

基于事理图谱的文本推理技术

丁 效

哈尔滨工业大学

社会计算与信息检索研究中心

2020.7.3



7月3日-4日

CSDN

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



7月3日-4日

CSDN

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。

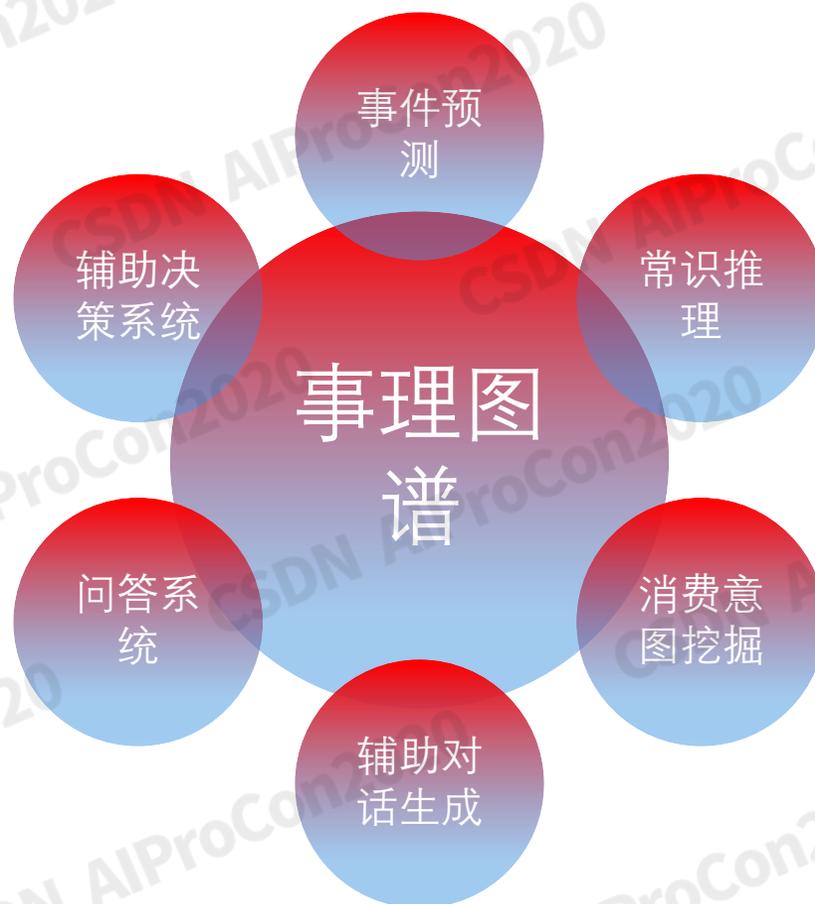


7月3日-4日

CSDN

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。



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

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



7月3日-4日

CSDN

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



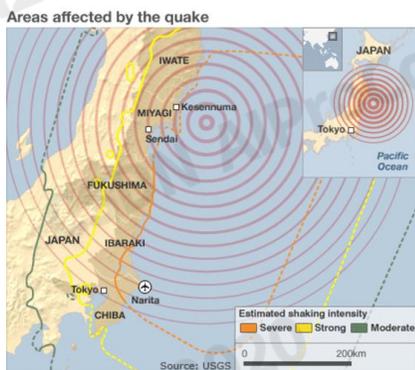
7月3日-4日

CSDN

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

百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

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**）



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

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。



7月3日-4日

CSDN

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。

核泄漏

因果

海洋污染

AI

2020
ProCon
万人开发者大会

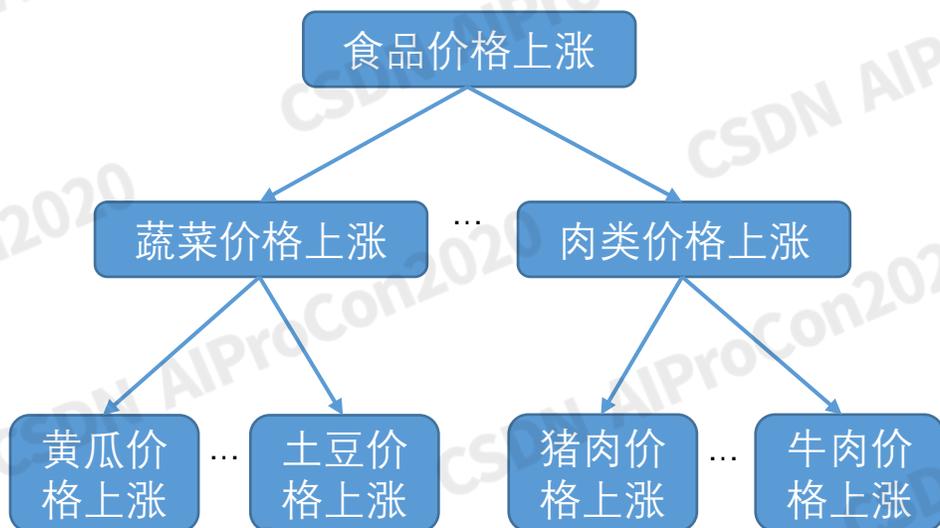
7月3日-4日

CSDN

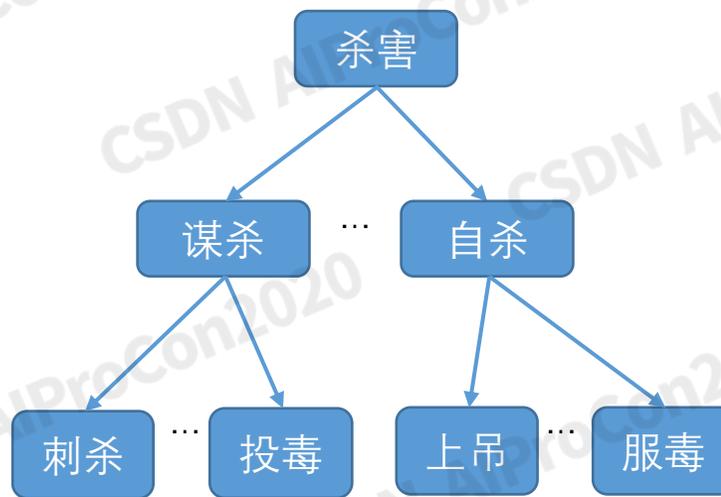
事件间上下位关系

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



7月3日-4日

CSDN

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。



7月3日-4日

CSDN

基于事理图谱的文本推理技术

1. 基于事件背景知识的IF-THEN类型常识推理
2. 基于事理图谱的脚本事件预测
3. 基于事理图谱的溯因推理
4. 基于事理图谱的因果推理
5. 大规模开放域因果生成



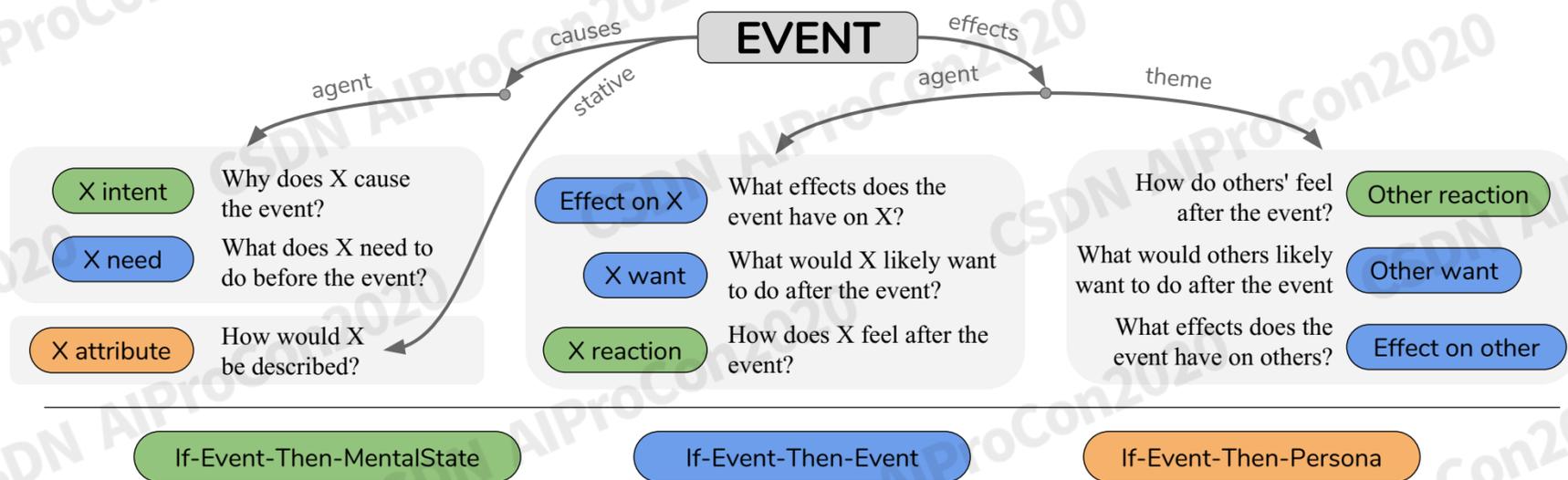
7月3日-4日

CSDN

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



Li Du, et al. Modeling Event Background for If-Then Commonsense Reasoning Using Context-aware Variational Autoencoder. In Proc. of EMNLP 2019



7月3日-4日

CSDN

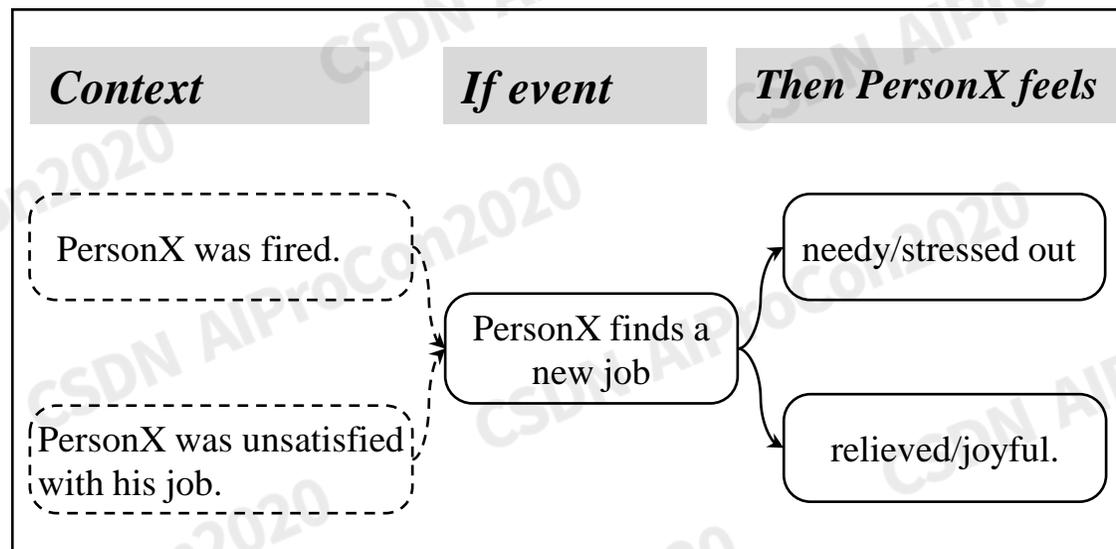
基于事件背景知识进行IF-THEN类型常识推理

- 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



7月3日-4日

CSDN

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



7月3日-4日

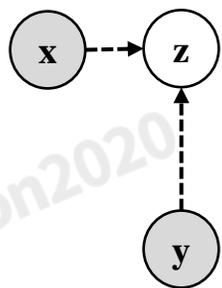
CSDN

上下文感知的变分自编码器和两阶段训练过程

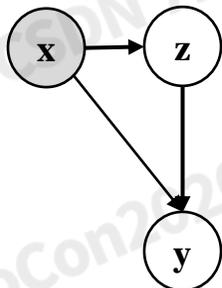
- 变分自编码器

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

(a) 训练



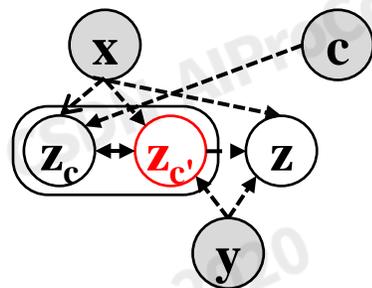
(b) 生成



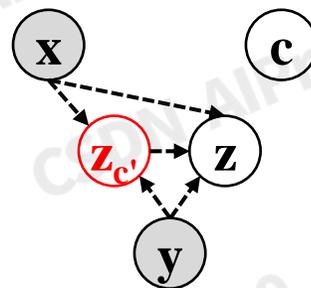
- 上下文感知的变分自编码器

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

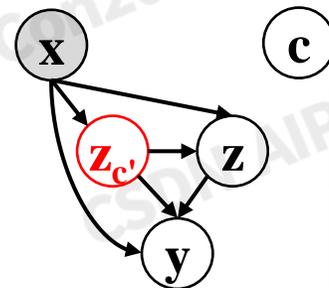
(a) 预训练



(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



7月3日-4日

CSDN

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	15.93	20.32	23.85	50.74	21.39	24.02	14.02	11.70	29.13
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	14.93
	CWVAE	12.12	15.67	5.63	14.64	8.13	15.01	13.83	8.58	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	0.0046	0.0021	0.0013	0.0018	0.0005	0.0006
	CWVAE	0.0055	0.0045	0.0142	0.0028	0.0043	0.0040	0.0021	0.0030	0.0033
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	0.0113	0.0043	0.0029	0.0041	0.0011	0.0009
	CWVAE	0.0162	0.0112	0.0146	0.0072	0.0013	0.0107	0.0044	0.0068	0.0093

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	29.23	23.17	11.04
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	5.65	12.98	6.97

- 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	0.0052	0.0033	0.0025
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	0.0146	0.0099	0.0063



7月3日-4日

CSDN

基于事理图谱的脚本事件预测

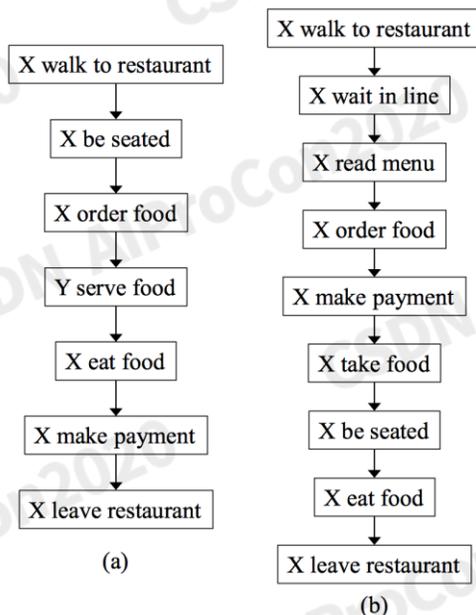
- Script (脚本) in artificial intelligence (Schank and Abelson, 1977)
 - 描述特定场景/上下文的事件序列
- 例子：餐厅脚本

场景1：进入餐馆

场景2：点单

场景3：吃饭

场景4：离开餐馆



Zhongyang Li, Xiao Ding, Ting Liu. Constructing Narrative Event Evolutionary Graph for Script Event Prediction. In Proc. of IJCAI, 2018.



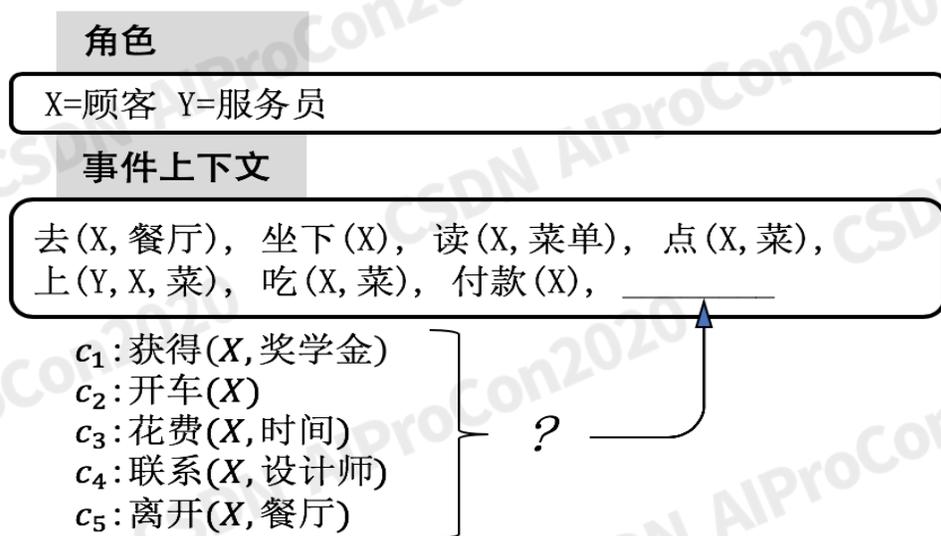
7月3日-4日

CSDN

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

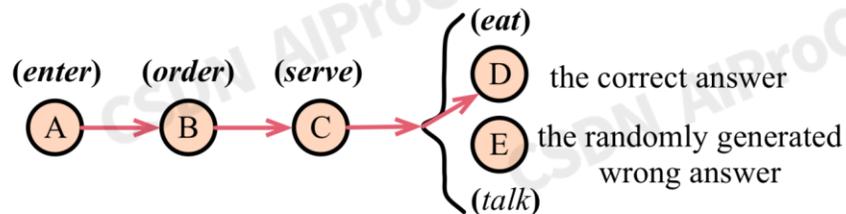


7月3日-4日

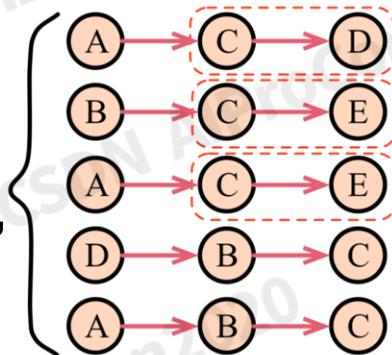
CSDN

基于事理图谱的脚本事件预测研究动机

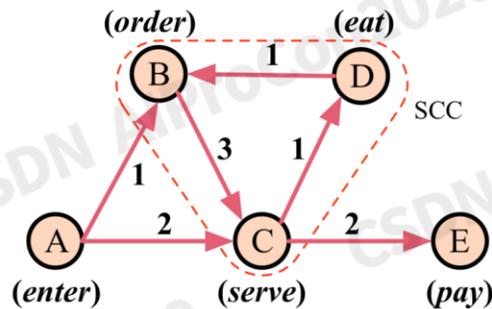
- 前人工作探索了如何利用event pair和event chain进行预测，我们的工作探索如何利用Eventic Graph的稠密连接信息来帮助事件预测
- 图结构学习到的事件表示更加利于预测
- 图结构能克服事件之间的不连通性，学习更好的事件关联



(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).



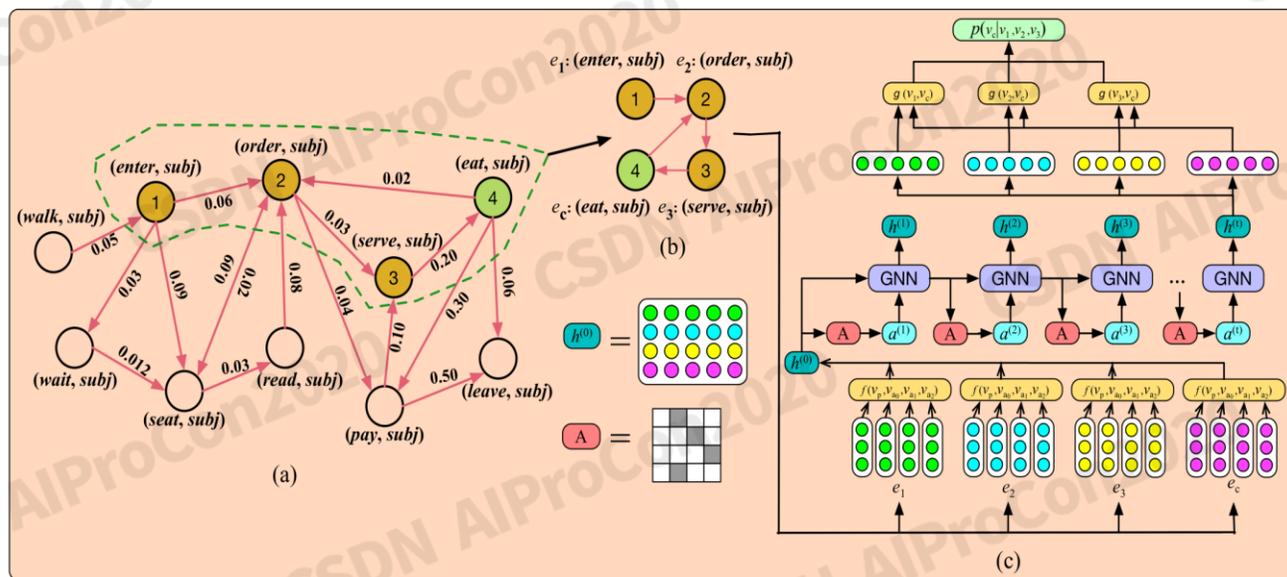
7月3日-4日

CS (talk)

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



7月3日-4日

CSDN

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

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



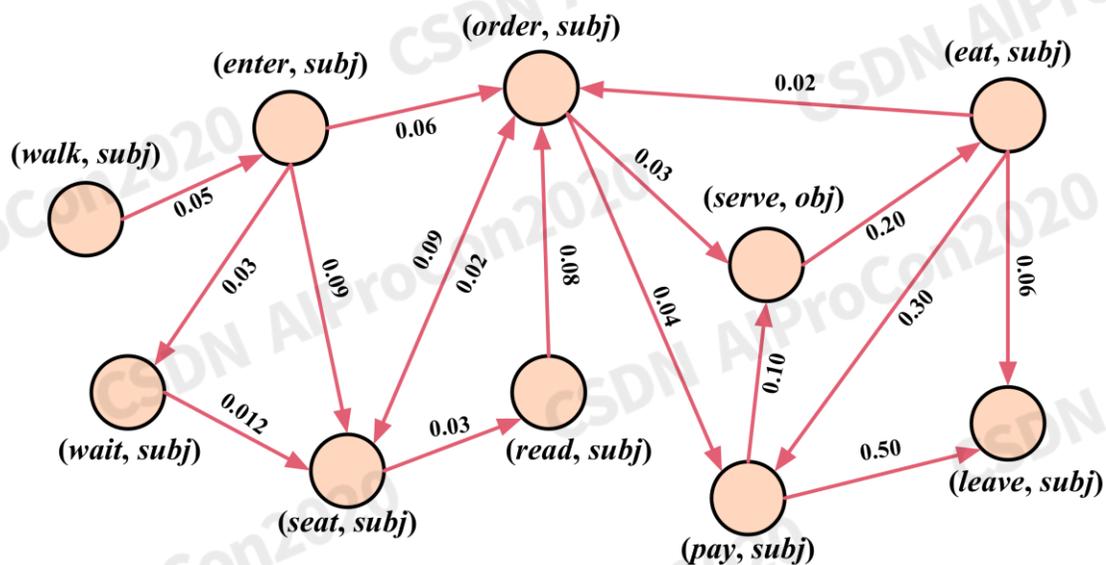
7月3日-4日

CSDN

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



7月3日-4日

CSDN

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 $e_i = \{p(a_0, a_1, a_2)\}$ 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： $v_{e_i} = f(v_p, v_{a_0}, v_{a_1}, v_{a_2})$

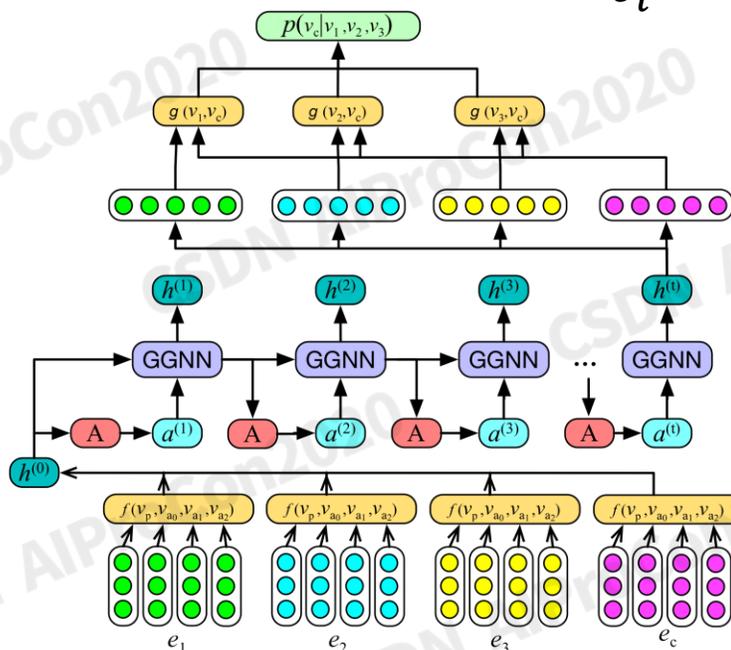
- **Average**: Use the mean value of the verb and all arguments vectors as the representation of the whole event.

- **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.



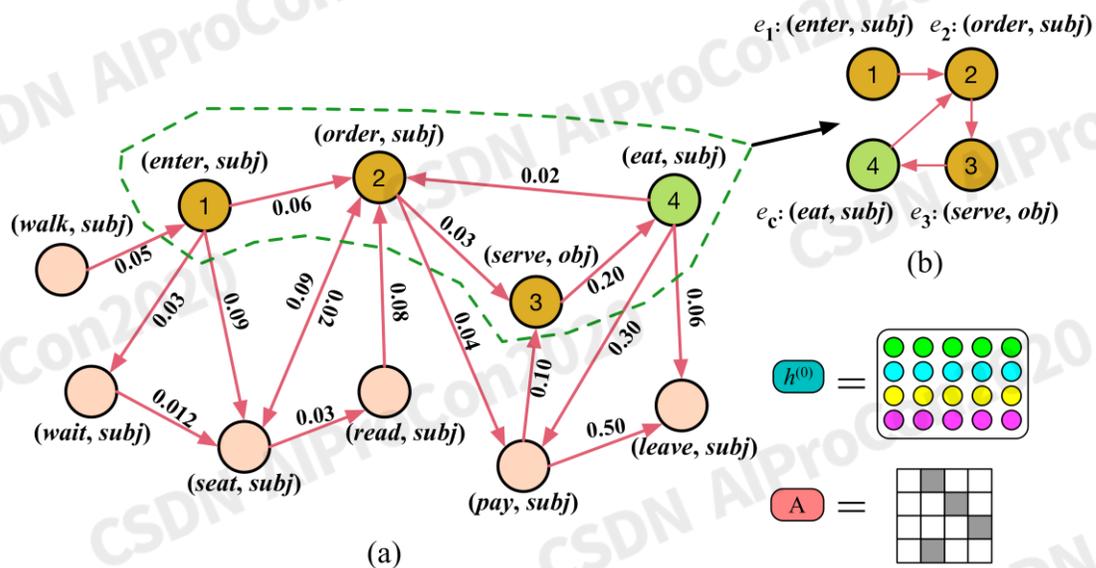
2020
ProCon
万人开发者大会

7月3日-4日

CSDN

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

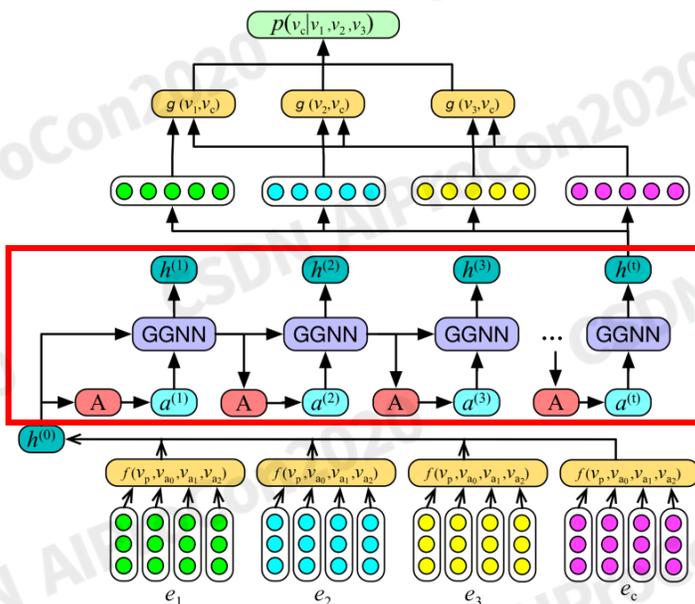
$$a^{(t)} = \mathbf{A}^\top h^{(t-1)} + b$$

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

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

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

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



2020
ProCon
万人开发者大会

7月3日-4日

CSDN

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：

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

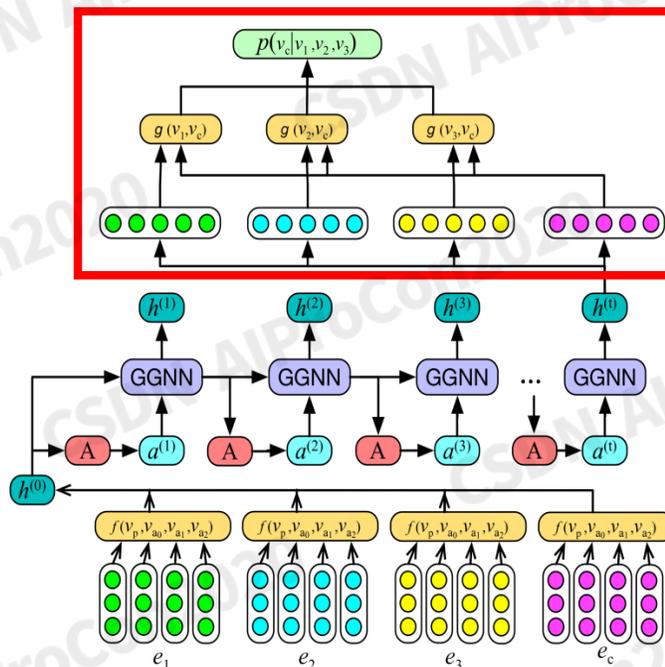
给出上文 e_1, e_2, \dots, e_n ，候选事件 e_{c_j} 的可能性打分为：

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

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

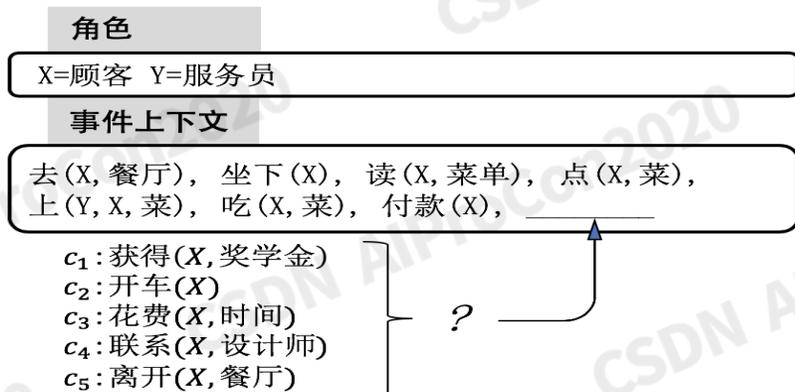


实验数据

- 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)



7月3日-4日

CSDN

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)	52.45
SGNN+PairLSTM	52.71
SGNN+EventComp	54.15
SGNN+EventComp+PairLSTM	54.93



7月3日-4日

CSDN

基于事理图谱的溯因推理

- 溯因推理任务:

- 给定已观测到的事件, 要求选择观测事件的可能原因

eg: O_1 : 离开时没关窗 \rightarrow O_2 : 屋子里一团糟

H_1 : 屋子里进了贼

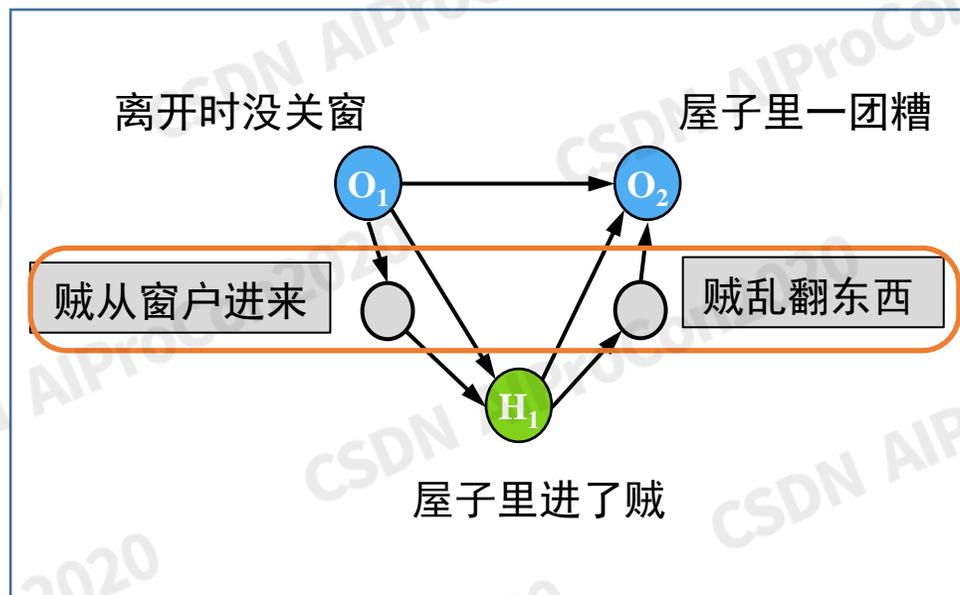
H_2 : 一阵微风吹来

- 溯因推理任务依赖于额外的常识知识以帮助理解事件间关系:

- 一阵微风很难导致屋子里一团糟,

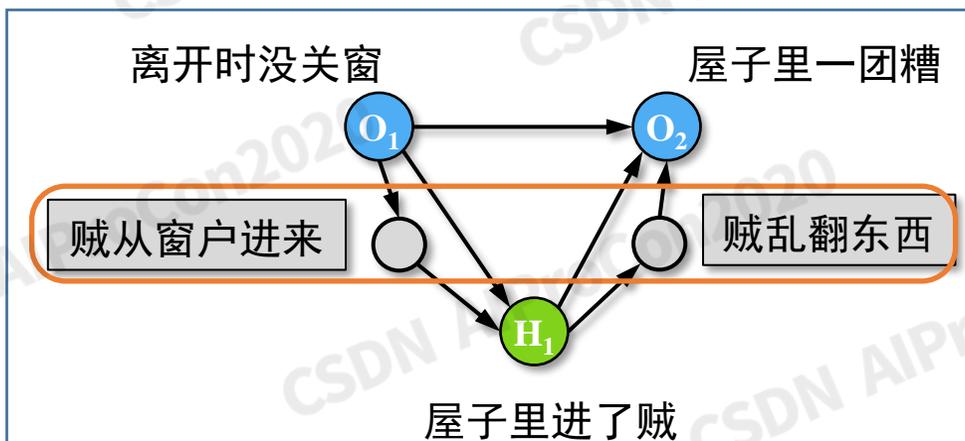
然而屋子里进了贼, 贼可能乱翻东西, 导致屋子里一团糟

- 事理图谱中可以获取此类常识知识

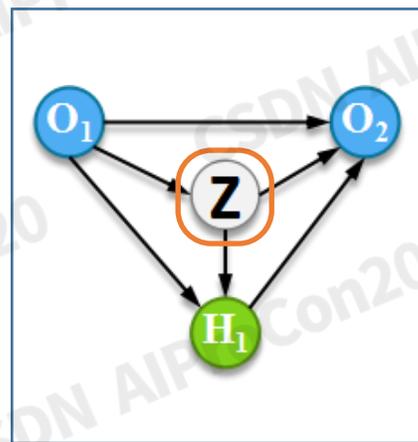


事理图知识增强的预训练模型

提出**事理图知识增强的预训练模型**(event-graph enhanced pretrained model, ege-RoBERTa)



事理图谱



ege-RoBERTa引入一个额外的隐变量 z 以捕获事理图谱知识

• 两阶段训练过程引入事件背景知识:

构建事理图谱

- 基于故事数据集构建事理图谱



预训练阶段

- 学习事理图谱知识



微调阶段

- 在标注数据集上进行微调

AI

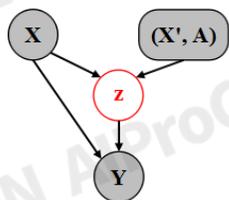
2020 ProCon 万人开发者大会

7月3日-4日

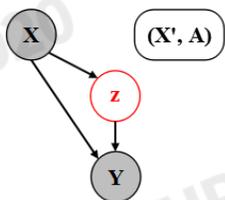
CSDN

事理图谱知识增强的预训练模型

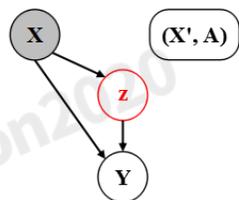
(a) 预训练



(b) 微调



(c) 生成



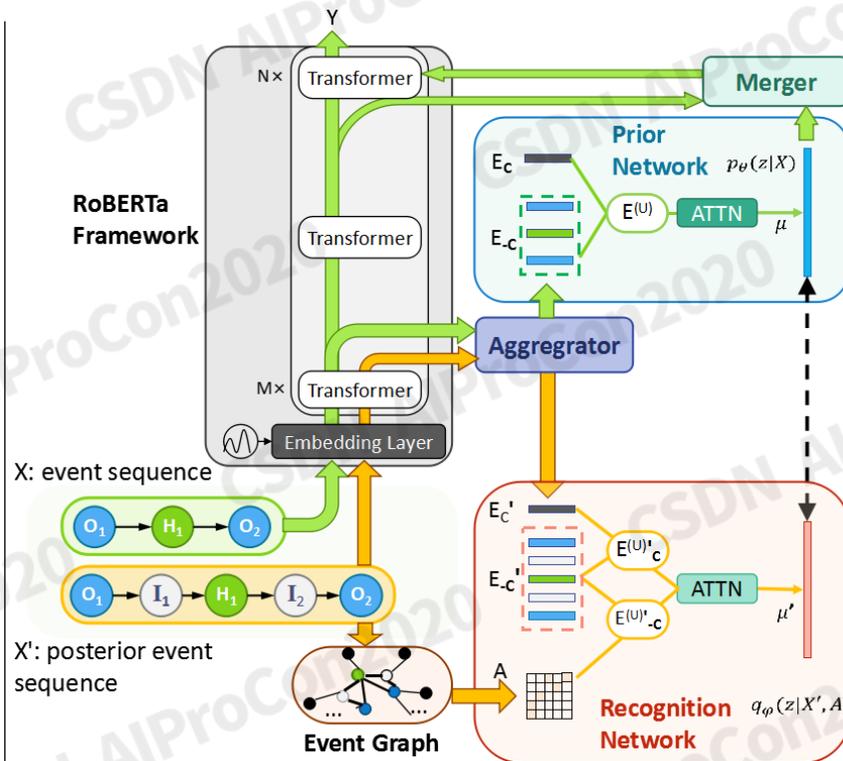
- X = 已观测事件 O +某一可能原因事件 H

- X' = 已观测事件 O +某一可能原因事件 H +**中间事件 I**

- A = 邻接矩阵, **描述 O, H, I 之间的邻接关系**

• 额外的常识知识:

中间事件 I +邻接矩阵 A



实验结果

- On aNLI dataset

Methods	Accuracy(%)
SVM	50.6
InferSent (Conneau et al., 2017)	50.8
ESIM+ELMo (Chen et al., 2017)	58.8
GPT (Radford et al., 2018)	63.1
BERT-base (Devlin et al., 2019)	63.3
RoBERTa-base (Devlin et al., 2019)	71.5
BERT-large (Devlin et al., 2019)	68.9
RoBERTa-large (Liu et al., 2019)	83.9
ege-RoBERTa-large _{unpretrained}	83.8
ege-RoBERTa-large _{$\lambda=0$}	84.2
ege-RoBERTa-base	75.9
ege-RoBERTa-large	86.0



7月3日-4日

CSDN

基于事理图谱的因果推理

- 人类陈述因果关系时往往不会陈述完整的因果推理链

因果对：【量化宽松】——【大宗商品价格上涨】

【降息】——【投资需求增加】

- 完整的因果链是

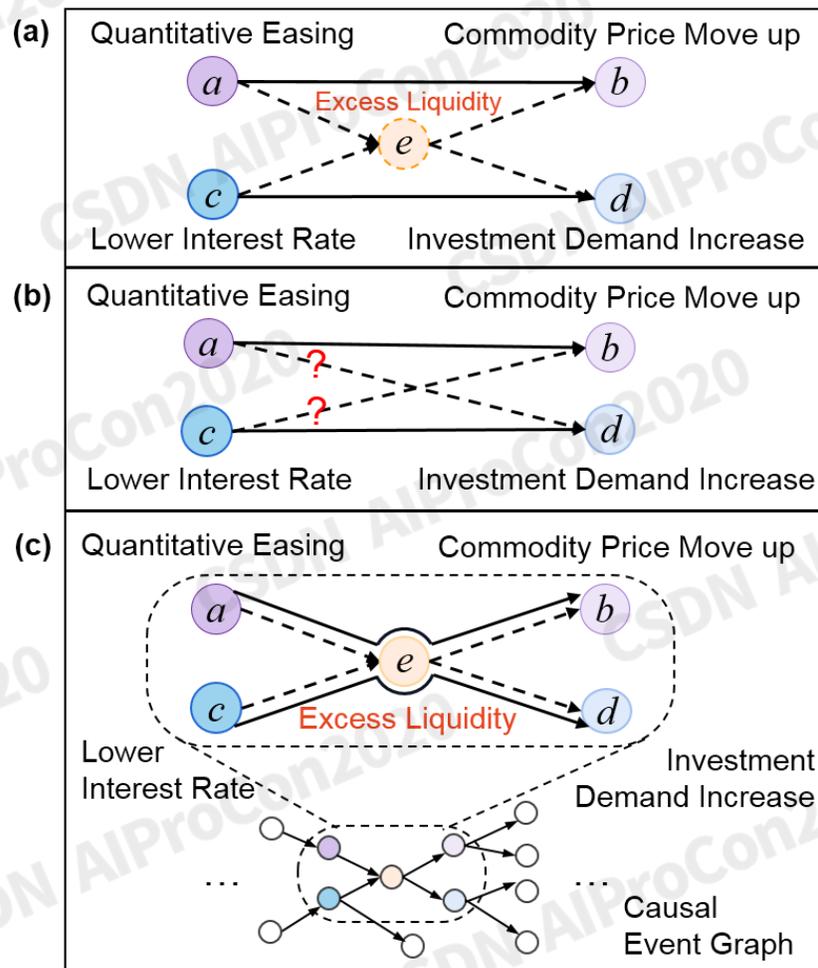
【量化宽松】—【货币超发】—【投资需求增加】

从因果对中难以直接学习到【原因】-【证据事件】-【结果】的逻辑链和推理过程

- 证据事件的忽略导致模型推理能力下降

- 基于事理图谱的因果推理

- 构建因果事理图谱
- 搜寻中间证据事件，形成证据图
- 还原逻辑链，基于逻辑链进行因果推理



AI

2020 ProCon 万人开发者大会

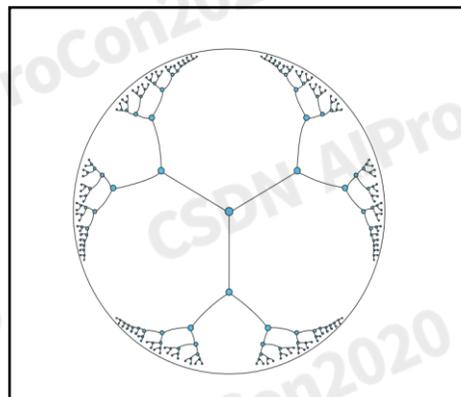
7月3日-4日

CSDN

挑战

- 因果事理图谱是树状图

- 一因多果、多因一果现象的大量存在导致因果事件图谱存在较明显的树状结构
- 在欧氏空间中进行树状图的表示与推理会带来结构的扭曲, 进而导致因果强度估计的误差

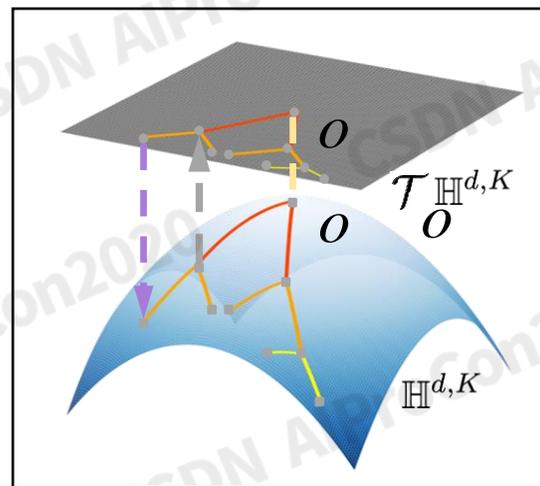


- 因果推理存在方向性

- 因果推理过程中需要引入方向性以避免混淆因果

- 有向的双曲图注意力机制

- 在双曲空间中进行节点表示与推理, 以减小因果强度估计误差
- 在图注意力机制中引入方向性



百万
人学AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

中间证据事件搜寻

- 基于文本相似度定位原因事件与结果事件

- 文本相似的事件可能具有类似的原因和结果事件

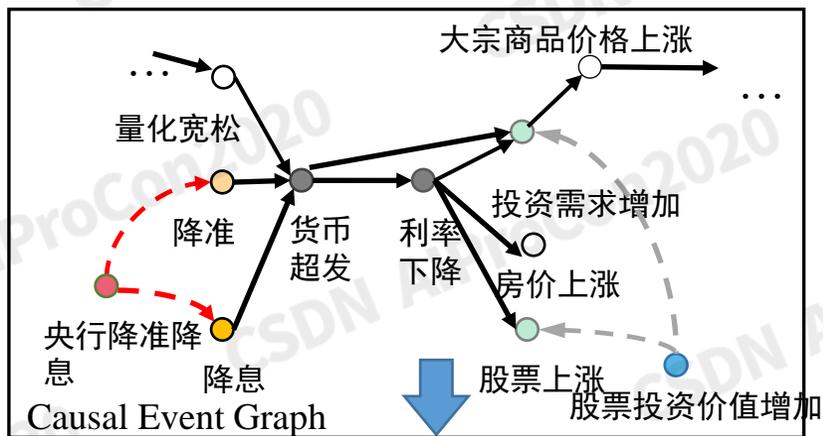
- 基于BM25算法，寻找原因事件/结果事件在图中的“锚”

- 搜寻证据事件

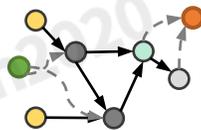
- 以“锚”事件为起点/终点

- 基于BFS

- 形成证据图



Causal Evidence Graph

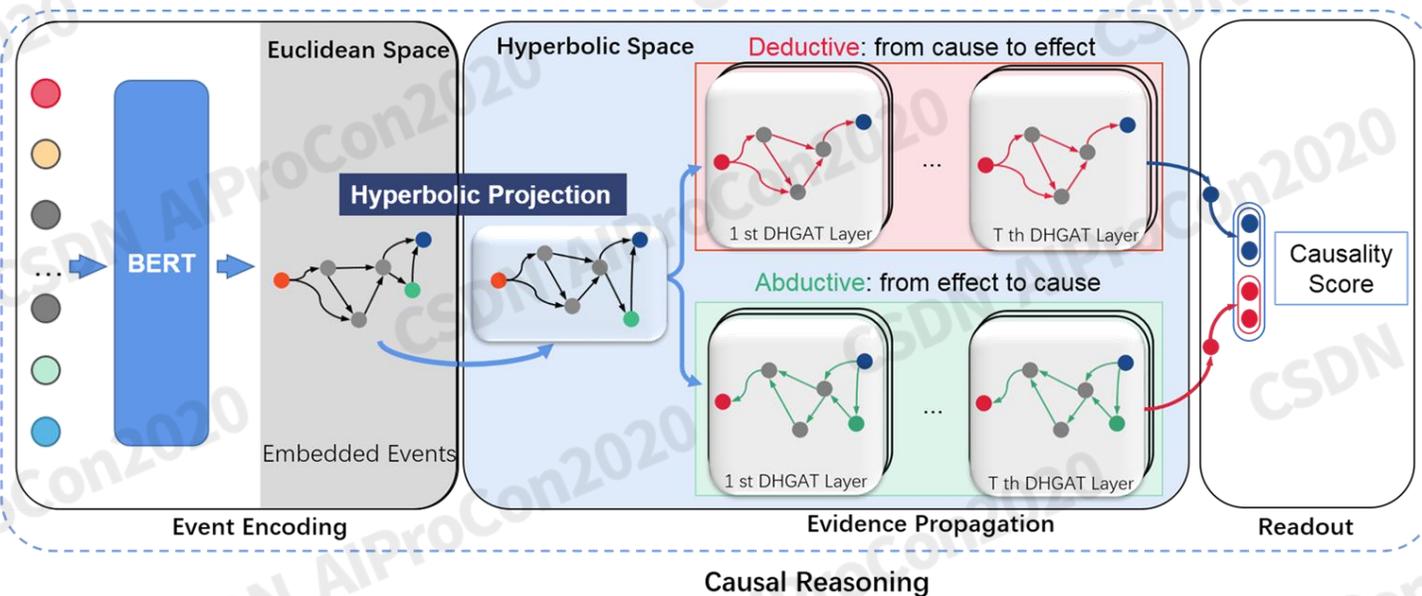
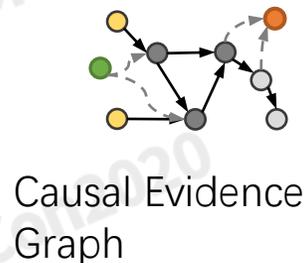


7月3日-4日

CSDN

基于有向双曲图注意力机制的推理模型

- 双曲空间中表示与推理，以减小因果强度误差
- 引入方向性，避免混淆因果



百万 AI

2020
ProCon
万人开发者大会

7月3日-4日

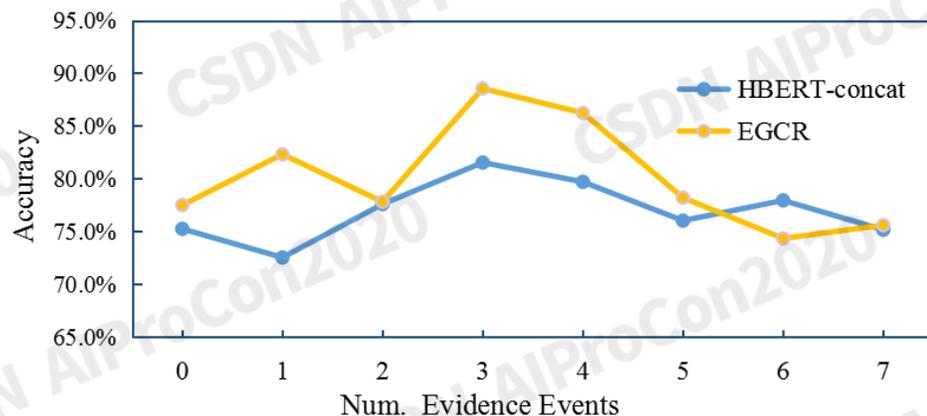
CSDN

实验结果

- 因果推理精度:

Dataset	COPA	C-COPA
δ -Hyperbolicity	6.0	2.5
Methods	Accuracy (%)	
BERT	70.4	75.5
BERT-concat	70.8	76.3
BERT-GAT	71.2	77.1
BERT-DGAT	71.6	78.0
HBERT	71.4	75.6
HBERT-concat	71.2	77.6
BERT-HGAT	72.4	78.3
EGCR	72.8	80.0

- 额外证据事件对因果精度的影响:



- 图神经网络层数对因果推理精度的影响:

Num. Layers	0	1	2	3	4
Accuracy	77.6	79.0	80.0	78.6	76.8



7月3日-4日

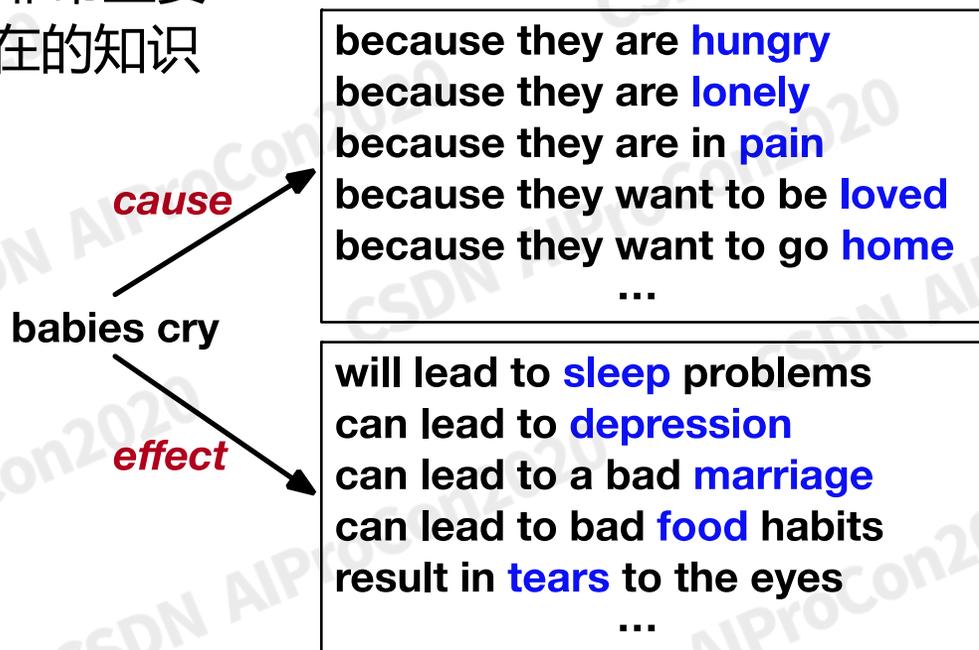
CSDN

大规模开放域因果生成

- 为什么要做因果生成
 - 因果推理对于可解释的机器学习非常重要
 - 生成新的因果知识 vs. 抽取已存在的知识
 - 对话生成, 决策支持系统

主要挑战

- 训练数据从哪里来?
 - 精确模式匹配
- 如何生成更好的原因和结果?
 - 因果关键词指导的生成



7月3日-4日

CSDN

CausalBank: 一个句子级的大规模因果对资源

- **CausalBank**: 一个大规模表达因果模式的英文句子和因果对集合
 - 模式匹配和过滤: 否定因果, 指代词, 句子长度, 连续片段, 显式因果
 - 包含 314 M 因果对, 37 GB 纯文本
 - 从预处理的英文 Common Crawl 语料得到 (5.14 TB), Buck et al., 2014

- 分为两个类别的因果对

- EPC (effect-pattern-cause) category

I am very sad **BECAUSE** *I lost my phone.*

- CPE (cause-pattern-effect) category

The earthquake **RESULTED IN** *many deaths*

Causal Pattern
as, as a consequence/result of, as long as, because, because of, caused by, due/owing to, in response to, on account of, result from
accordingly, consequently, bring on/about, give rise to, induce, in order to, lead to, result in, prevent/stop...from, and for this reason, cause, for the purpose of, if...then, ,..so, so that, thereby, therefore, thus, hence

Table 1: Causal patterns (their morphological variants are ignored) used to get the CausalBank corpus. The first row of patterns belong to the EPC category, while the second row belong to the CPE category.



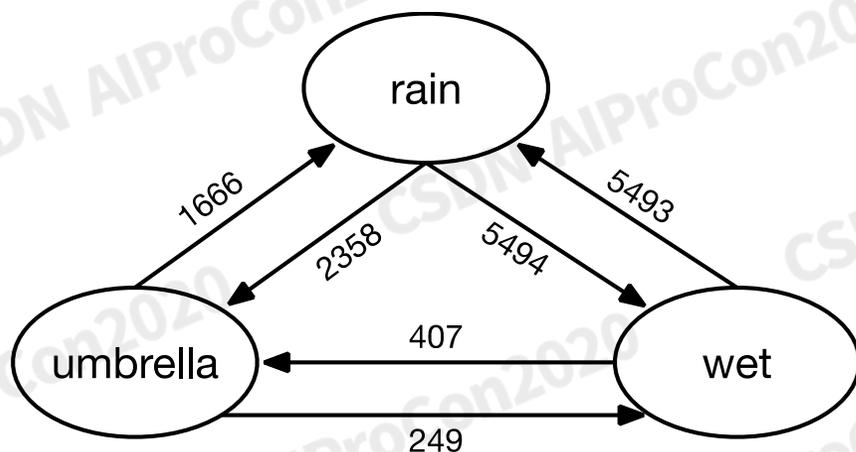
7月3日-4日

CSDN

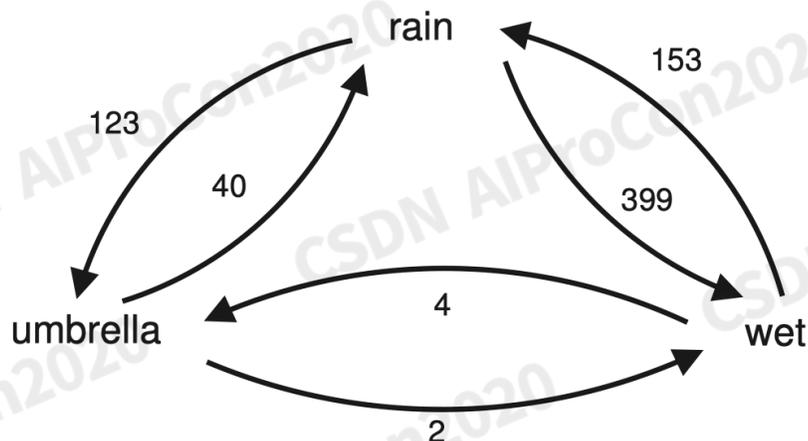
Cause Effect Graph: 一个词汇级别的因果图谱

对前人构建词汇级因果知识图谱的改进

- Commonsense Causal Reasoning between Short Texts, Luo et al., KR 2016
- Word co-occurrence in cause-effect pairs



Our *Cause Effect Graph*:
89.1M directed edges



CausalNet from Luo et al. 2016 :
13.3M directed edges



7月3日-4日

CSDN

我们构建的因果资源与前人工作的对比

Sentential Causal Resource	# CE Pairs
TCR [Ning <i>et al.</i> , 2018]	172
SemEval-2007 Task4 [Girju <i>et al.</i> , 2007]	220
Causal-TimeBank [Mirza <i>et al.</i> , 2014]	318
CaTeRS [Mostafazadeh <i>et al.</i> , 2016]	488
EventCausalityData [Do <i>et al.</i> , 2011]	580
RED [O’Gorman <i>et al.</i> , 2016]	1,147
SemEval2010 Task8 [Hendrickx <i>et al.</i> , 2009]	1,331
BEcauSE 2.0 [Dunietz <i>et al.</i> , 2017b]	1,803
EventStoryLine [Caselli and Vossen, 2017]	5,519
PDTB 2.0 [Prasad <i>et al.</i> , 2008]	8,042
Altlex [Hidey and McKeown, 2016]	9,190
PDTB 3.0 [Webber <i>et al.</i> , 2019]	13 K
DisSent [Nie <i>et al.</i> , 2019]	167 K
CausalBank (Ours)	314 M

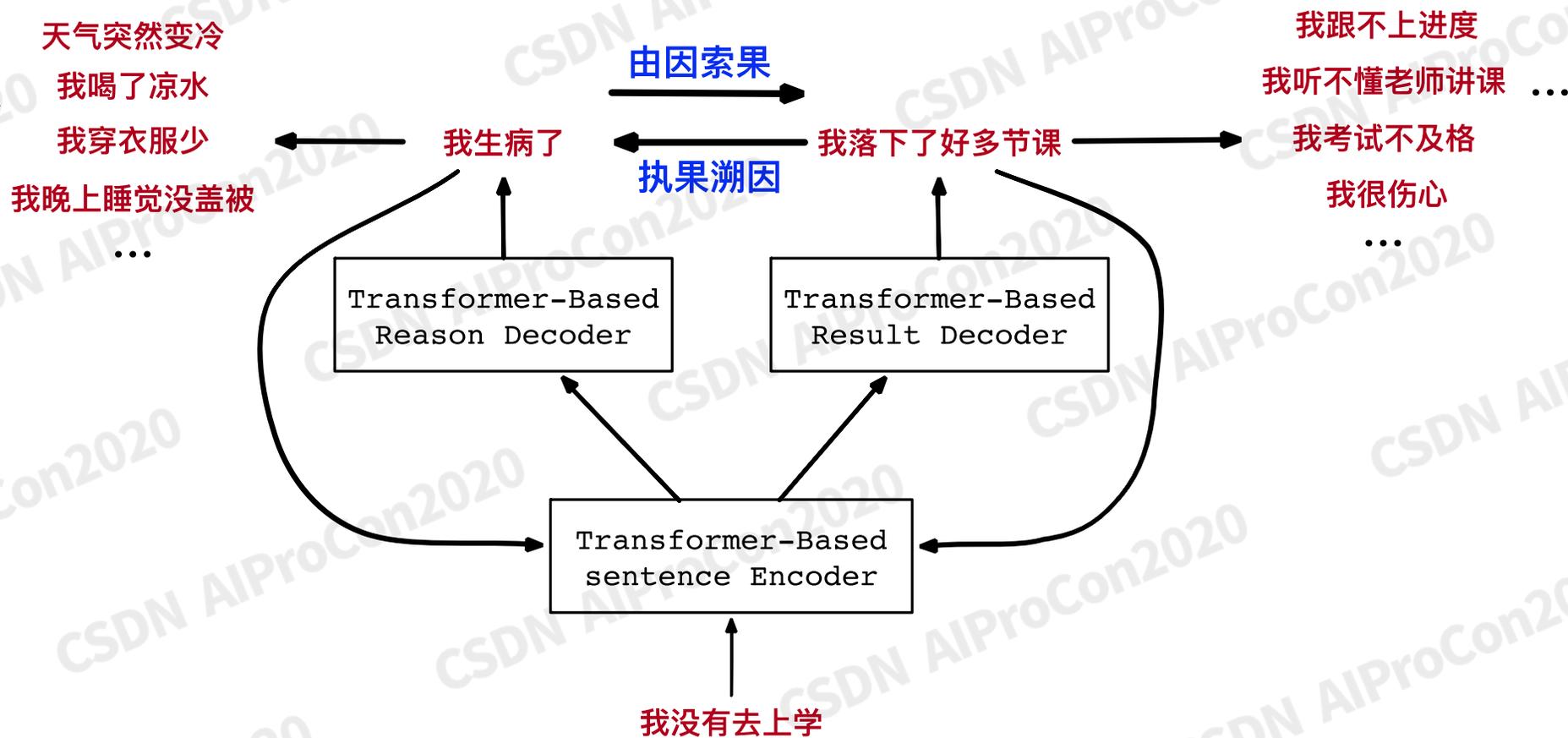
Causal Knowledge Graph	# CE Edges
Event2mind [Rashkin <i>et al.</i> , 2018]	25 K
ConceptNet 5.7 [Speer <i>et al.</i> , 2017]	473 K
ASER Core [Zhang <i>et al.</i> , 2019]	494 K
Atomic [Sap <i>et al.</i> , 2019]	877 K
CausalNet [Luo <i>et al.</i> , 2016]	13.3 M
Cause Effect Graph (Ours)	89.1 M



7月3日-4日

CSDN

方法整体框架



7月3日-4日

CSDN

因果关键词指导的生成

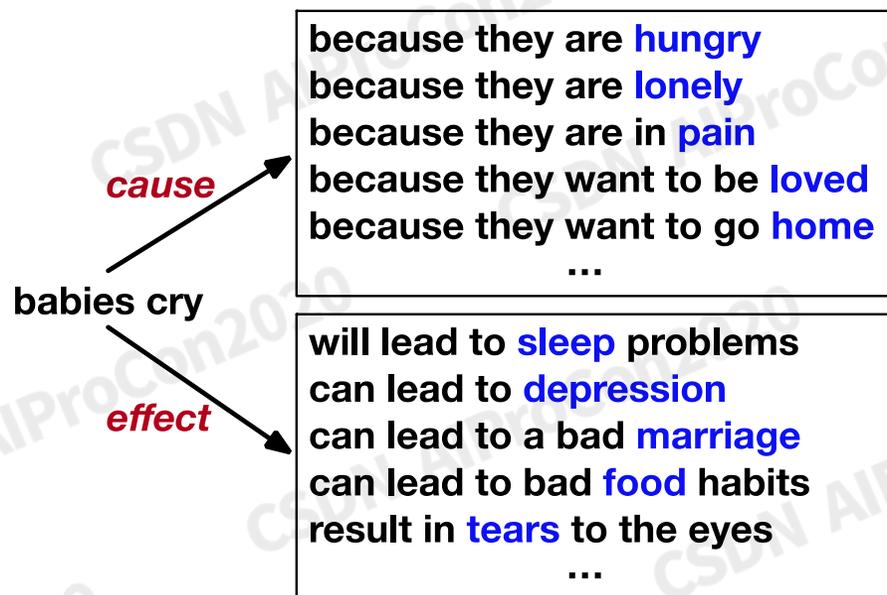
- 在Beam Search阶段中融入关键词约束

- Positive Constraints
- Negative Constraints
- “AND” Constraints
- “OR” Constraints

- Disjunctive Positive Constraints Decoding
 - {smile | smiles | smiling | smiled}
 - {eat | eating | eats | ate}

- Sockeye工具包

- A MxNet based seq2seq toolkit by Amazon
- Incorporate some constraints generation implementation



Fast lexically constrained decoding with dynamic beam allocation for neural machine translation. M. Post and D. Vilar. In NAACL, 2018.

Improved lexically constrained decoding for translation and monolingual rewriting. J E. Hu, Huda Khayrallah, et al. In NAACL, 2019.

Sockeye: A toolkit for neural machine translation. F. Hieber, T. Domhan, et al. arXiv, 2017.



7月3日-4日

CSDN

因果生成评估效果

Method	Dataset	Cause		Effect	
		Per	Acc	Per	Acc
RNN-LSTM	CB_10M	66.0	29.6	55.2	32.2
RNN-GRU	CB_10M	67.6	29.5	48.0	33.7
CNN	CB_10M	37.6	36.1	39.5	35.4
Conv-Transformer	CB_10M	29.5	38.9	31.1	38.2
Transformer	CB_10M	28.3	39.1	29.9	38.4
Transformer	CB_all	31.4	38.0	27.6	39.7
Transformer_BIG	CB_all	29.9	38.5	26.4	39.8

Table 2: Dev-set results: perplexity (**Per**), word accuracy (**Acc (%)**).

Method	Cause				Effect				
	P@1	P@3	H	Div	P@1	P@3	H	Div	
TrainSub	KNN	89.0	67.3	0.85	0.11	98.0	71.3	0.90	0.02
	GPT-2	31.0	22.3	0.39	0.13	8.0	9.3	0.30	0.11
	N-Best	59.0	45.3	0.53	0.15	63.0	42.7	0.53	0.11
	Random	68.0	59.3	0.66	0.11	74.0	61.7	0.70	0.09
	CN-Cons	72.0	71.3	0.79	0.02	66.0	67.0	0.76	0.02
Gold-Cons	78.0	75.3	0.83	0.12	71.0	73.0	0.80	0.10	
COPA_Dev	KNN	10.0	8.0	0.53	0.10	4.0	2.7	0.26	0.01
	GPT-2	40.0	34.0	0.45	0.12	38.0	32.0	0.46	0.10
	Random	66.0	53.7	0.65	0.09	62.0	46.7	0.57	0.08
	N-Best	69.0	65.0	0.77	0.08	72.0	68.0	0.82	0.07
	CN-Cons	74.0	70.0	0.81	0.02	72.0	72.0	0.87	0.02
Gold-Cons	73.0	73.0	0.87	0.09	72.0	71.3	0.87	0.09	

Table 3: Human evaluation results of cause and effect generation.



7月3日-4日

CSDN

在线Demo:

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

输入

车速太快 样例▼

包含

请用；分割各个需包含的词或词组 ...

排除

请用；分割各个需排除的词或词组 ...

生成目标

结果 原因

语言

中文 英文

确认

输出

提示：点击删除按钮可替换低质量生成结果

两车相撞	删除
事故	删除
侧翻	删除
刹车失灵	删除
后车追尾	删除
打滑	删除

输入

车速太快 样例▼

包含

请用；分割各个需包含的词或词组 ...

排除

请用；分割各个需排除的词或词组 ...

生成目标

结果 原因

语言

中文 英文

确认

输出

提示：点击删除按钮可替换低质量生成结果

在高速公路上行驶	删除
驾驶员操作不当	删除
转向系统故障	删除
长时间的驾驶	删除
后轮驱动系统	删除
前轮爆胎	删除
油门踏板	删除
过于紧张	删除
超载	删除
发动机故障	删除

百万
人学AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

在线Demo:

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

输入

The book was deemed inappropriate for children 样例▼

包含

请用;分割各个需包含的词或词组 ...

排除

请用;分割各个需排除的词或词组 ...

生成目标

结果 原因

语言

中文 英文

确认

输出

提示: 点击删除按钮可替换低质量生成结果

resulting in the death of a child	删除
therefore not allowed to be published	删除
therefore not included in the report	删除
therefore not included in this review	删除
therefore they were not allowed to read it	删除

输入

I love my mother 样例▼

包含

请用;分割各个需包含的词或词组 ...

排除

请用;分割各个需排除的词或词组 ...

生成目标

结果 原因

语言

中文 英文

确认

输出

提示: 点击删除按钮可替换低质量生成结果

because she takes care of me	删除
because she is my mother	删除
because she makes me happy	删除
because she loves me unconditionally	删除
because she is always there for me	删除
because she reminds me of myself	删除
because she makes me laugh	删除



7月3日-4日

CSDN

- 2018年9月10日，HIT-SCIR正式发布《金融事理图谱V1.0版本》
- 《金融事理图谱》是在金融领域自动构建的基于因果关系的大规模中文事理逻辑知识库



哈工大SCIR

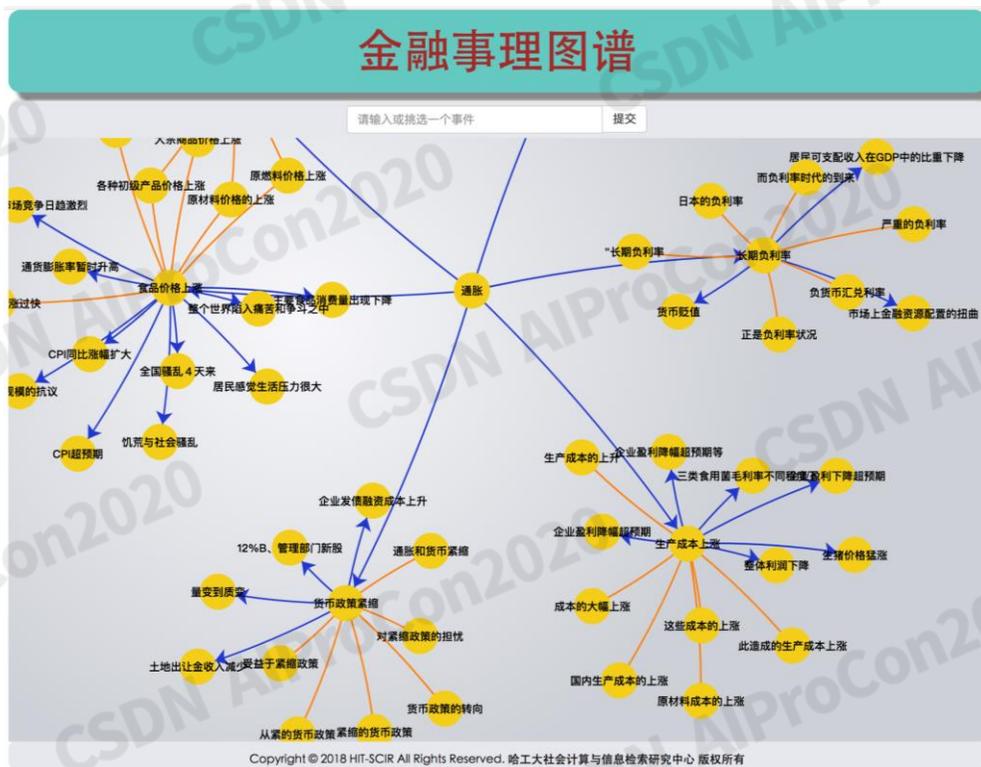
18-9-10 来自专业版微博

哈工大社会计算与信息检索研究中心提出了事理图谱的概念，并基于大规模财经新闻文本构建了一个金融领域事理图谱。本文中我实验室正式对外发布该金融事理图谱Demo ([网页链接](#))。链接：

事理图谱
谓词性事件及其关系
有向图
辑关系，以及概率转移信息
的演化关系多数是不确定的

赛尔新闻 | 哈工大SCIR正式对外发布金融事理图谱Demo V1.0

知识图谱等现有知识库普遍以名词性实体为核...



7月3日-4日

CSDN

- 2019年7月， HIT-SCIR正式发布《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、Resnet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、Resnet、 新浪、中财网、中金在线、证券之星 等网站的金融新闻、 金融领域微信公众号、研报
数据规模（文档数）	11,653,062	12,121,279（ 增加100余万篇文档 ）
具体事件数	1,348,459	2,187,086（ 增加80余万个事件 ）
具体因果系数	1,410,642	1,607,638（ 增加近20万个因果关系 ）
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%（ 提升近16% ）

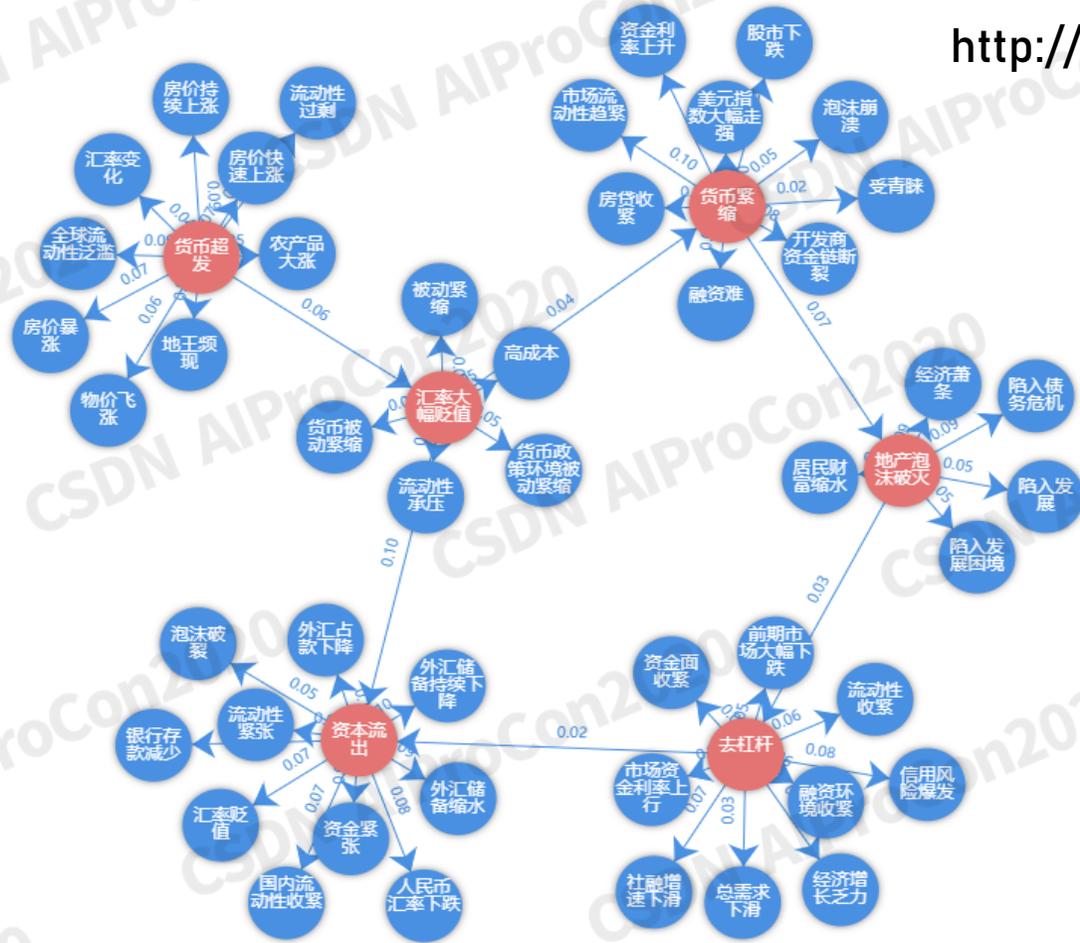


7月3日-4日

CSDN

金融事理图谱演示系统

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



7月3日-4日

CSDN

PEL: Perpetual Eventics Learner

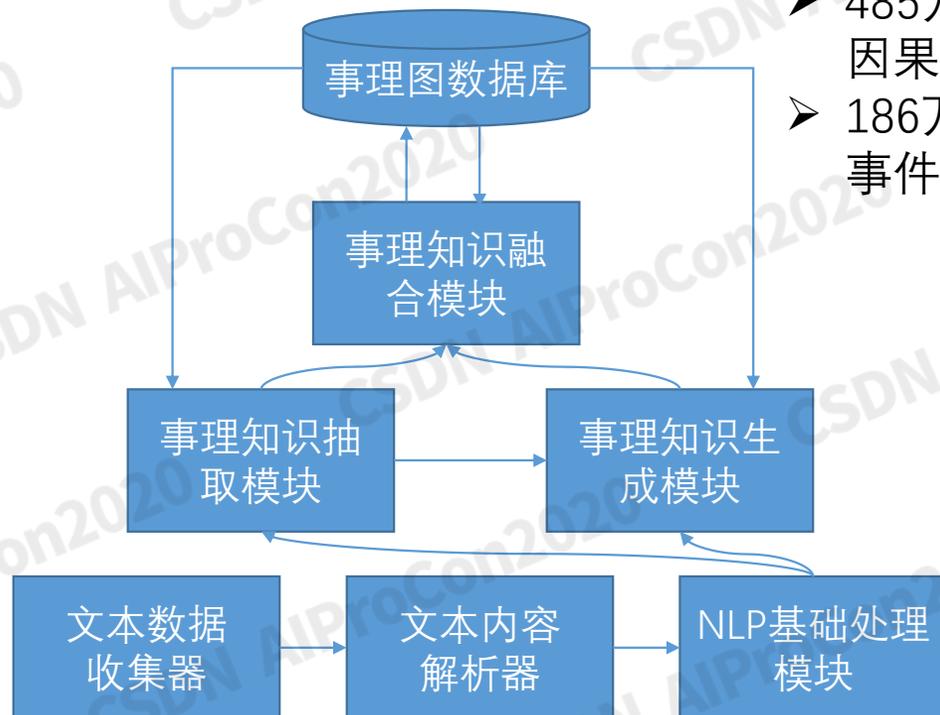
<http://openeg.8wss.com/main>

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



现有数据量:

- 485万具体事件, 417万具体因果边;
- 186万抽象事件 (含上下位事件), 131万抽象因果;

百万 AI

2020 ProCon 万人开发者大会

7月3日-4日

CSDN

事理图谱的潜在应用

• 消费意图识别与推荐

- 事件可以和“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



7月3日-4日

CSDN

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



7月3日-4日

CSDN

事理图谱研究基础

- Zhongyang Li, Xiao Ding, Ting Liu, Edward Hu, Benjamin Van Durme. Guided Generation of Cause and Effect. In Proc. **IJCAI** 2020
- 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.
- 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.
- Xiao Ding, Yue Zhang, Ting Liu, Junwen Duan. Knowledge-Driven Event Embedding for Stock Prediction. In Proc. COLING 2016.
- Xiao Ding, Yue Zhang, Ting Liu, Junwen Duan. Deep Learning for Event-Driven Stock Prediction. In Proc. IJCAI 2015.
- 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.



7月3日-4日

CSDN

结论

- 文本推理能力是人工智能的一个重要体现
- 事理图谱可以给文本推理提供重要的知识支撑
- 因果知识的融入是当前人工智能的一个瓶颈
- 图神经网络与预训练模型相结合可以推动文本推理技术的发展



7月3日-4日

CSDN

谢谢!



百万
学AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN