

# On Generating Characteristic-rich Question Sets for QA Evaluation

## 研究动机

问题语句本身包含各种属性从而决定了问题的难度，比如数值分析，结构复杂（多条问题），多个答案，如果对于问题特征能够更加明确，可以对于问答系统进行更加细粒度的检测，在这项工作中，首次尝试生成具有明确指定特征的问题，并检查各种问题特征在QA中的影响。

## 研究贡献

- 1.提出的框架通过结合知识库中的结构化信息和来自Web的统计信息来过滤生成的低质量查询图。首先，我们在图形查询中识别冗余组件，并开发移除它们的技术。此外，基于从Web挖掘的实体、类和关系的频率，我们量化了图形查询的共性，并过滤掉了太少的查询。
- 2.采用半自动的方法将图形查询转换为自然语言问题，这提供了两个级别的释义：从网络中挖掘的实体的常见词汇形式（例如伊丽莎白女王和伊丽莎白女王陛下）用作实体释义，问题的其余部分由注释者解释。
- 3.使用提出的框架基于Freebase构建一个包含5000多个问题的新数据集，并广泛评估多个QA系统。

## 研究框架

提出了一个半自动化的框架，用于构建因子式问答（QA）数据集，其中一系列问题特征被形式化，包括结构复杂性、功能性、通用性、答案分类和释义。我们没有收集问题并手动描述它们，而是采用相反的过程，首先从知识库生成一种图形结构的逻辑形式，然后将其转换为问题

