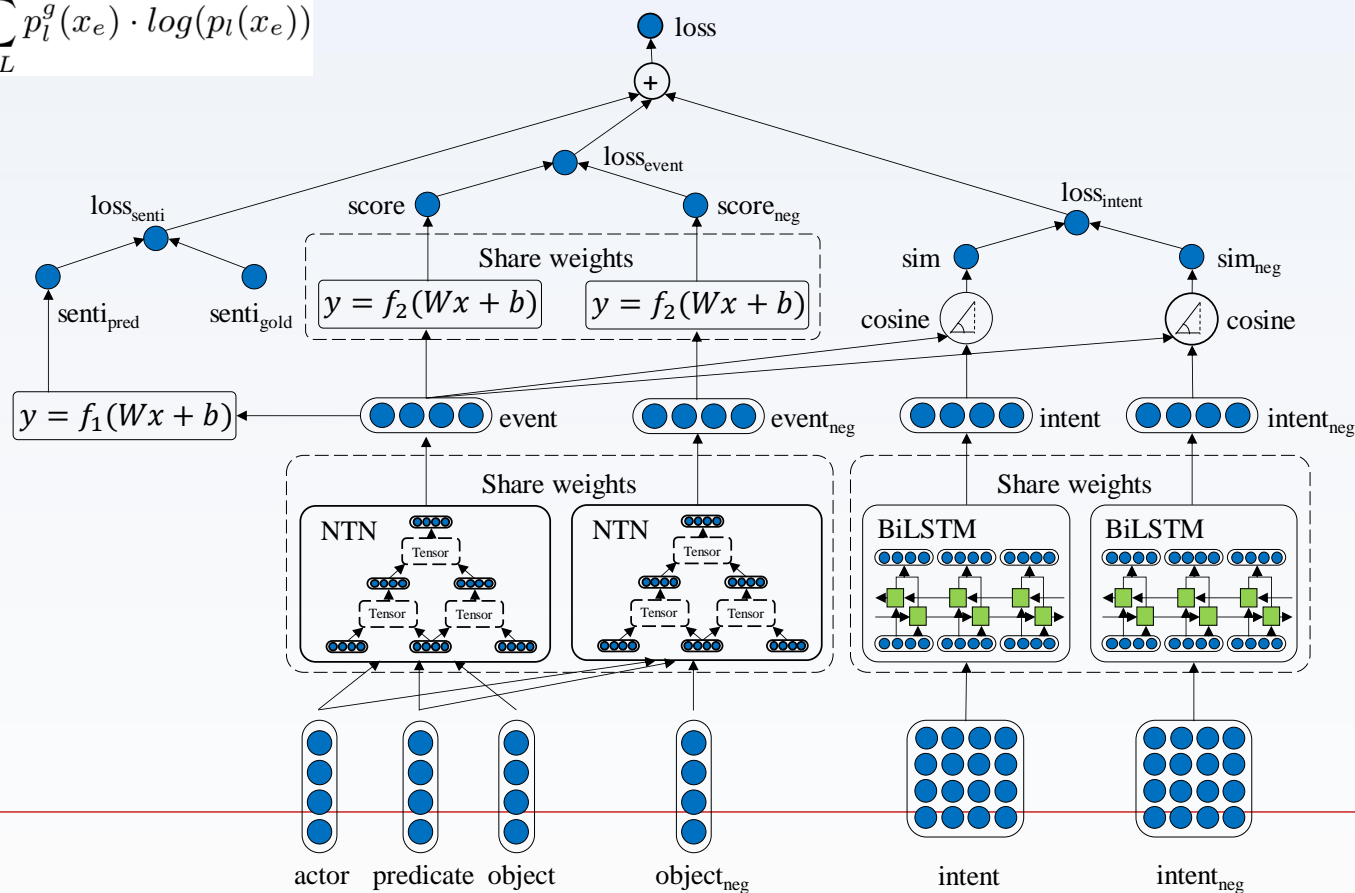
# 常识知识（情感、意图）增强的事件表示学习模型

多任务学习模型:  $\mathcal{L} = \alpha\mathcal{L}_E + \beta\mathcal{L}_I + \gamma\mathcal{L}_S$

事件表示学习:  $\mathcal{L}_E = \text{loss}(E, E^r) = \max(0, 1 - g(E) + g(E^r)) + \lambda\|\Phi\|_2^2$

意图表示学习:  $\mathcal{L}_I = \max(0, 1 - \text{cosine}(\mathbf{v}_e, \mathbf{v}_i) + \text{cosine}(\mathbf{v}_e, \mathbf{v}'_i))$

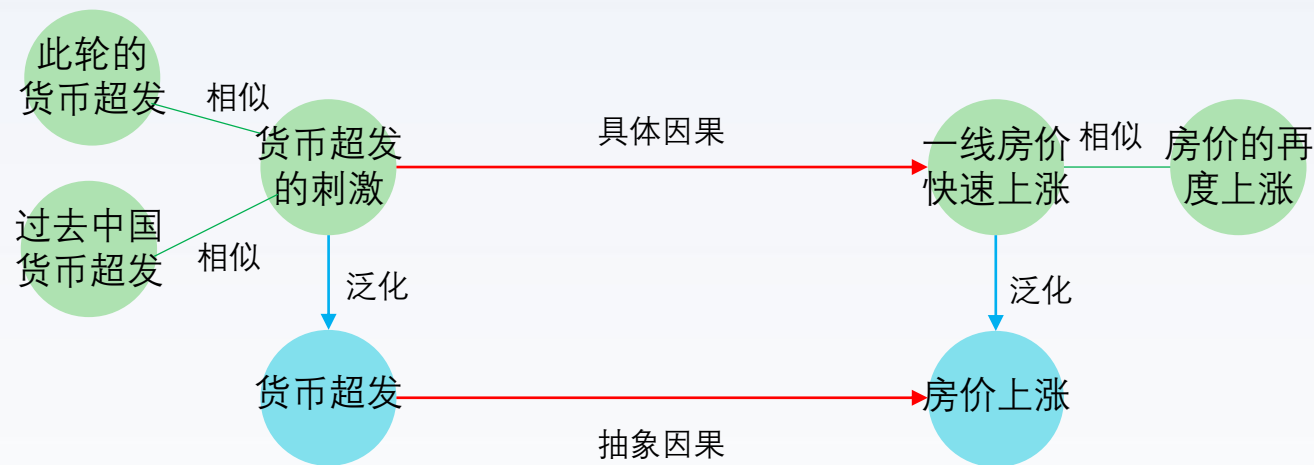
情感表示学习:  $\mathcal{L}_S = - \sum_{x_e \in C} \sum_{l \in L} p_l^g(x_e) \cdot \log(p_l(x_e))$





# 事件泛化：聚类

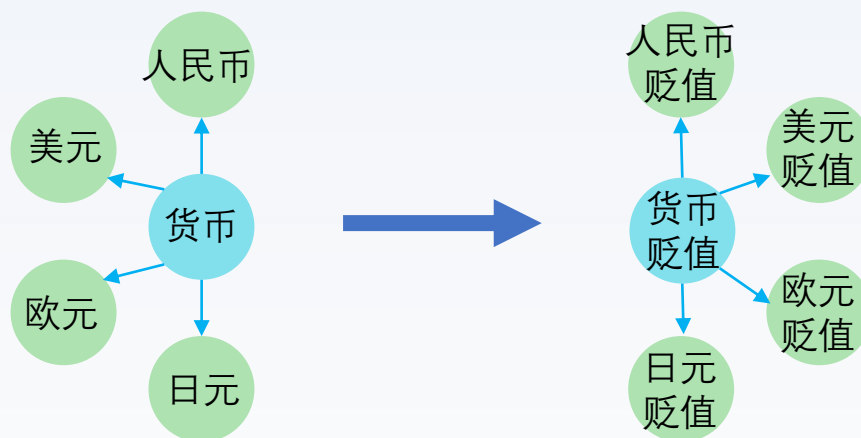
- 基于事件聚类的泛化方法
  - 相似事件聚成一类，取类别中所有相似事件中公共成分作为泛化事件
  - 同时“具体因果边”也被泛化成“抽象因果边”





# 事件泛化：事件上下位关系获取

- 基于《大词林》的事件名词性上下位关系获取
  - 对事件中的主体/客体，在大词林中寻找其上位词
  - 若找到，将该词替换为上位词，替换后的事件作为上位事件



基于《大词林》获取到的上位事件数量： 1,883,792

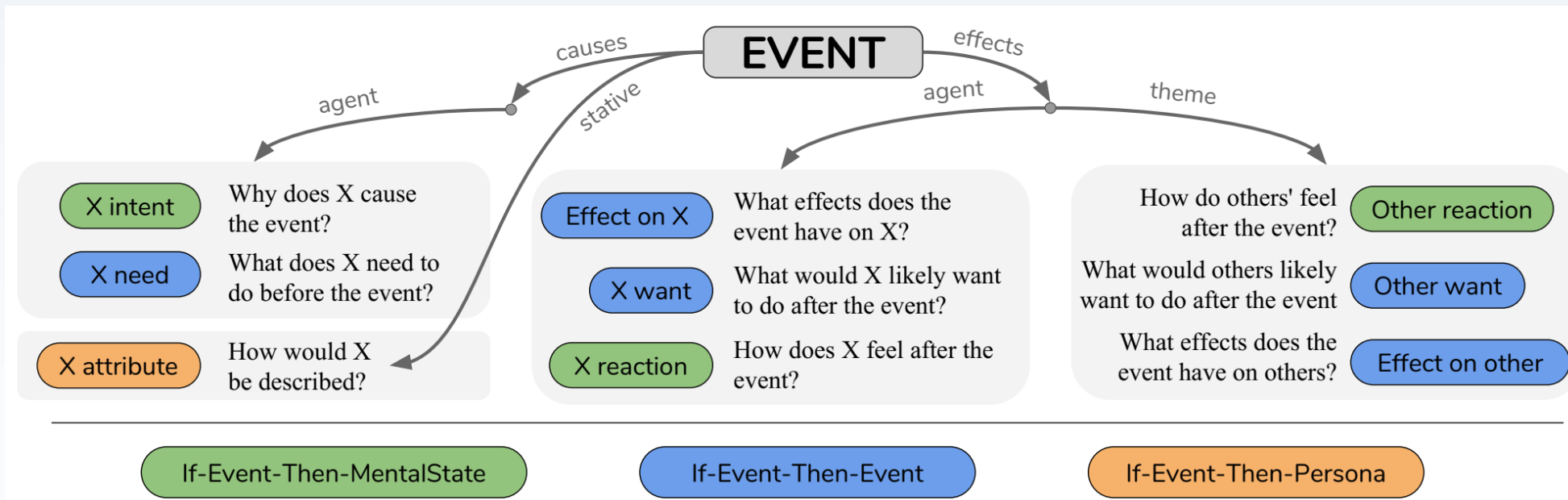




# 基于事件背景知识进行IF-THEN类型常识推理

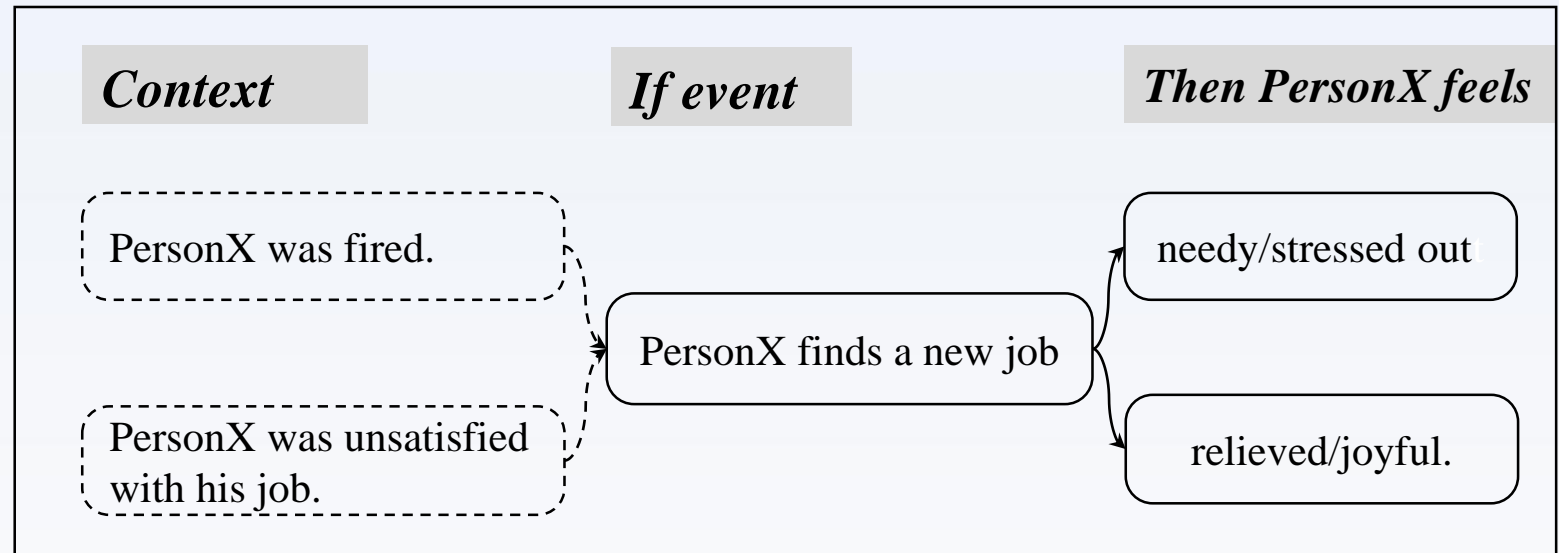
- IF-THEN类型常识推理：

- 华盛顿大学Noah Smith团队提出的一种常识推理任务，并通过众包方式标注数据集
- 给定事件，要求生成事件参与者的心理状态（意图、需求），后续事件等



# 基于事件背景知识进行IF-THEN类型常识推理

- IF-THEN常识推理存在的问题：
  - 事件的演化是多种多样的，需要根据背景/上下文信息帮助推理事件的发展脉络
- 一对多生成问题：
  - 同一事件对应多种可能结果
  - 基于RNN的Seq2Seq模型倾向于给出某些“泛泛”的回答



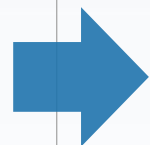


# 上下文感知的变分自编码器和两阶段训练过程

- 我们提出**上下文感知的变分自编码器**(Context-aware Variational Autoencoder, CWVAE)和相应的**两阶段训练过程**以解决上述问题：
- 上下文感知的变分自编码器：
  - 变分自编码器可以帮助解决生成多样性
  - 相比于传统变分自编码器，引入额外的上下文感知隐变量，以学习事件背景知识
- 两阶段训练过程引入事件背景知识：

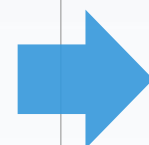
## 构建辅助数据集

- 基于故事结尾预测语料构建辅助数据集
- 该语料包含丰富事件背景知识



## 预训练阶段

- CWVAE在辅助数据集上学习事件背景知识



## 微调阶段

CWVAE在标注数据集上进行微调

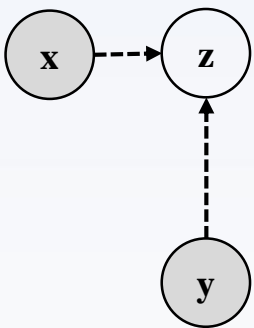


# 上下文感知的变分自编码器和两阶段训练过程

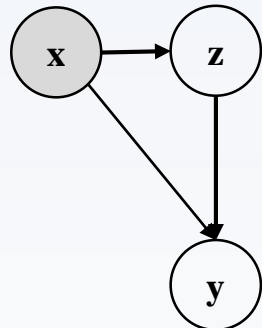
- 变分自编码器

- [事件 $x$ ]  $\rightarrow$  事件属性知识 $y$
- $z$ : 隐变量

(a) 训练



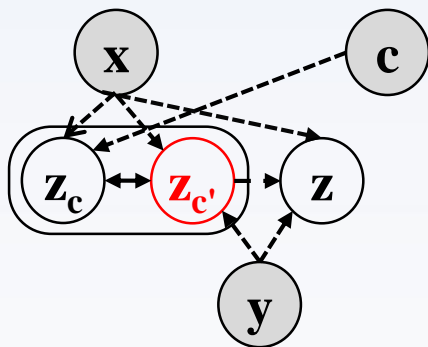
(b) 生成



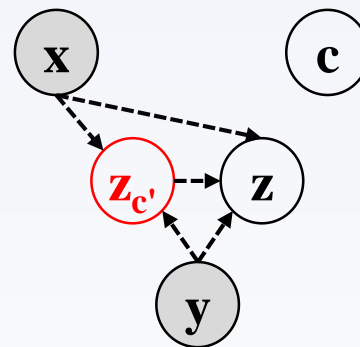
- 上下文感知的变分自编码器

- [(背景 $c$ ), 事件 $x$ ]  $\rightarrow$  事件属性知识 $y$
- $z_{c'}$ : 上下文感知隐变量

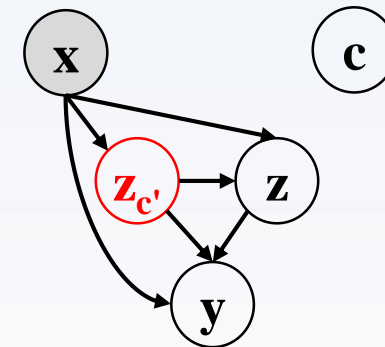
(a) 预训练



(b) 微调



(c) 生成



- 在辅助数据集上, 利用上下文感知隐变量, 学习事件背景知识

- 标注数据集上进行微调





# 实验结果

## On Atomic

- 准确性:

Metric	Methods	xIntent	xNeed	xAttr	xEffect	xReact	xWant	oWant	oReact	oEffect
PPL	RNN-based Seq2Seq	22.54	24.69	33.54	65.13	29.52	26.63	16.76	14.99	35.17
	Variational Seq2Seq	26.48	28.31	33.00	68.62	29.93	29.50	16.98	14.25	34.20
	VRMNT	21.04	24.28	24.87	61.05	26.62	28.57	14.45	14.86	30.12
	CWVAE-Unpretrained	20.73	23.72	25.80	60.62	25.75	26.71	15.93	12.82	32.00
	CWVAE	<b>15.93</b>	<b>20.32</b>	<b>23.85</b>	<b>50.74</b>	<b>21.39</b>	<b>24.02</b>	<b>14.02</b>	<b>11.70</b>	<b>29.13</b>
BLEU	RNN-based Seq2Seq	8.17	12.35	2.96	5.26	3.43	13.44	7.08	4.09	6.42
	Variational Seq2Seq	8.31	12.05	2.13	6.07	2.52	11.71	7.40	4.08	6.38
	VRMNT	9.52	13.35	4.87	4.42	7.64	9.80	10.79	5.28	13.71
	CWVAE-Unpretrained	11.37	14.64	4.07	14.11	7.86	12.70	12.09	8.16	<b>14.93</b>
	CWVAE	<b>12.12</b>	<b>15.67</b>	<b>5.63</b>	<b>14.64</b>	<b>8.13</b>	<b>15.01</b>	<b>13.83</b>	<b>8.58</b>	11.63

- 多样性:

Metric	Methods	xIntent	xNeed	xAttr	xEffect	xReact	xWant	oWant	oReact	oEffect
dist-1	RNN-based Seq2Seq	0.0012	0.0029	0.0004	0.0019	0.0001	0.0022	0.0006	0.0001	0.0006
	Variational Seq2Seq	0.0006	0.0018	0.0002	0.0002	0.0001	0.0013	0.0007	0.0001	0.0002
	VRMNT	0.0002	0.0001	0.0053	0.0005	0.0018	0.0022	0.0005	0.0001	0.0004
	CWVAE-Unpretrained	0.0019	0.0036	0.0119	<b>0.0046</b>	0.0021	0.0013	0.0018	0.0005	0.0006
	CWVAE	<b>0.0055</b>	<b>0.0045</b>	<b>0.0142</b>	0.0028	<b>0.0043</b>	<b>0.0040</b>	<b>0.0021</b>	<b>0.0030</b>	<b>0.0033</b>
dist-2	RNN-based Seq2Seq	0.0036	0.0081	0.0002	0.0018	0.0002	0.0006	0.0013	0.0001	0.0011
	Variational Seq2Seq	0.0013	0.0042	0.0001	0.0003	0.0002	0.0026	0.0002	0.0003	0.0006
	VRMNT	0.0002	0.0011	0.0002	0.0005	0.0001	0.0034	0.0005	0.0001	0.0004
	CWVAE-Unpretrained	0.0060	0.0088	0.0136	<b>0.0113</b>	0.0043	0.0029	0.0041	0.0011	0.0009
	CWVAE	<b>0.0162</b>	<b>0.0112</b>	<b>0.0146</b>	0.0072	<b>0.0013</b>	<b>0.0107</b>	<b>0.0044</b>	<b>0.0068</b>	<b>0.0093</b>

## On Event2Mind

- 准确性:

Metric	Methods	xIntent	xReact	oReact
PPL	RNN-based Seq2Seq	44.12	29.18	14.08
	Variational Seq2Seq	42.06	28.22	12.62
	VRMNT	33.45	25.54	11.93
	CWVAE-Unpretrained	31.32	24.07	11.37
	CWVAE	<b>29.23</b>	<b>23.17</b>	<b>11.04</b>
BLEU	RNN-based Seq2Seq	2.75	2.11	5.18
	Variational Seq2Seq	2.84	2.43	2.08
	VRMNT	3.94	4.81	6.61
	CWVAE-Unpretrained	5.52	7.36	5.33
	CWVAE	<b>5.65</b>	<b>12.98</b>	<b>6.97</b>

- 多样性:

Metric	Methods	xIntent	xReact	oReact
dist-1	RNN-based Seq2Seq	0.0002	0.0002	0.0001
	Variational Seq2Seq	0.0006	0.0003	0.0001
	VRMNT	0.0002	0.0002	0.0003
	CWVAE-Unpretrained	0.0023	0.0017	0.0004
	CWVAE	<b>0.0052</b>	<b>0.0033</b>	<b>0.0025</b>
dist-2	RNN-based Seq2Seq	0.0005	0.0002	0.0002
	Variational Seq2Seq	0.0014	0.0002	0.0001
	VRMNT	0.0005	0.0003	0.0001
	CWVAE-Unpretrained	0.0061	0.0040	0.0013
	CWVAE	<b>0.0146</b>	<b>0.0099</b>	<b>0.0063</b>





# 事理图谱应用——脚本事件预测

- Script (**脚本**) in artificial intelligence (Schank and Abelson, 1977)
  - 描述特定场景/上下文的事件序列

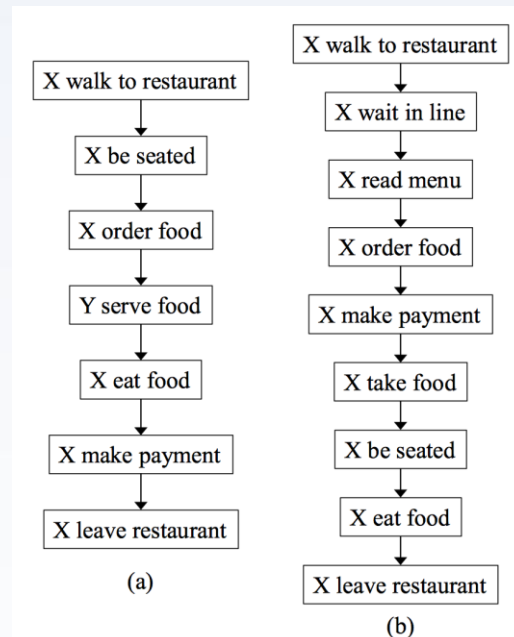
- 例子：餐厅脚本

场景1：进入餐馆

场景2：点单

场景3：吃饭

场景4：离开餐馆

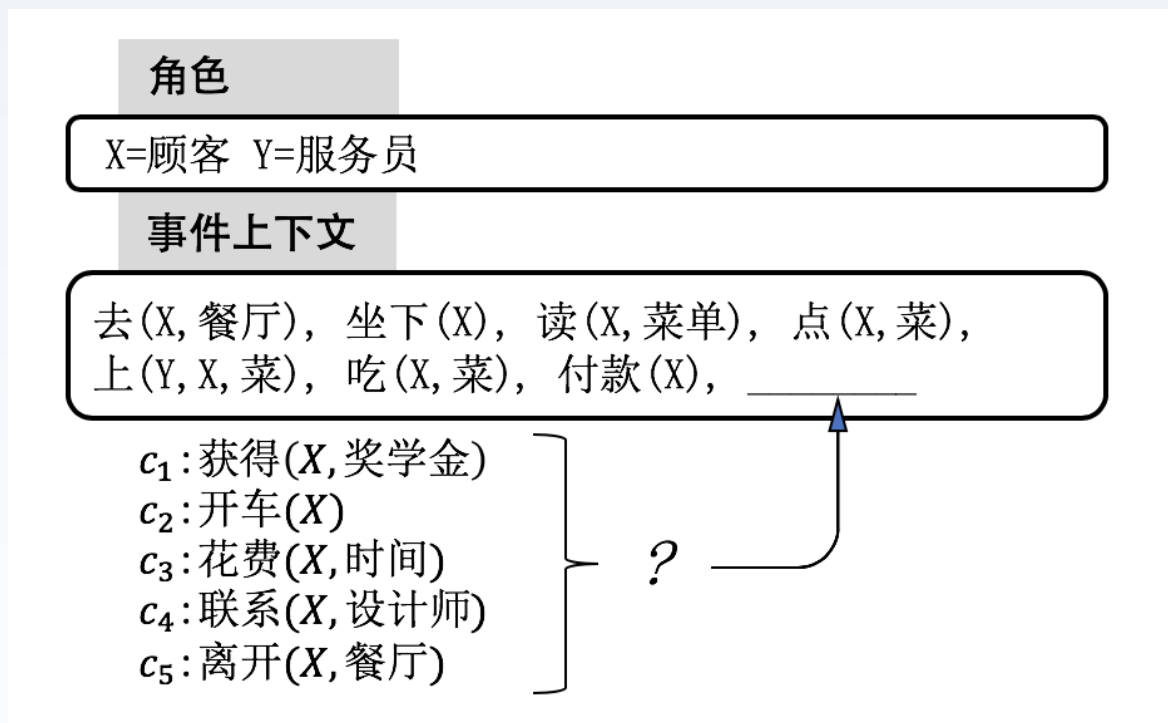




# 事理图谱应用——脚本事件预测

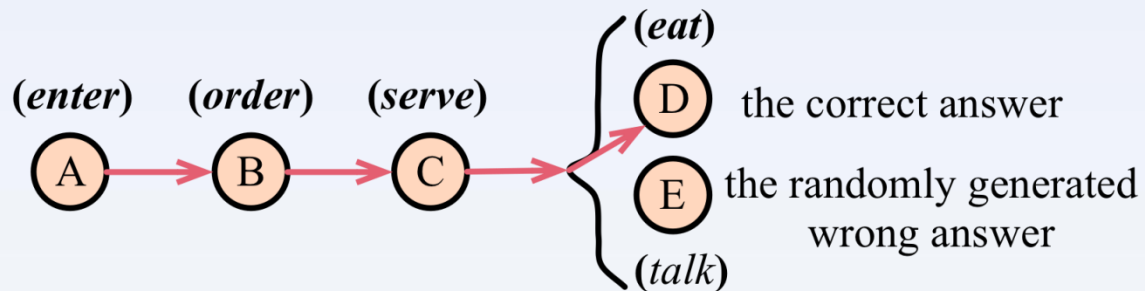
## • 任务

- 给定事件上文，从候选事件列表中选出接下来最有可能发生的事件

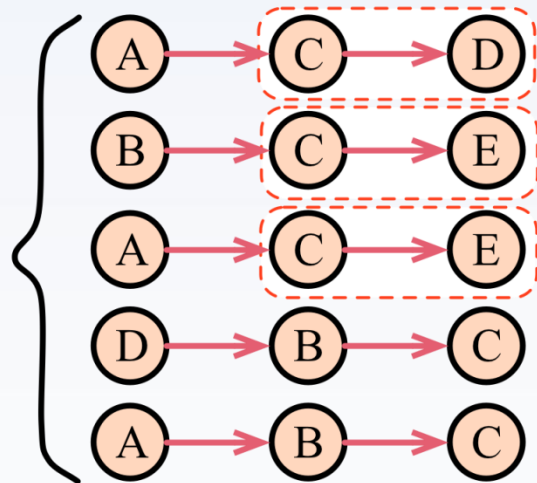


# 基于事理图谱的脚本事件预测研究动机

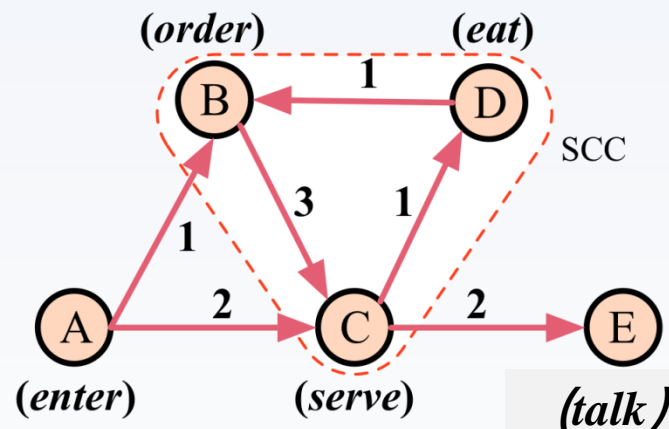
- 前人工作探索了如何利用event pair和event chain进行预测，我们的工作探索如何利用Eventic Graph的稠密连接信息来帮助事件预测
- 图结构学习到的事件表示更加利于预测
- 图结构能克服事件之间的不连通性，学习更好的事件关联



(a) Given an event context (A, B, C), choose the subsequent event from D and E.



(b) Training event chains.

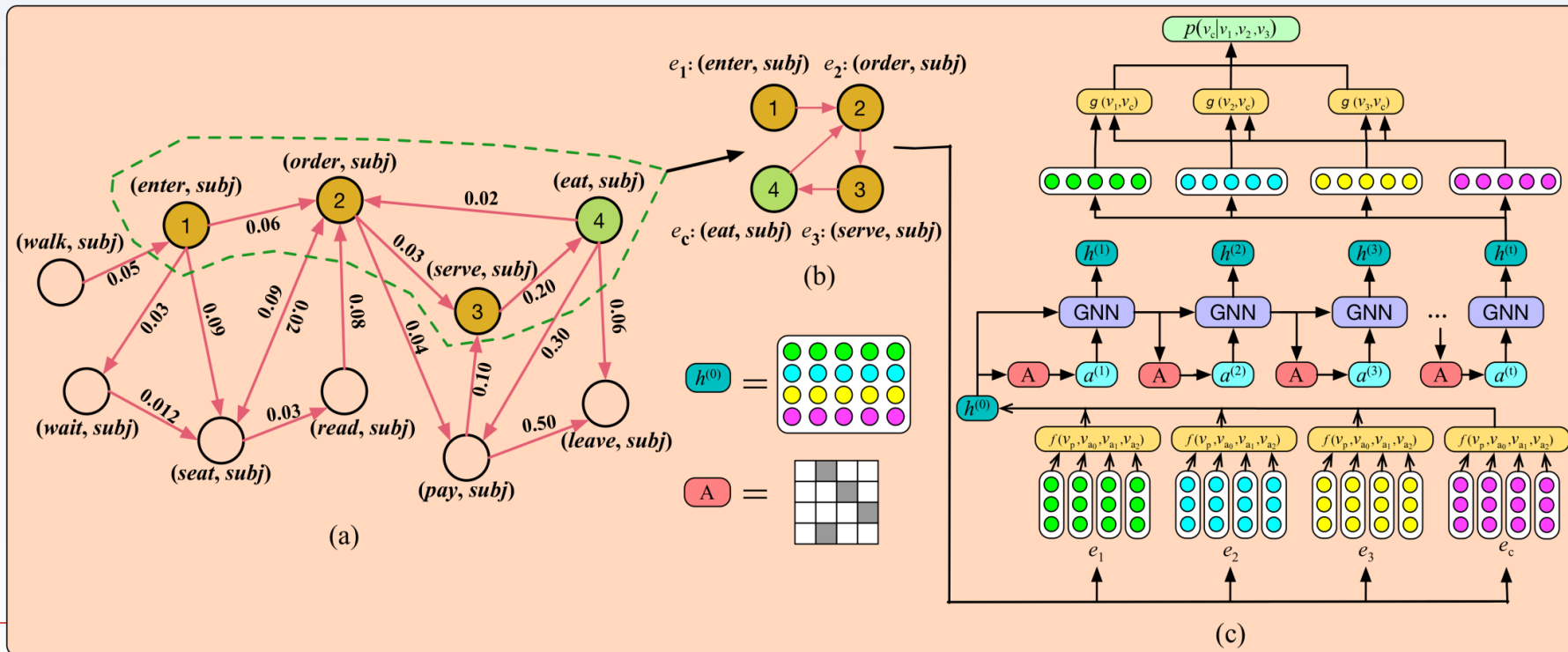


(c) Narrative event graph based on event chains in (b).

# 基于可扩展的图神经网络脚本事件预测方法

- 我们提出了可扩展的图神经网络(Scaled Graph Neural Network-SGNN)来解决图上的预测、推理问题
  - 重点解决了GNN在大规模图结构上无法进行推理的问题

步骤一：基于自动抽取的叙事事件链条构建叙事推理图谱



步骤二：提出可扩展图神经网络进行叙事图谱网络表示学习，用以解决大规模图上的推理问题





# 步骤一：构建叙事事理图谱

- 抽取叙事脚本事件链条
  - 基于前人提出的抽取步骤
- 采用一种合适的事件表示方式
  - Predicate-GR: (eats, subj)
  - 动词和它与事件链参与者的语法依存关系
  - 是很多具体事件的一种抽象表示形式
- 构建叙事事理图谱
  - 把事件链中的每个事件Bigram当做一条有向边，并计算事件转移概率

$$w(\mathbf{v}_j | \mathbf{v}_i) = \frac{\text{count}(\mathbf{v}_i, \mathbf{v}_j)}{\sum_k \text{count}(\mathbf{v}_i, \mathbf{v}_k)}$$

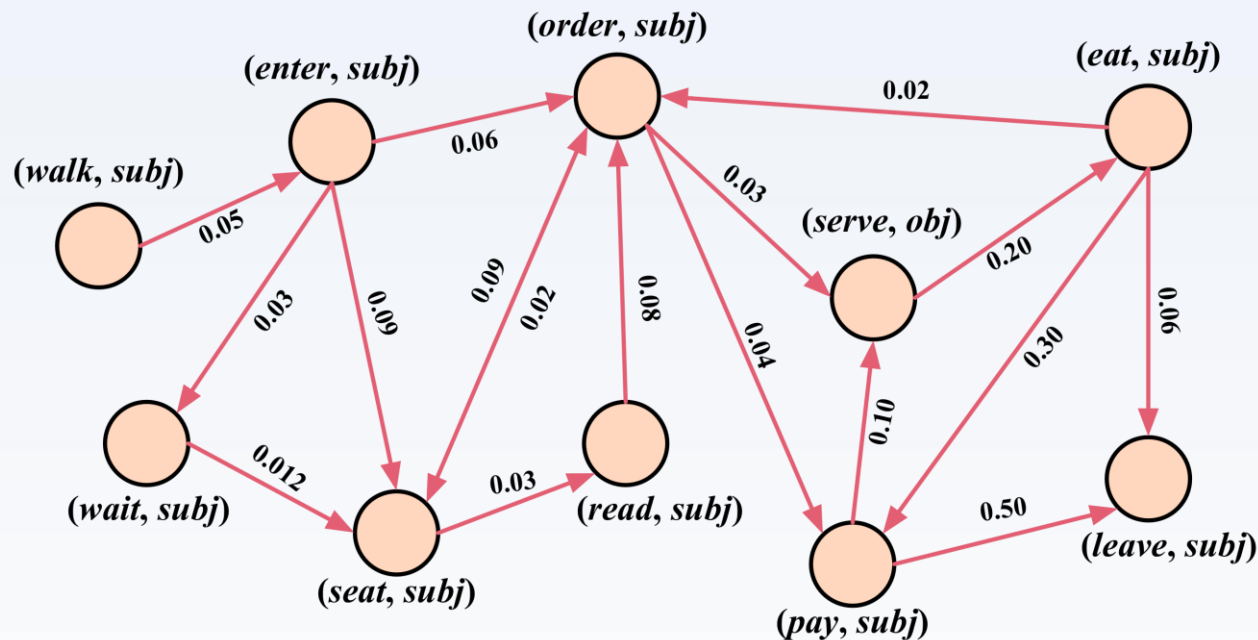






# 步骤一：构建完成的叙事事理图谱

- 构建完成的叙事事理图谱包含
  - 10万个节点
  - 618万条有向边





# 步骤二：网络表示学习—学习初始事件表示

## • 学习初始事件表示

- 给出一个事件  $e_i = \{p(a_0, a_2, a_2)\}$  以及动词和各事件元素的词向量： $v_p, v_{a_0}, v_{a_1}, v_{a_2} \in \mathbb{R}^d$ ，得到整个事件的表示： $v_{e_i} = f(v_p, v_{a_0}, v_{a_1}, v_{a_2})$

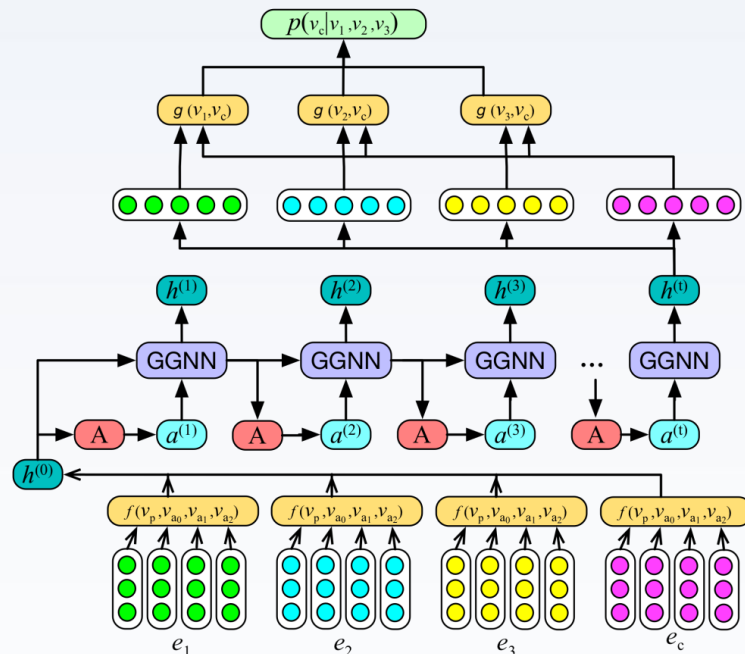
- **Average:** Use the mean value of the verb and all arguments vectors as the representation of the whole event.

- **Nonlinear Transformation** [Wang *et al.*, 2017]:

$$v_e = \tanh(W_p \cdot v_p + W_0 \cdot v_{a_0} + W_1 \cdot v_{a_1} + W_2 \cdot v_{a_2} + b) \quad (2)$$

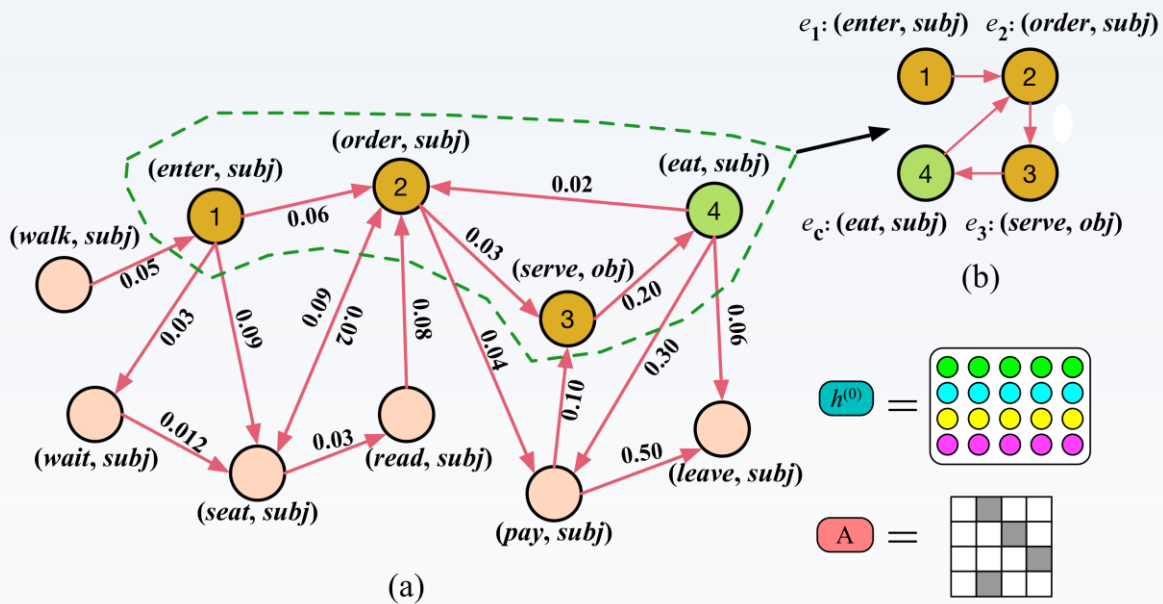
where  $W_p, W_0, W_1, W_2, b$  are model parameters.

- **Concatenation** [Granroth-Wilding and Clark, 2016]: Concatenate the verb and all argument vectors as the representation of the whole event.



# 步骤二：网络表示学习—输入

- 通过SGNN更新事件表示（基于GGNN, Li et al. 2015改进得到）
  - SGNN的输入包括两个矩阵 $h^{(0)}$  和  $A$



$h^{(0)}$ 表示事件上文以及所有候选事件的向量表。

$A$ 是相应的子图邻接矩阵：

$$A[i, j] = \begin{cases} w(\mathbf{v}_j | \mathbf{v}_i), & \text{if } \mathbf{v}_i \rightarrow \mathbf{v}_j \in E, \\ 0, & \text{others.} \end{cases}$$



# 步骤二：网络表示学习—更新参数

- 通过SGNN更新事件表示（基于GGNN, Li et al. 2015改进得到）
  - SGNN的更新部分与GGNN和GRU的门控制类似

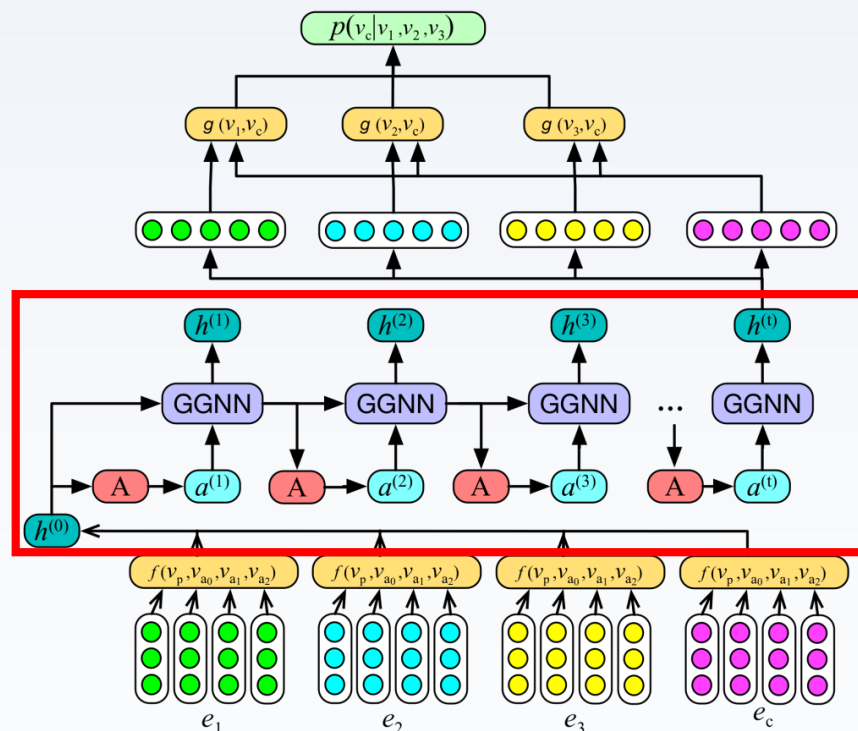
$$a^{(t)} = \mathbf{A}^\top h^{(t-1)} + b$$

$$z^{(t)} = \sigma(W^z a^{(t)} + U^z h^{(t-1)})$$

$$r^{(t)} = \sigma(W^r a^{(t)} + U^r h^{(t-1)})$$

$$c^{(t)} = \tanh(W a^{(t)} + U(r^{(t)} \odot h^{(t-1)}))$$

$$h^{(t)} = (1 - z^{(t)}) \odot h^{(t-1)} + z^{(t)} \odot c^{(t)}$$



# 步骤二：预测正确事件结尾

## • 挑选正确的后续事件

第*i*个上文事件与第*j*个候选事件的相关性得分计算：

$$s_{ij} = g(h_i^{(t)}, h_{c_j}^{(t)})$$

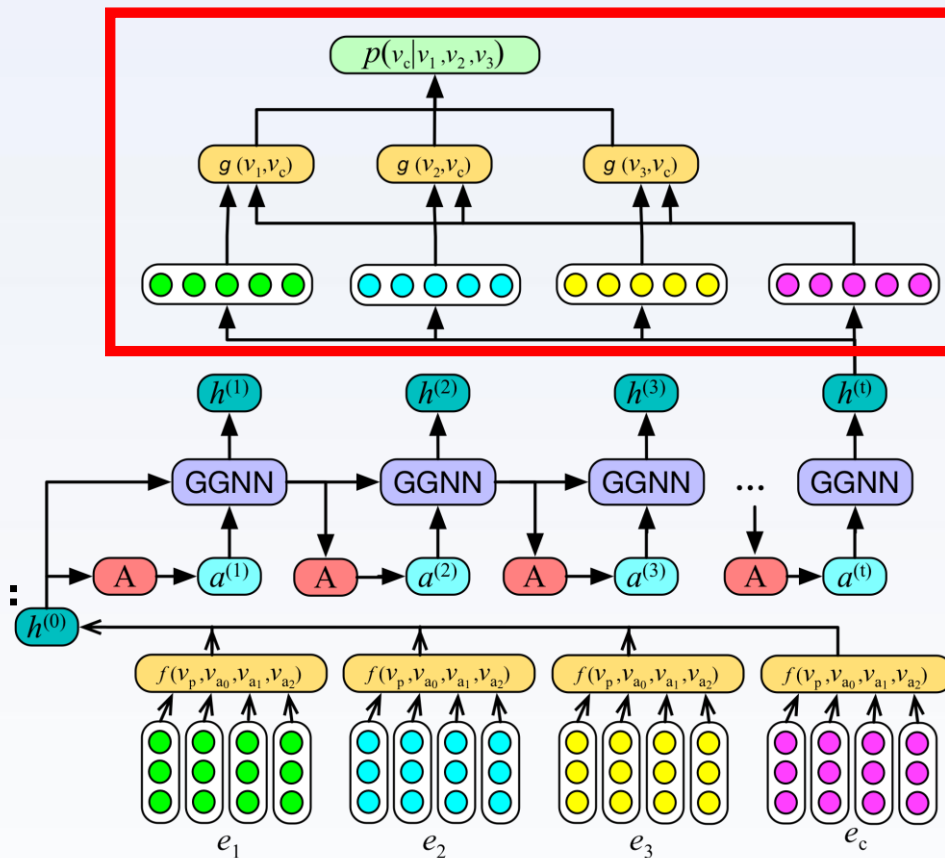
给出上文  $e_1, e_2, \dots, e_n$ ，候选事件  $e_{c_j}$  的可能性打分为：

$$s_j = \frac{1}{n} \sum_{i=1}^n s_{ij}$$

然后通过选择得分最高的候选事件作为正确答案：

$$c = \max_j s_j$$

损失函数： 
$$L(\Theta) = \sum_{I=1}^N \sum_{j=1}^k (\max(0, margin - s_{Iy} + s_{Ij})) + \frac{\lambda}{2} \|\Theta\|^2$$



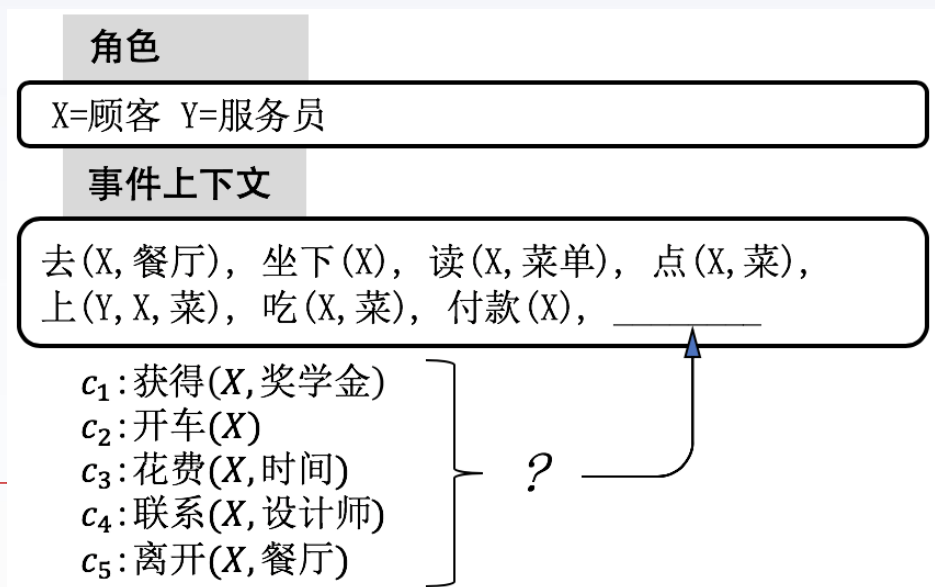


# 实验数据

- 数据集: Gigaword语料NYT部分

	Training	Development	Test
#Documents	830,643	103,583	103,805
#Chains for NEEG	5,997,385	-	-
#Chains for SGNN	140,331	10,000	10,000

- 评估: Multiple Choice Narrative Cloze (MCNC)





# 实验结果

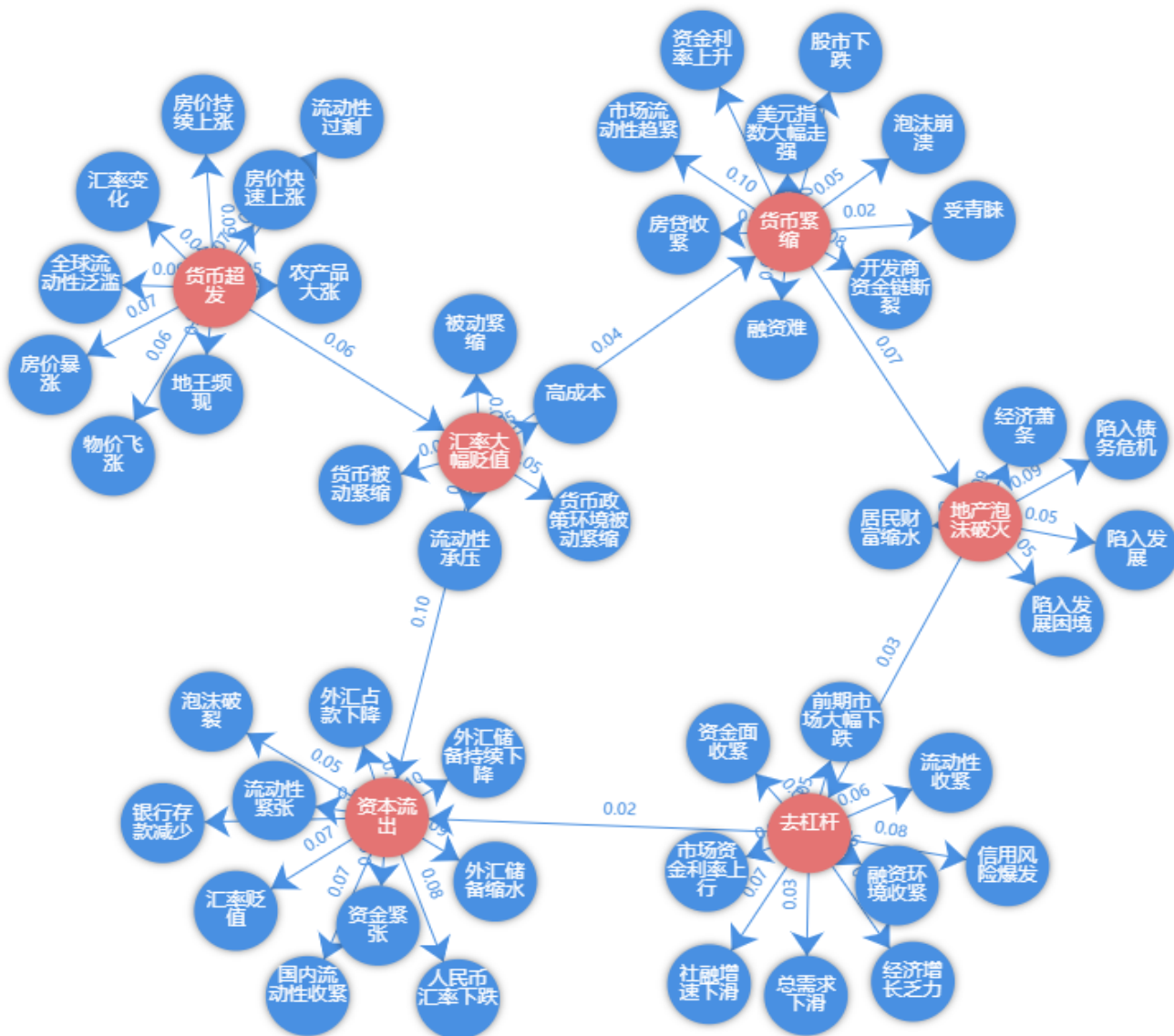
Methods	Accuracy
Random	20.00
PMI [Chambers and Jurafsky, 2008]	30.52
Bigram [Jans <i>et al.</i> , 2012]	29.67
Word2vec [Mikolov <i>et al.</i> , 2013]	42.23
DeepWalk [Perozzi <i>et al.</i> , 2014]	43.01
EventComp [Granroth-Wilding and Clark, 2016]	49.57
PairLSTM [Wang <i>et al.</i> , 2017]	50.83
SGNN-attention (without attention)	51.56
SGNN (ours)	<b>52.45</b>
SGNN+PairLSTM	52.71
SGNN+EventComp	54.15
SGNN+EventComp+PairLSTM	<b>54.93</b>





# 金融事理图谱样例

<http://eeg.8wss.com/>









# 金融事理图谱V2.0版本

	金融事理图谱v1.0	金融事理图谱v2.0
数据源	北京语言大学新闻语料、腾讯、网易、股吧、和讯、Resset等网站的金融新闻	北京语言大学新闻语料、腾讯、网易、股吧、和讯、Resset、 <b>新浪、中财网、中金在线、证券之星</b> 等网站的金融新闻、 <b>金融领域微信公众号、研报</b>
数据规模（文档数）	11,653,062	12,121,279（ <b>增加100余万篇文档</b> ）
具体事件数	1,348,459	2,187,086（ <b>增加80余万个事件</b> ）
具体因果系数	1,410,642	1,607,638（ <b>增加近20万个因果关系</b> ）
抽象事件数	-	2,021,289
抽象因果系数	-	414,777
上下位关系数	-	1,883,792
因果关系抽取F值	59.54%	85.12%（ <b>提升近16%</b> ）





# NEEL: Never-Ending Eventics Learning System

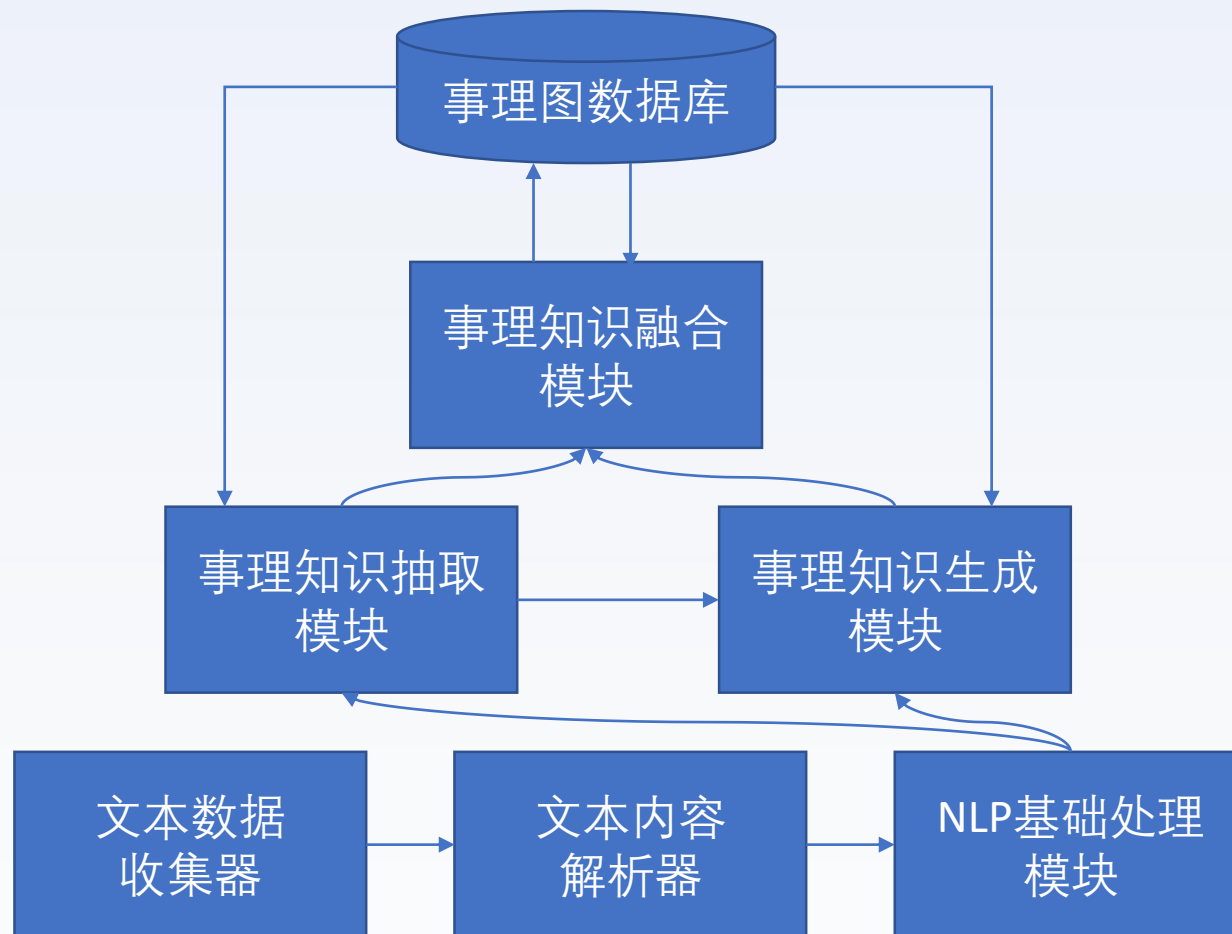
<http://openeg.8wss.com/main>

事理知识存储层

事理知识融合层

事理知识学习层

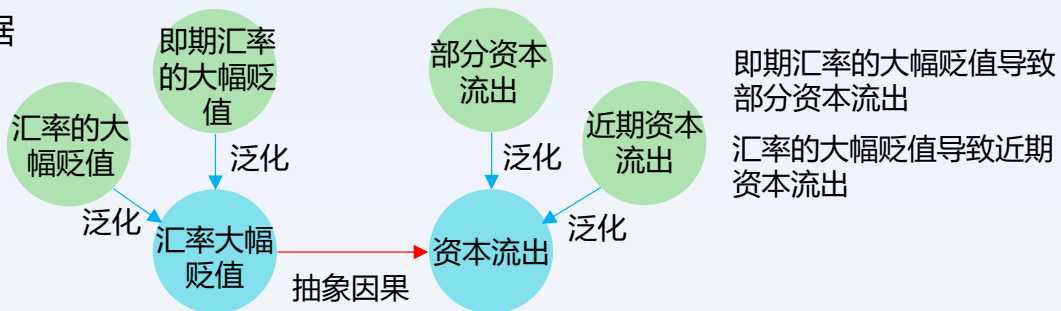
底层基础处理层



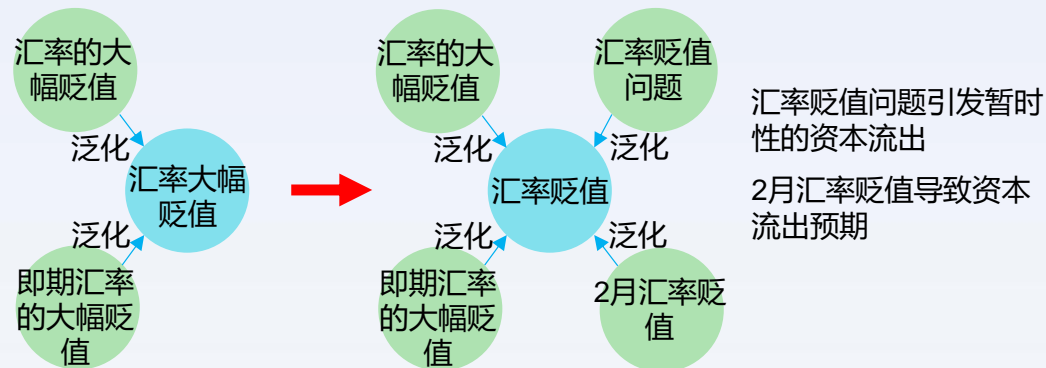


# 事理知识融合模块

原有数据

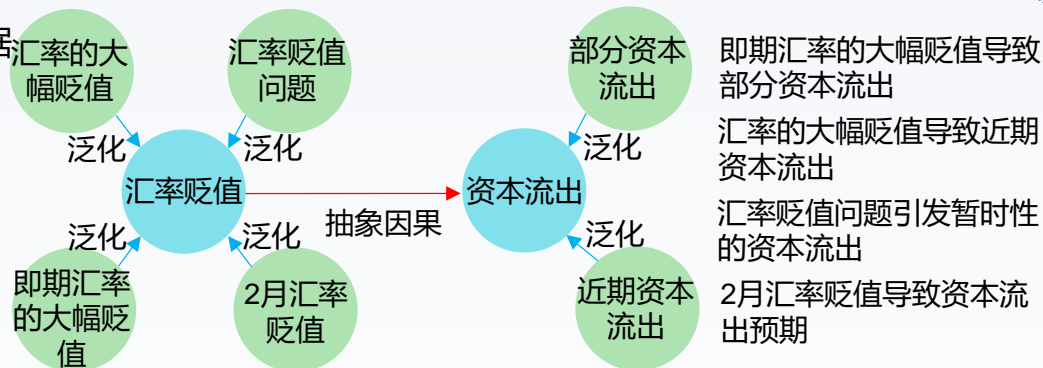


新增数据



事件泛化

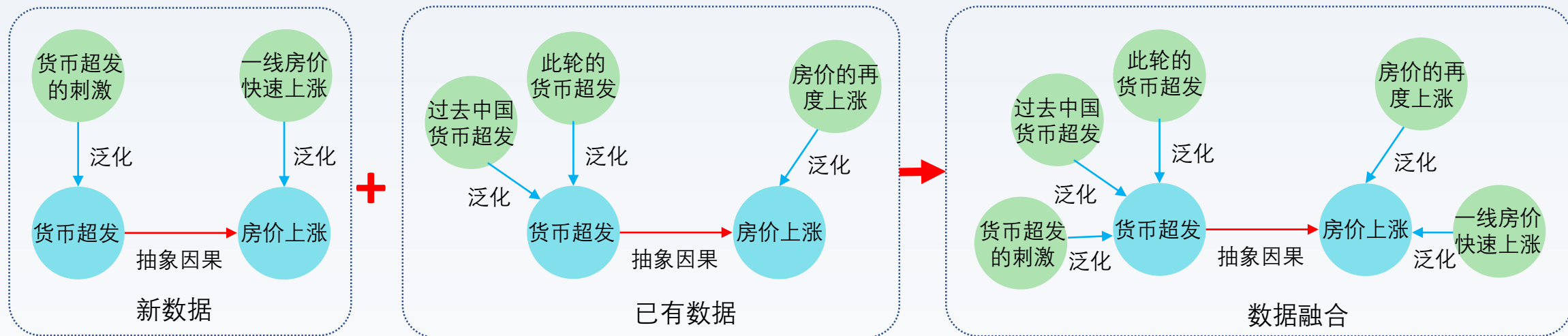
融合后数据





# 事理知识融合模块

## 数据融合





# NEEL: Never-Ending Eventics Learning System

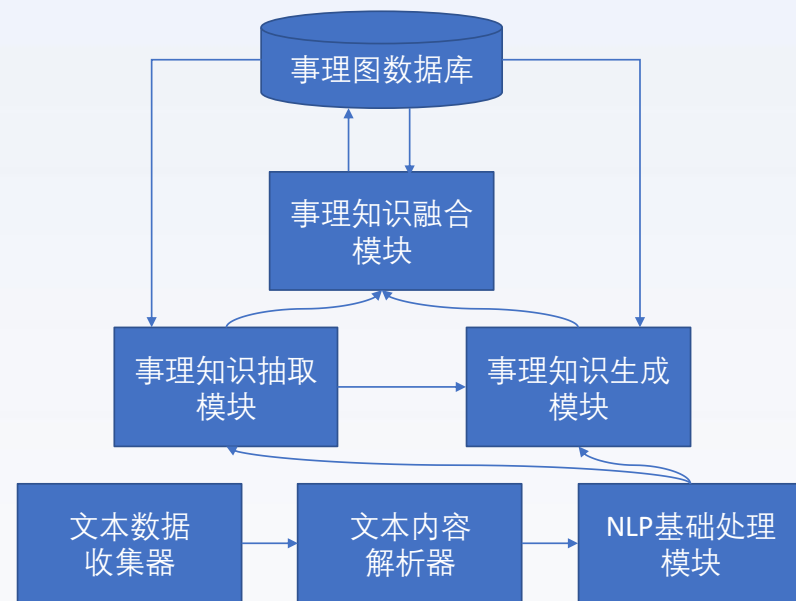
	NEEL
URL	<a href="http://openeg.8wss.com/">http://openeg.8wss.com/</a>
数据源	北京语言大学新闻语料、北京日报网、央广网、中国政务网、共产日报、光明日报、环球日报、北方新闻等20余个来源，仍在持续补充
数据规模（文档数）	8,538,575
具体事件数	1,579,609
具体因果系数	1,148,814
抽象事件数	1,241,229
抽象因果系数	230,157
上下位关系数	1,034,916

事理知识存储层

事理知识融合层

事理知识学习层

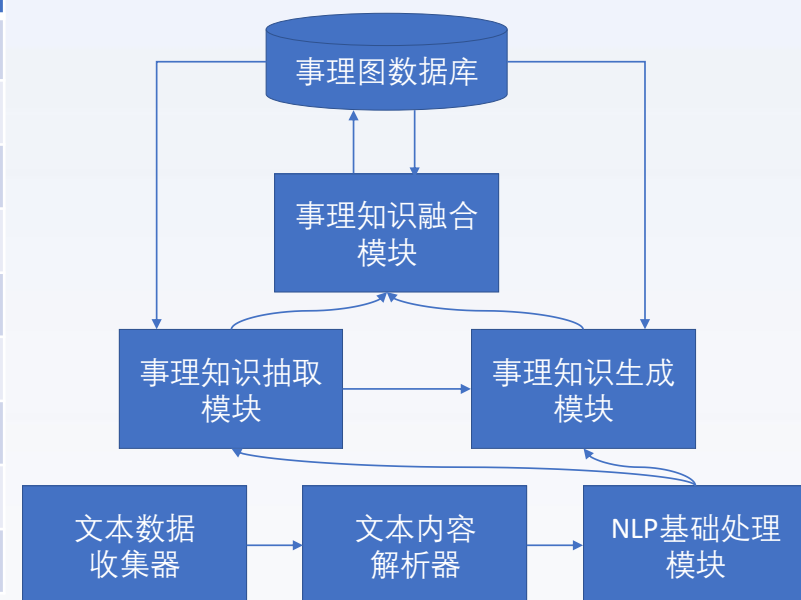
底层基础处理层





# NEEL: Never-Ending Eventics Learning System

日期	文档数	具体事件	具体因果边	抽象事件	抽象因果边	上下位边
2019-11-18	+399	+137	+71	+159	+26	+106
2019-11-19	+536	+196	+103	+196	+20	+133
2019-11-20	+644	+234	+125	+235	+28	+157
2019-11-21	+737	+269	+140	+280	+31	+199
2019-11-22	+955	+378	+197	+393	+47	+266
2019-11-23	+1122	+385	+199	+400	+53	+265
2019-11-24	+1206	+455	+235	+506	+74	+341
合计	+5599	+2054	+1070	+2169	+279	+1467
平均	+799.86	+293.43	+152.86	+309.86	+39.86	+209.57





# 事理图谱的潜在应用

## • 消费意图识别与推荐

- 几乎所有事件都可以和“想”、“打算”、“计划”连起来形成一个意图事件
  - 想去北京，想看电影，计划去爬泰山，想吃火锅
- 部分节点是消费意图显著节点，可以触发一系列后续消费事件，找到这类节点是很有价值的
  - 逛街，看电影，去旅游等等







# 事理图谱的潜在应用

## • 辅助对话生成

- 去北京→买票
- A: 我打算去北京玩。 B: 那你买票了吗?
- 去泰安→爬泰山
- A: 我想去爬泰山。 B: 那你得先到泰安。

## • 问答系统

- 爬泰山有什么注意事项?
- 租大衣、带水、带手电等等





# 事理图谱研究基础

- Zhongyang Li, Xiao Ding, Ting Liu. Story Ending Prediction by Transferable BERT. In Proc. of **IJCAI**, 2019.
- Xiao Ding, Kuo Liao, Ting Liu, Junwen Duan, Zhongyang Li. Event Representation Learning Enhanced with External Commonsense Knowledge. **EMNLP**, 2019.
- Li Du, Xiao Ding, Ting Liu, Zhongyang Li. Modeling Event Background for If-Then Commonsense Reasoning Using Context-aware Variational Autoencoder. **EMNLP**, 2019.
- Zhongyang Li, Xiao Ding, Ting Liu. Constructing Narrative Event Evolutionary Graph for Script Event Prediction. In Proc. of **IJCAI**, 2018: 4201-4207.
- Xiaocheng Feng, Lifu Huang, Duyu Tang, Bing Qin, Heng Ji, Ting Liu. A Language-Independent Neural Network for Event Detection. **ACL** 2016.
- Sendong Zhao, Quan Wang, Sean Massung, Bing Qin, Ting Liu, Chengxiang Zhai. Constructing and Embedding Abstract Event Causality Networks from Text Snippets. In Proc. of ACM **WSDM** 2017.
- Zhongyang Li, Sendong Zhao, Xiao Ding, Ting Liu. EEG: Knowledge Base for Event Evolutionary Principles and Patterns[C]. Chinese National Conference on Social Media Processing. Springer, Singapore, 2017: 40-52.
- Xiao Ding, Bing Qin, Ting Liu. BUEES: a Bottom-Up Event Extraction System. Frontiers of Information Technology & Electronic Engineering, 2015, 16(7): 541-552.
- Xiao Ding, Bing Qin, Ting Liu. *Building Chinese Event Type Paradigm Based on Trigger Clustering*. In Proc. of IJCNLP, 2013: 311-319.
- Sendong Zhao, Ting Liu, Sicheng Zhao, Yiheng Chen, Jian-Yun Nie. Event causality extraction based on connectives analysis. *Neurocomputing* (2016)173: 1943-1950.





# 总结

- 知识图谱在各个领域精耕细作，逐渐显露价值
- 知识表示形式有待突破，推理能力有待提高
- 以“谓词性短语”为节点，以事件演化（顺承、因果、上下位等）为边的事理图谱方兴未艾
- 事理图谱必将在预测、推荐、对话等领域发挥重要作用，有力地提升人工智能系统的可解释性
- 知识图谱与事理图谱的融合更有利推动相关研究和产业应用





# 谢谢!

