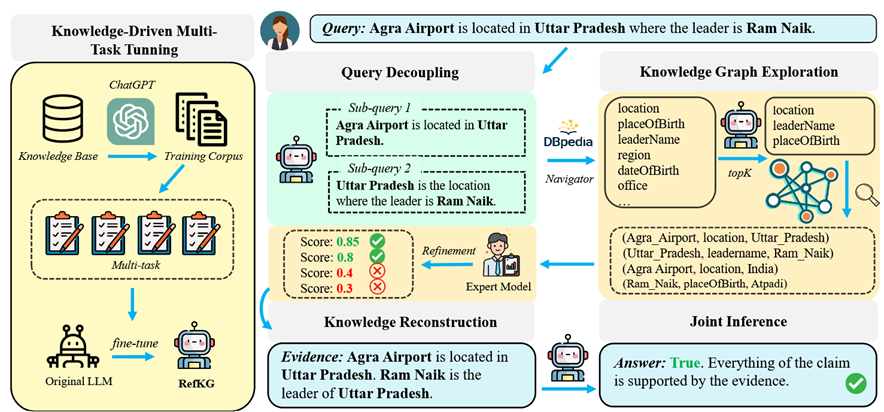
**Reflection on Knowledge Graph for Large Language Models Reasoning**

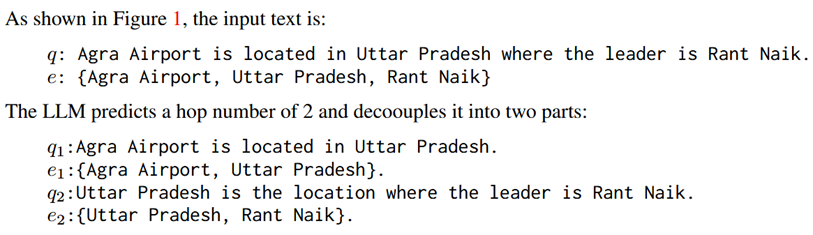


* 动机：用传统方法从知识图谱查询到的知识容易包含噪声，需要一种方法能更好的评估并精炼知识
* 方法：

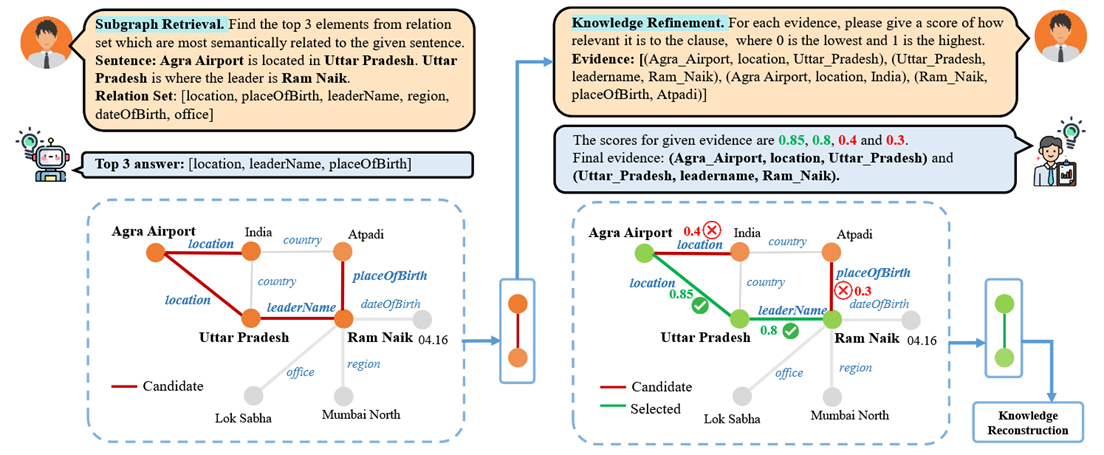
本文提出RefKG，该模型包含三个处理步骤：

1. 查询分解：

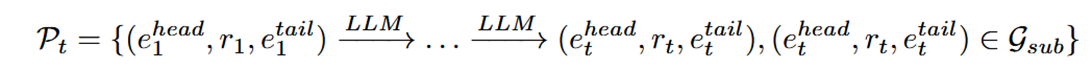
将多跳查询分解为多个单跳查询，并且提取每个单跳查询的topic entity。



1. 知识图谱检索和精炼：



* 子图检索：用大模型检索出Top-k个关系，由topic entity检索包含Top-k关系和topic entity的子图

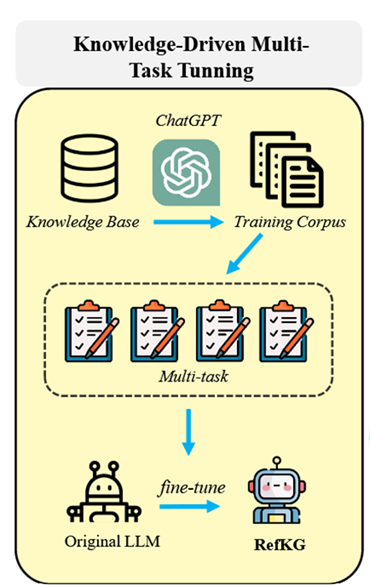


* 知识精炼：训练MLP对关系进行评分，保留高于预设阈值的关系



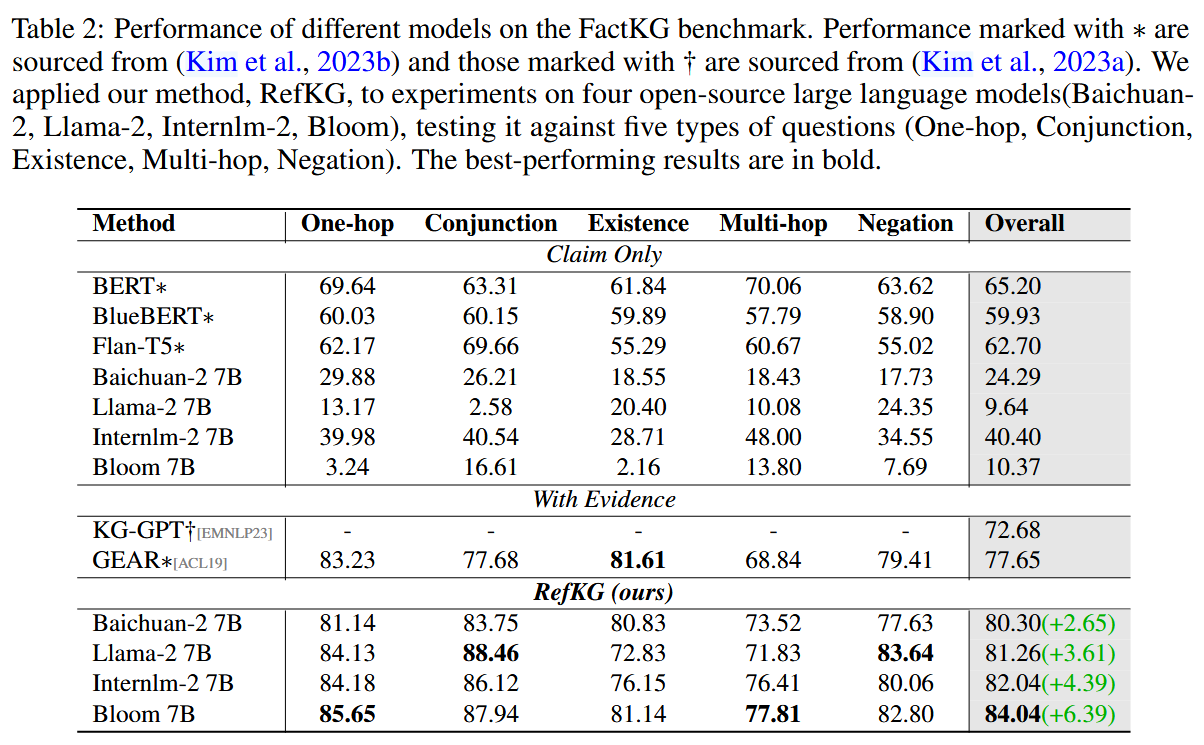
1. 用重构的知识推理：将三元组重构成自然语言，作为证据输入大模型

多任务微调：通过GPT对数据进行标注，再针对Decoupling, Retrieval和Evidence Generation三项任务对大模型进行微调。

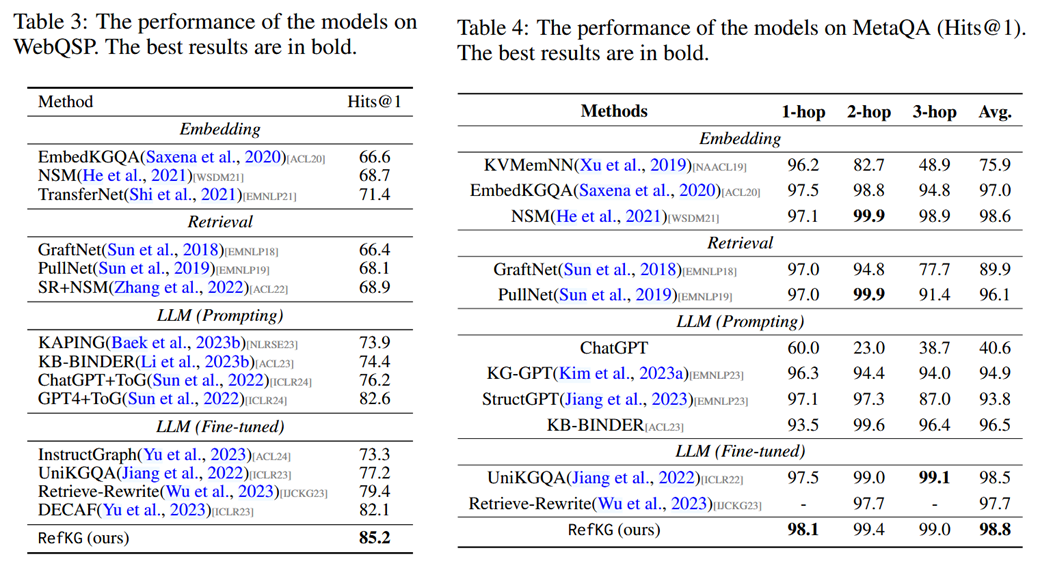


* 实验：

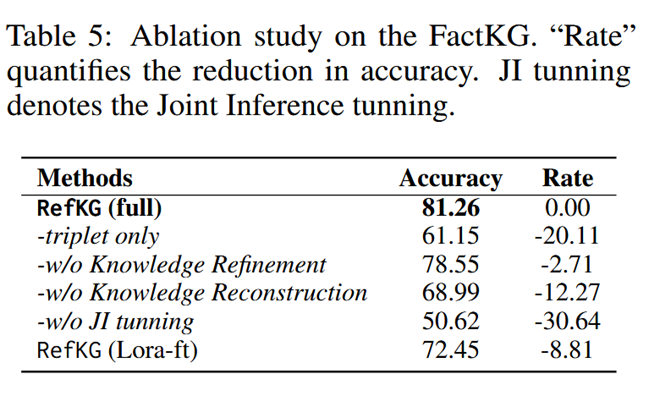
1）对不同大模型进行微调，性能均大幅超过微调前，同时高于对比模型中最佳的GEAR



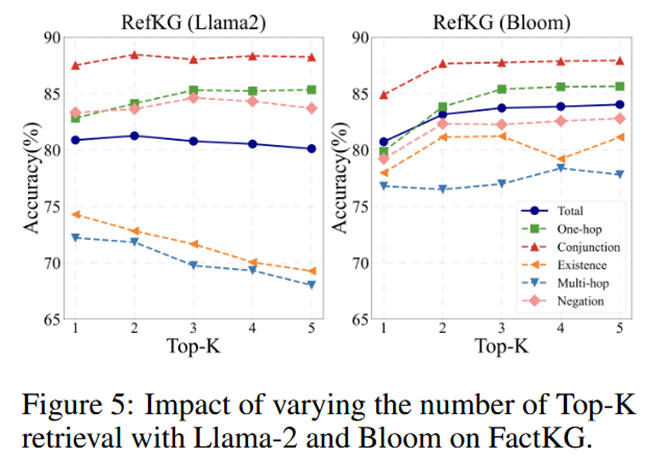
2）在WebQSP和MetaQA数据集上的实验表明RefKG效果良好



3) 通过消融实验发现，将三元组转化为文本形式对推理有很大帮助

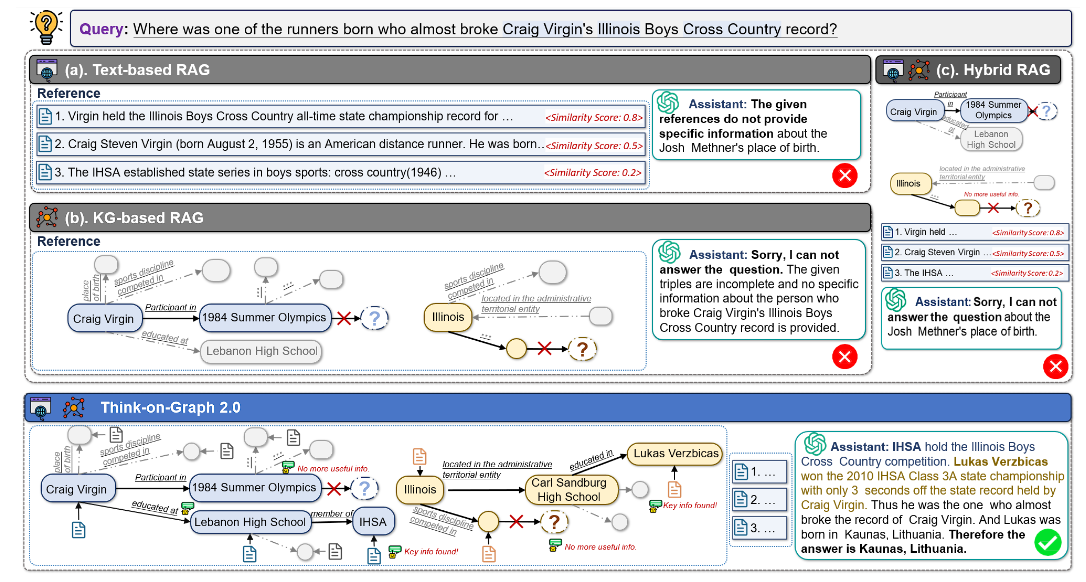


4) 增加Top-k数量对Llama的作用是负面的，但对Bloom是正面的

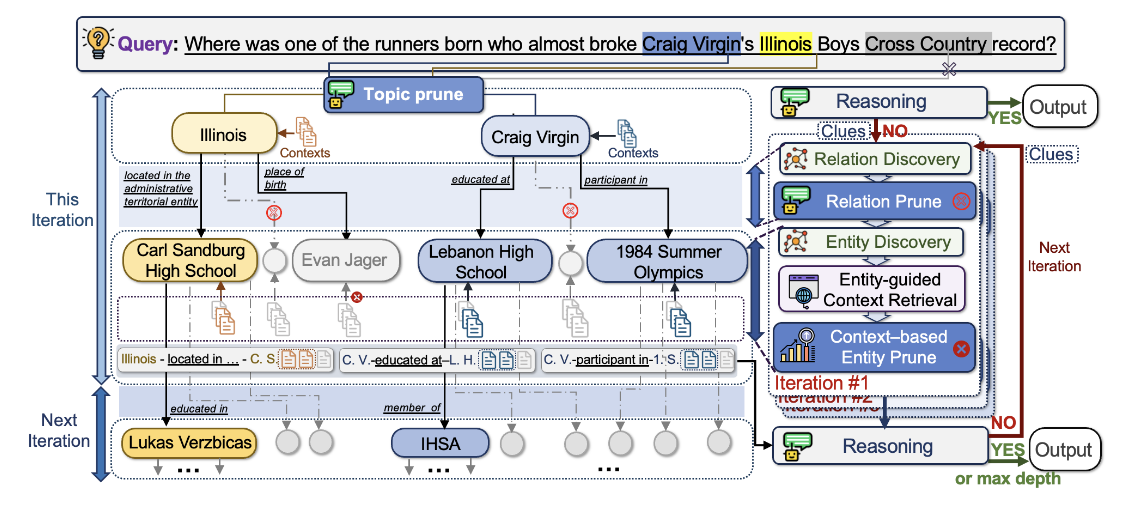


**Think-on-graph 2.0: Deep and faithful large language model reasoning with knowledge-guided retrieval augmented generation**

* 动机：基于文本的RAG无法从文本发现知识的关联，而基于知识图谱的RAG存在知识图谱不完整的问题，现有的混合RAG又没有有效结合使用两部分知识进行检索。因此，基于Think-on-Graph的研究，Think-on-Graph 2.0研究能同时利用知识图谱和文档解答问题的模型，该模型可以用文档知识增强在知识图谱上的检索。



* 方法：



1. Topic prune：由大模型根据问题和其中的实体筛选topic entity

2. 知识指引的图检索：

* Relation Discovery：提取topic entity的邻接关系
* Relation Prune：用大模型对提取的关系进行打分并筛选
* Entity Discovery：提取与topic entity相连的、以上一步筛选的关系相连的实体。



3. 知识指引的上下文检索

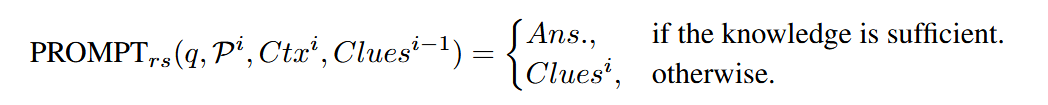
* Entity-guided Context Retrieval：使用dense retrieval models(DRMs)判断尾部实体文档资料与问题的相关分数，其中会被转化为自然语言与尾部实体的文档资料相连



* Context-based Entity Prune：根据相关分数计算该尾部实体与问题的相关度，选择分数较高的尾部实体作为下一轮迭代的topic entity加入推理路径

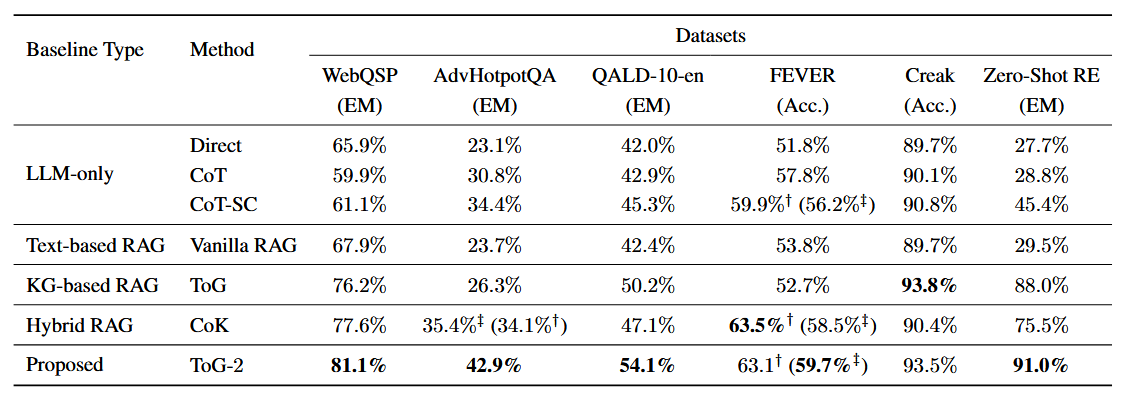
4. 推理：

将问题、当前轮次得到的推理路径、文档资料和上一轮生成的Clue(对当前资料是否足够的分析)输入大模型：

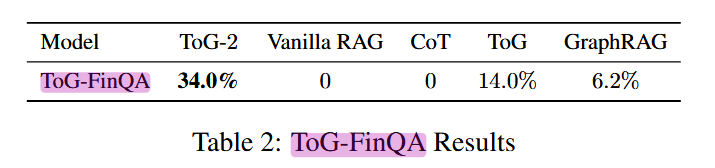


* 实验：

1）实验结果显示ToG-2的性能高于其他模型：

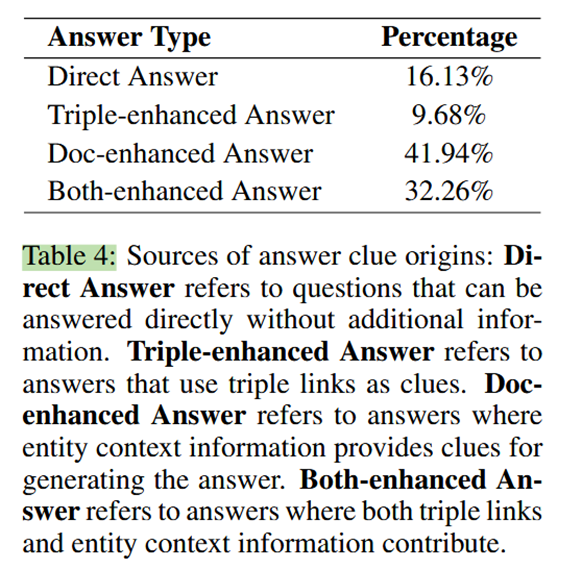


2）构建ToG-FinQA检验模型解决涉及训练时未接触数据问题的能力，测试显示ToG-2表现高于其他模型



3) 实验发现，大模型依赖文档解决的问题最多（根据大模型返回的clue进行判断）





**FiDeLiS: Faithful Reasoning in Large Language Model for Knowledge Graph Question Answering**

* 方法：本文构建的模型FiDeLis包含两个重要结构：Path-RAG和Deductive Verification guided Beam Search(DVBS)，迭代使用两个结构处理问题

1. Path-RAG：

1）通过大模型从问题中提取关键词

2）从知识图谱中检索与关键词最语义相关的实体和关系

* 计算初始嵌入：

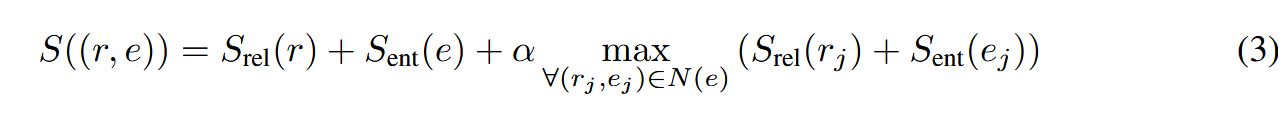


* 将关键词连接得到K，计算实体、关系与K的相似度，取top-m的实体和关系：



3）计算关系-实体对的分数，选择得分高的关系-实体对进入DVBS阶段



DVBS:

1）大模型提出推理计划

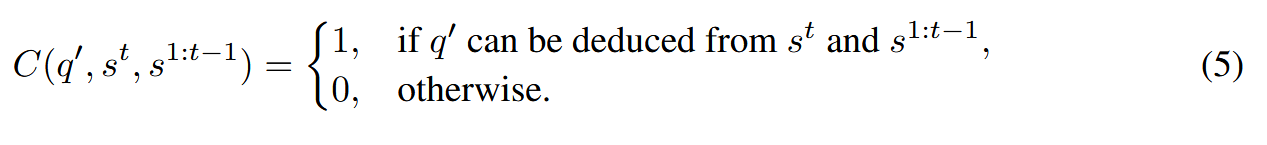
2）基于Beam Search拓展推理路径

* 用大模型判断推理步s（即关系-实体对），其中w为大模型生成的推理计划。由大模型直接决定属于top-k的推理路径：



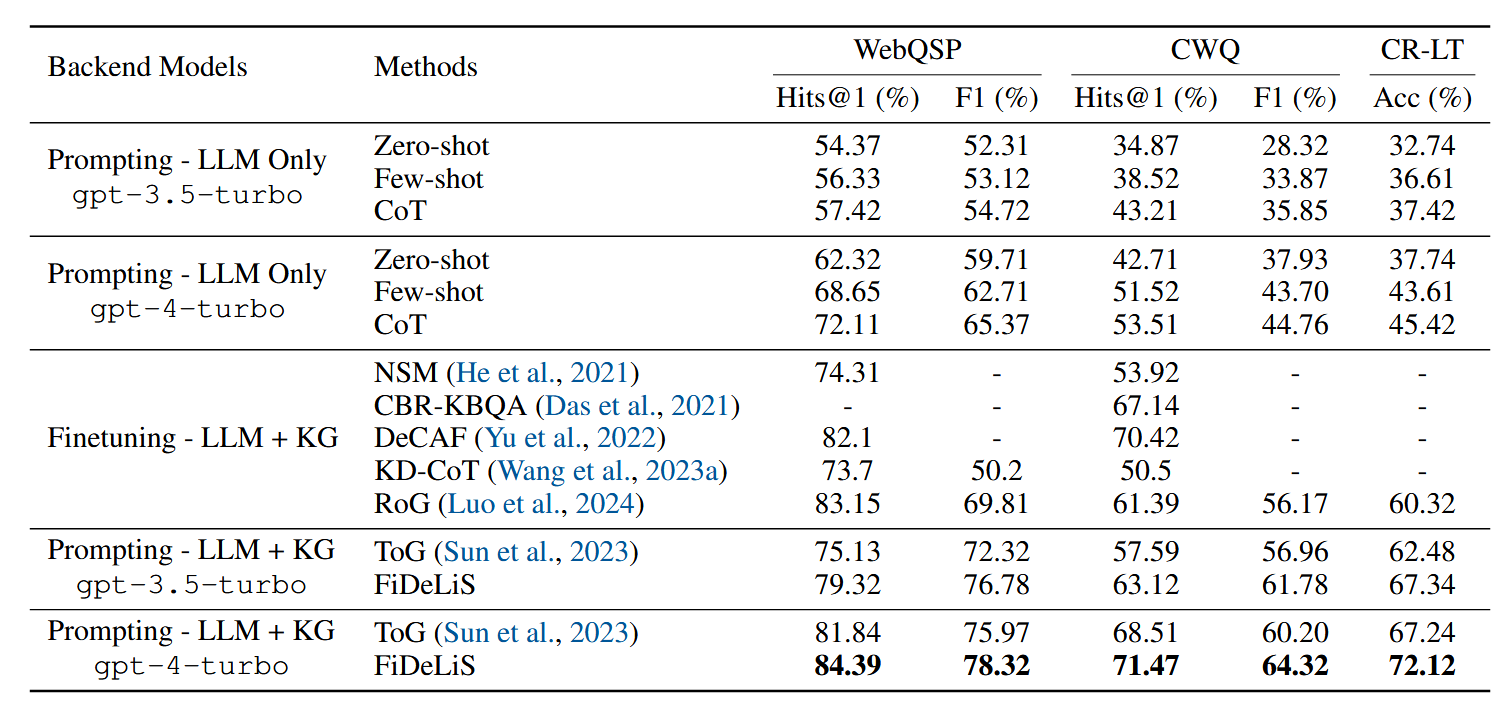
3）对现有推理路径进行评估，判断是否应该结束推理

* 大模型判断是否应停止推理，其中q’是问题的陈述形式。大模型需要判断由当前推理路径是否能推导出q’：

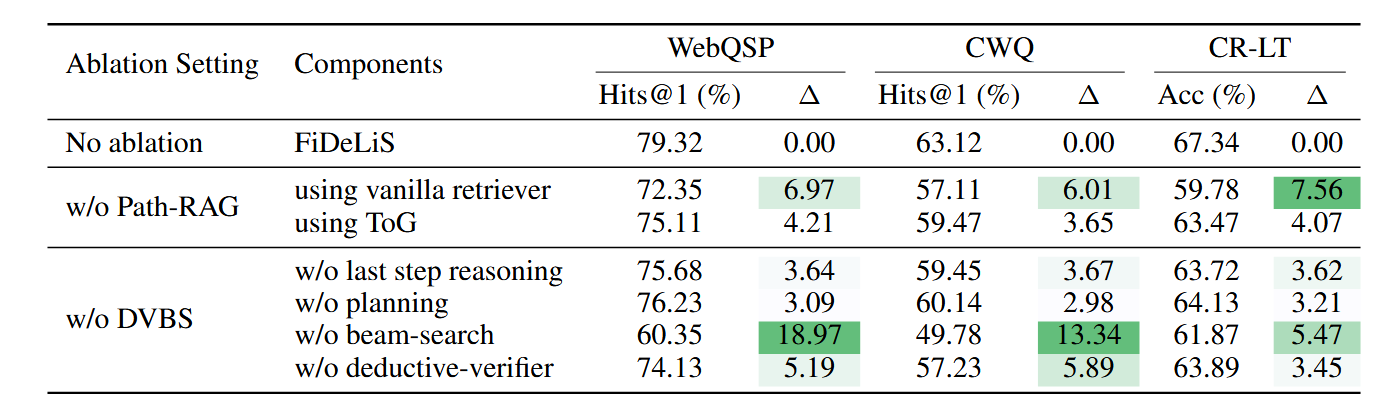


* 实验：

1）实验结果显示模型在使用gpt-4时达到sota，在使用gpt-3时也在多数情况下超过其他模型



2）消融实验表明，Path-RAG和DVBS都具有显著作用，其中beam search影响最大



3）针对效率的实验表明，使用Path-RAG能有效降低运行时间和消耗token数，使用GPT-4o和GPT-4o-mini效果更好

