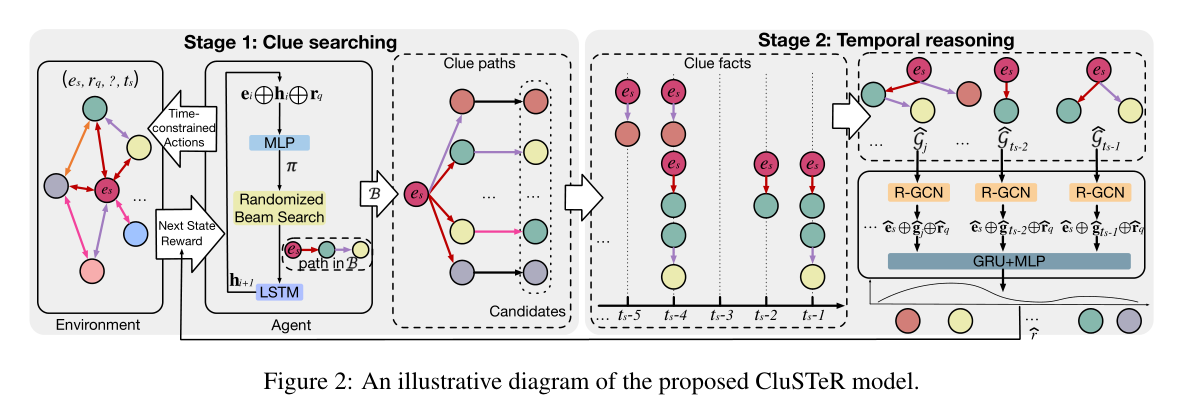
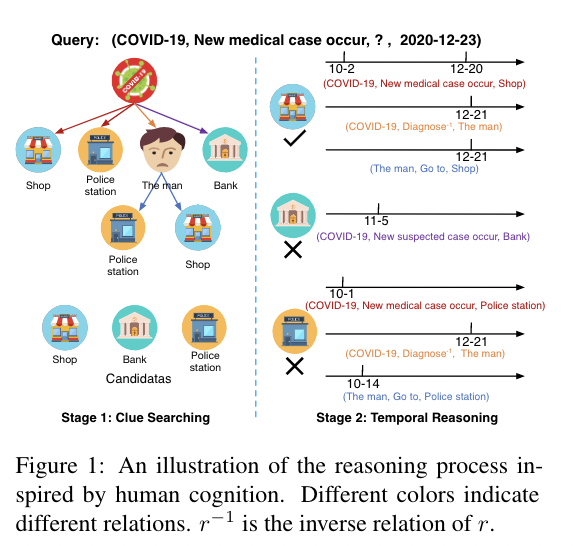
Search from History and Reason for Future: Two-stage Reasoning on

Temporal Knowledge Graphs

**方法总叙：**

本文提出了一个新的模型，称为CluSTeR，包括两个阶段，线索搜索(阶段1)和时间推理(阶段2)。在第一阶段，CluSTeR将线索搜索形式化为马尔可夫决策过程(MDP)(萨顿和巴尔托，2018)，并学习波束搜索策略来解决它。在第二阶段，CluSTeR将第一阶段发现的线索重新组织成一系列图形，然后使用图形卷积网络(GCN)和门控递归单元(GRU)从图形中推导出准确的答案。



如上图所示，该模型由两个阶段组成，线索搜索和时间推理。这两个阶段被协调来执行快速和慢速思维(Daniel，2017)，分别解决TKG推理任务，受到人类认知的启发。具体来说，阶段1主要集中在搜索线索路径上，这些线索路径的组合语义信息与具有时间约束的给定查询相关。然后，为第二阶段的推理提供线索路径和随后的候选实体，第二阶段的推理主要集中在细致地建模线索事实之间的时间信息并获得最终结果。在CluSTeR模型中，这两个阶段在训练阶段相互作用，在推理阶段共同决定最终答案。

**第一阶段：**阶段1的目的是搜索和归纳与给定查询相关的线索路径(es，rq，？，ts)从历史中。前期研究(金等，2019，2020；朱等，2020)使用启发式策略来提取1跳重复路径，丢失了大量其他信息线索路径。此外，历史上有大量的事实。因此，一个可学习的、高效的线索搜索策略是非常必要的。受这些观察的激励，阶段1可以被视为一个顺序决策问题，并由RL系统解决

**第二阶段：**为了更深入地探究不同时间戳的线索事实之间的时间信息以及并发线索事实之间的结构信息，阶段2将所有线索事实重新组织成一系列图形G = { G0，...，Gj，...，Gts-1 }，其中每个Gj是由时间戳j=0,...ts-1的线索事实组成的多关系图。我们使用一个ω层RGCN(Schlichtkrull等人，2018)来建模.

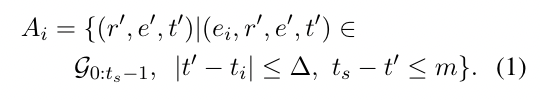
**Stage 1 Clue Searching**

RL系统由两部分组成，代理和环境。我们将RL系统表述为一个**MDP**，它是一个从代理和环境之间的交互中学习的框架，以找到最佳的线索路径。从es开始，代理通过随机波束搜索策略依次选择输出边，并遍历到新实体，直到到达最大步长I。

RL系统描述：

**状态**：每个状态i=（ei、ti、es、rq、ts）∈ 是一个元组，

**动作**（时间限制动作）：与静态KG相比，TKG的时间维度导致了极大的动作空间。此外，人类的记忆集中在最近发生的事件上。因此，我们将每个事实的时间戳与时间戳之间的时间间隔限制为不超过m。前一个操作的时间戳和每个可用操作之间的时间间隔不超过∆. 因此，可能的交流电压的设置为：Ai∈A（A是所有可用动作的集合）在步骤i中，包括时间限制支出ei，

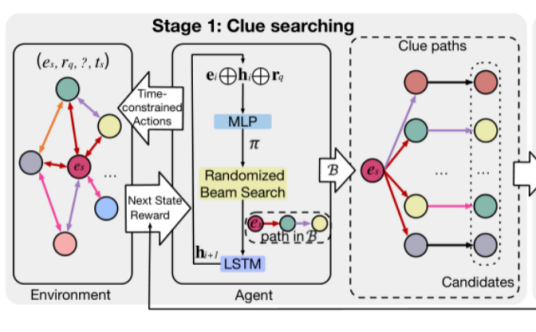
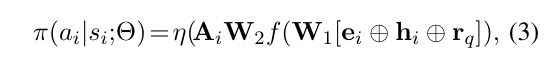


**状态转移**：一个跃迁函数δ：S×A→S在TKG的情况下是确定性的，只需将状态更新为与代理选择的操作相关的新实体。

**奖励**：代理只在搜索结束时收到一个最终报酬R，这个报酬R是两个部分的组合，二进制奖励和真实的价值奖励。二进制奖励是如果目的实体Ei是正确的目标实体eo，否则是0，此外，如果ei是目标实体，则代理会从阶段2获得一个真实的价值奖励r，

语义策略网络（注意各个变量的含义和维度）：

它根据当前状态和搜索历史hi= (es，a0，...，ai-1)而不考虑阶段1中的时间戳，计算所有动作的策略分布。其中ai=（ri+1，ei+1，ti+1）这是每一步i采取的动作，实际上，没有时间戳的搜索历史是一个候选线索路径。其中动作ai的embedding和ri+1和ei+1的连接操作 ，然后用LSTM比编码一个候选路径作为一个连续的hi向量。 截屏2021-08-17 下午2.19.16

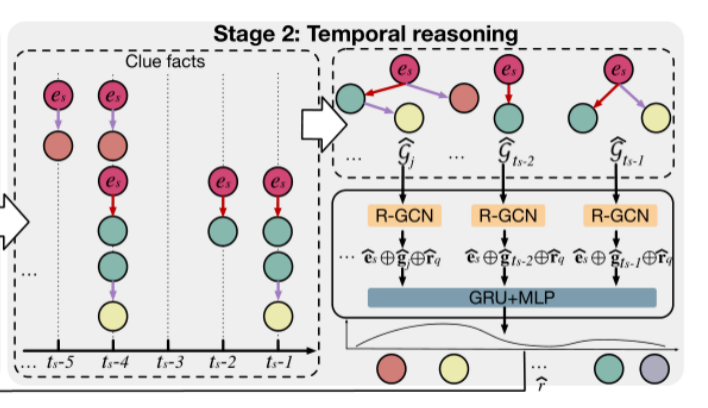
对于步骤i，动作空间是通过叠加所有动作的嵌入来编码的，这些嵌入被表示为Ai∈R | Ai |×2d。这里，讨论了实体嵌入和关系嵌入的维数。然后，策略网络通过W1和W2参数化的多层感知器（MLP）计算所有动作的分布π，如下所示：

随机束波搜索：

具体就是，在步骤i时，一个束波包含B个候选路径，对每个候选路径，我们

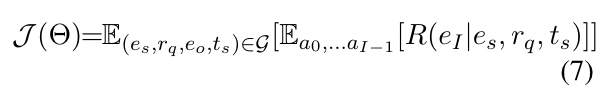
我们将最有可能的操作B（根据等式3）附加到路径的末尾，从而产生一个新的路径池，大小为sizeB×B。然后，我们要么选择概率为μ的得分最高的路径，要么均匀地选择概率为1−μ的随机路径重复了b次。每个候选线索路径在步骤i的得分等于**截屏2021-08-17 下午3.00.10**

**Stage 2: Temporal Reasoning**

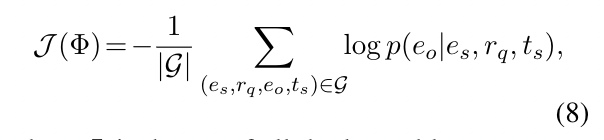
****

**训练策略：**

**Stage1:随机束波策略网络是通过最大化奖励期望进行训练的**

****

**Stage2：使用交叉上定义损失函数进行训练**

****