

SEU 知识推理

2019年12月3日 11:25

- 知识推理背景场景
- 知识推理方法简介
 - 基于逻辑的方法
 - 基于机器学习的方法
 - 基于神经网络的方法

- 知识推理的问题背景

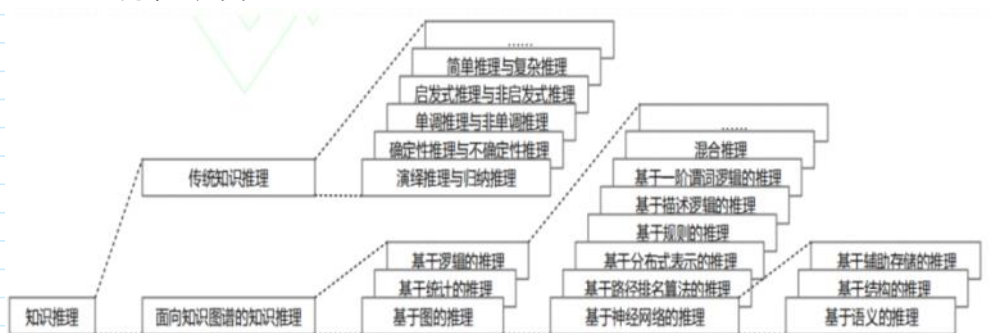
- 目前大部分的开放知识图谱，例如Freebase、DBpedia都是有人工或者半自动的方式进行搭建。这些知识图谱比较稀疏，大量实体之间隐含的关系没有被充分的挖掘出来
- 知识图谱补全（knowledge graph completion KGC）技术就是为了补全知识图谱中不完整的知识、关系。其中最主要的方法就是知识推理--利用知识图谱中现有的显性知识来预测图谱中尚未存储的隐性知识，逐步将知识图谱补充完整

- 知识推理的应用场景

- 金融反诈骗
- 智能问答与搜索
- 军事辅助决策
- 智能医疗
-

- 知识推理的简介

- 简单的分类



- 基于逻辑--一阶为此逻辑、描述逻辑等利用规则推理
- 基于统计--机器学习、统计学习
- 基于图--path ranking、神经网络
- 基于逻辑的知识推理
 - E.G.

问题：谁养鱼？

条件：

- 1、在一条街上，有5座房子，喷了5种颜色。
- 2、每个房里住着不同国籍的人
- 3、每个人喝不同的饮料，抽不同品牌的香烟，养不同的宠物

要求：

- 1、英国人住红色房子
- 2、瑞典人养狗
- 3、丹麦人喝茶
- 4、绿色房子在白色房子左面
- 5、绿色房子主人喝咖啡
- 6、抽Pall Mall香烟的人养鸟
- 7、黄色房子主人抽Dunhill香烟
- 8、住在中间房子的人喝牛奶
- 9、挪威人住第一间房
- 10、抽Blends香烟的人住在养猫的人隔壁
- 11、养马的人住抽Dunhill香烟的人隔壁
- 12、抽Blue Master的人喝啤酒
- 13、德国人抽Prince香烟
- 14、挪威人住蓝色房子隔壁
- 15、抽Blends香烟的人有一个喝水的邻居

■ 基于逻辑规则：

- 自然性
- 适宜于精确性的知识的表示，而不适宜于不确定性知识的表示
- 易实现
- 与谓词逻辑表示法相对应



一阶逻辑的字母表

- 个体常项: $a, b, c, \dots, a_1, b_1, c_1, \dots$
- 个体变项: $x, y, z, \dots, x_1, y_1, z_1, \dots$
- 函数符号: $f, g, h, \dots, f_1, g_1, h_1, \dots$
- 谓词符号: $F, G, H, \dots, F_1, G_1, H_1, \dots$
- 量词符号: \exists, \forall
- 联结词符号: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$
- 括号与逗号: $(,), ,$

■ 基于统计学习的知识推理

- 马尔科夫逻辑网
- 概率软逻辑
- 贝叶斯推断
-

■ 基于图的知识推理

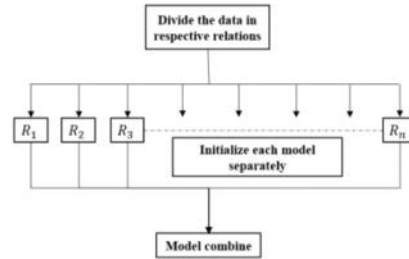
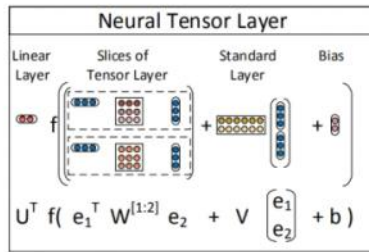
- 路径排序算法 (Path Ranking Algorithm--PRA)
- 不完备知识库的关联规则挖掘 (Association Rule Mining under Incomplete Evidenc--AMIE)
-

■ 基于神经网络的知识推理

- 基于语义的推理
- 基于结构的推理
-
- 神经网络中的推理
 - ◆ 基于语义的推理建立在挖掘和利用语义信息的基础上，例如实体和关系的名称、描述以及上下文信息等
 - ◆ 基于结构的推理是指利用知识库中的三元组内部或相互之间的结构联系进行推理，常用于多步推理问题当中。
 - ◆ 神经网络中基于语义的推理
 - ◇ 用双线性张量层代替传统的标准线性神经网络层 (NTN)

$$g(e_1, R, e_2) = u_R^T f \left(e_1^T W_R^{[1:k]} e_2 + V_R \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + b_R \right)$$

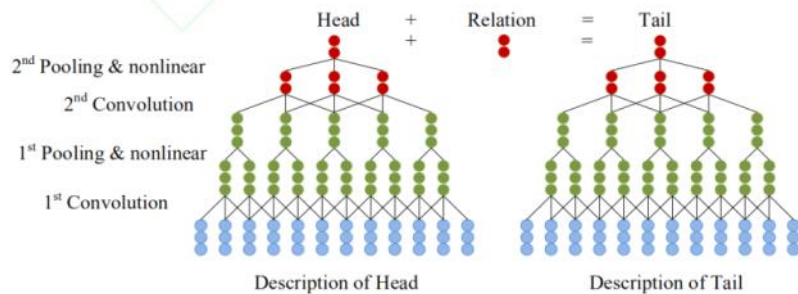
$$J(\Omega) = \sum_{i=1}^N \sum_{c=1}^C \max \left(0, 1 - g \left(T^{(i)} \right) + g \left(T_c^{(i)} \right) \right) + \lambda \|\Omega\|_2^2,$$



◇ 知识表示模型

- TransE, NIPS2013, Translating embeddings for modeling multi-relational data
- TransH, AAAI2014, Knowledge graph embedding by translating on hyperplanes
- TransR, AAAI2015, Learning Entity and Relation Embeddings for Knowledge Graph Completion
- TransD, ACL2015, Knowledge graph embedding via dynamic mapping matrix
- TransA, arXiv2015, An adaptive approach for knowledge graph embedding
- TranSparse, AAAI2016, Knowledge graph completion with adaptive sparse transfer matrix
- TransG, arXiv2015, A Generative Mixture Model for Knowledge Graph Embedding
- KG2E, CIKM2015, Learning to represent knowledge graphs with gaussian embedding

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'} [\gamma + d(h+r,t) - d(h'+r',t')]_+$$



- **ProjE**: 将实体预测视为多候选项排名问题，取其中排名得分最高者的候选想为实体预测结果
- **ConMask**: 为了能够应对知识图谱外部的位置新实体，进一步定义了“Open-world KGC”，即“开放世界知识图谱”
 - 依赖于关系的内容屏蔽：筛选文本信息，删去无关信息，仅留下与任务有关的内容，其中利用了注意力机制
 - 目标融合：使用全卷积神经网络从相关文本抽取目标实体的嵌入向量
 - 目标实体解析：生成候选实体和抽取实体嵌入之间的相似度排名，相似度最高的认为是最佳结果

■ 神经网络中基于结构的推理

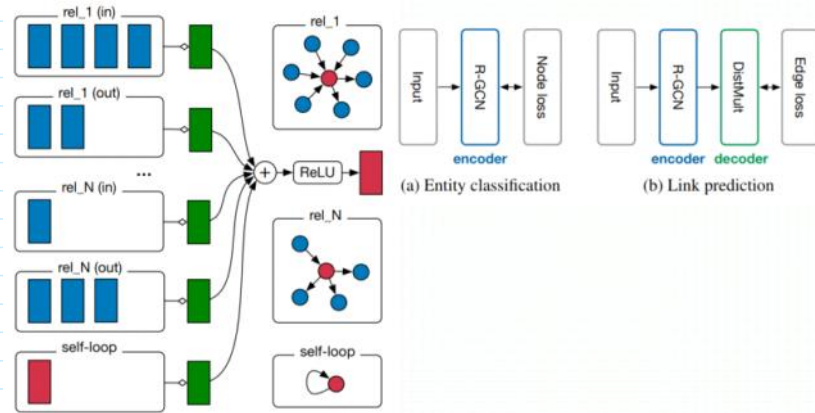
□ 基于相邻实体的推理

- ◆ 相邻实体：关系图卷积神经网络 (R-GCN)

◇ R-GCN从图的角度讲GCN网络引入知识推理，思路新颖

◇ 但是该方法基于GCN网络的演化设计较为粗糙，针对关系类型较少的数据集处理能力偏弱

- ◇ 半监督的方法还是比较难超越projE
- ◇ 实验结果不够稳定：在AIFB和AM标准数据集上取得的效果不好，而在MUTAG和BGS数据集上的实验结果不佳

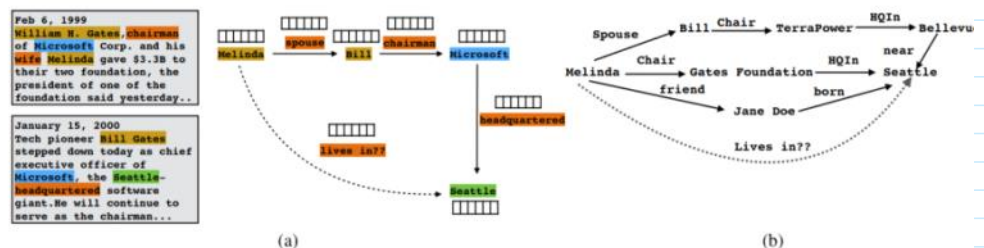


□ 基于多跳关系的推理

- ◆ 三元组之间并非相互独立
- ◆ E.G.一个三元组的尾实体可能是另一个三元组的头实体。这样的两个三元组可以构成一个五元组路径，而此时该五元组的头尾实体之间可能隐含了某种可以表示的关系
- ◆ E.G.PTransE：依次经过BornInCity、CityInState和StateInCountry的路径，通过关系的组合操作建模，组合操作可以是路径上的关系的加和、连乘等形式，然后将组合后的路径看成是头实体SteveJobs和尾实体UnitedStates之间的转移

□ 基于组合路径的推理

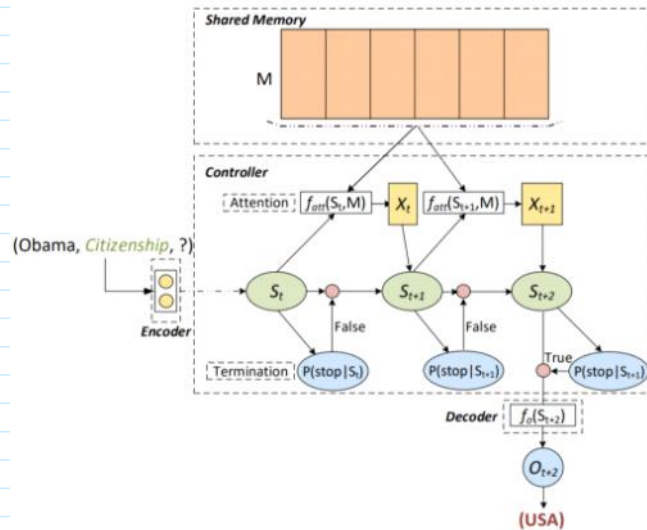
- ◆ 针对多个三元组链接而成的多元组，其首尾实体之间可能存在多条路径。即，首尾实体之间可能由若干多元组构成多种不同的路径形式



- ◆ (组合路径&RNN) 缓解了路径排序算法 (path ranking algorithm) 无法共享参数及参数数量大的共性问题，允许跨多个关系类型共享参数进行推理
- ◆ 除建模的每步关系，还建模对应的实体，共同推理所有的关系、实体和实体类型

□ 基于辅助存储的推理

- ◆ 基于辅助存储的推理类比人类大脑对知识的存储和读写。利用共享记忆组件或外部存储矩阵来存储推理所需中间结果或必要信息，模拟了人的推理思考过程。通过辅助存储能够便于推理过程获取隐含信息，提高推理效率。
- ◆ 隐性推理网IRN (Implicit Reasoning Nets)
 - ◇ 设计了共享记忆组件来隐式地存储知识库信息，以此模仿人脑对知识的存储
 - ◇ 在没有人干预的情况下能够通过对共享记忆组件的读取来隐式学习多步推理过程，模拟了人脑在进行推理判断时读取记忆的过程



◆ DNC (Differentiable Neural Computer) 模型

- ◇ 与IRN模型区别在于其**不包含**一个LSTM神经网络控制器和一个存储矩阵，而且该存储矩阵可以**选择性的进行读写操作**，可以通过迭代修改存储内容（IRN模型的记忆组件只支持读取数据，而不能随时写入或更改）
- ◇ DNC通过对共享外部矩阵的读写来模拟人脑对知识记忆的“增删改查”过程，此方法在于外部存储矩阵所存储信息的可解释性不强
 - ▶ 修改知识记忆的依据不详
 - ▶ 未提出区分修改前后知识正误的判断方法
 - ▶ 每次与矩阵交互读取及修改的隐式向量包含哪些内容不明
- ◇ 严格意义上DNC模型的实验针对的是**更加直接意义上的推理**，而且是直接和LSTM性能和效果对比

□ 多种混合方法的推理

- ◆ 混合PRA与分布式表示的推理：PRAlatent
- ◆ 混合规则与分布式表示的推理：通过矩阵分解直接学习一阶逻辑规则的代表
- ◆ 混合规则与神经网络的推理：Neural NLP (neural logic programming)
- ◆ Neural LP的**混合方法**为目前一个很高的水平

• 知识推理总结与展望

- 面向多元关系的知识推理
- 融合多源信息与多种方法的知识推理
- 基于小样本学习的知识推理
- 动态知识推理
-