[Data Science](#)[Natural Language Processing](#)[NLP Papers Summary](#)

Day 128: NLP Papers Summary – Neural Approaches To Conversational AI – KB-QA (Symbolic Methods)

By Ryan 7th May 2020 No Comments

Conversational QA system has become very popular in recent years and many techniques have been explored to build a good reliable QA system. Most QA systems allow users to either query from a knowledge base (KB-QA) or from a collection of documents (text-QA). We will first review the KB-QA space, both symbolic and neural methods and then the text-QA space.

In terms of KB-QA space, we will go over the symbolic methods in this blog post and then the neural methods in tomorrow's post. We will first review symbolic approaches in knowledge graphs and

showcase some of its limitations including scalability. We will then review different symbolic methods that have been proposed to address these issues. We will also be discussing multi-turn and conversational KB-QA agents which are not as well-studied.

What is a Knowledge Base?

A standard KB is a structured database that stores all the different entities and relationships in the form of subject-predicate-object triplets (s, r, t). In this case, s and t are the entities and r is the relation between the two entities. KB can be represented in the form of a knowledge graph as shown below, where entities are the nodes and relations are the directed edges.

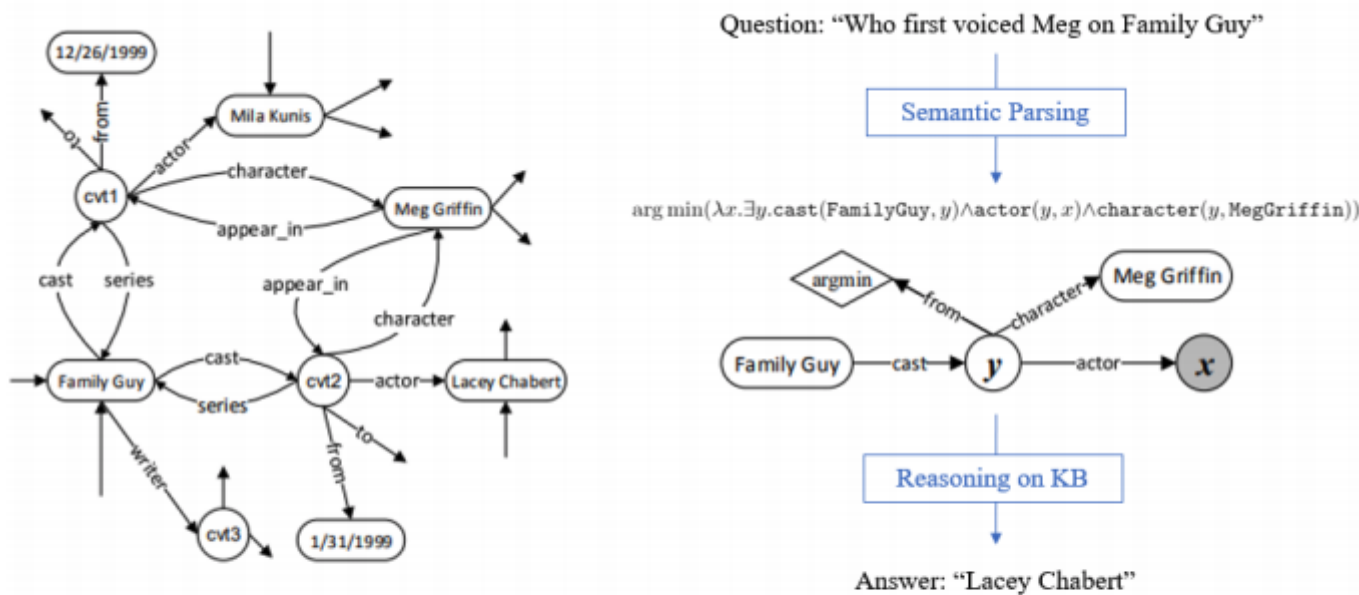


Figure 3.1: An example of semantic parsing for KB-QA. (Left) A subgraph of Freebase related to the TV show Family Guy. (Right) A question, its logical form in λ -calculus and query graph, and the answer. Figures adapted from Yih et al. (2015a).

Semantic Parsing for KB-QA

Figure above (on the right) showcase semantic parsing for KB-QA, which it's the core of many symbolic methods SOTA KB-QA systems. Questions are mapped to its formal meaning representations (logic form) and its equivalent graph representation (query graph), before it's feed into the KB. Reasoning will then happen within the KB through finding a sequence of paths hat match the query and retrieve the end nodes.

With the figure above, the two rounded rectangle represents the entities. The shared circle node x represents the answer node. The circle node y represents that there should be an entity

that describe the queried casting relations like character, actor, and time. The diamond \wedge de argmin provides a constraint to the answer. This query graph without the constraint would return “Lacey Chabert” and “Mila Kunis”. With the constraint, “Lucy Chabert” was returned as she started the role first.

There are two main challenges of symbolic KB-QA systems:

1. *Paraphrasing in language*. With semantic parsing, it is vulnerable to paraphrasing in natural language as different semantically equivalent questions can lead to different query graphs and subsequently different inaccurate answers
2. *Search complexity*. The number of paths grow exponentially with path length

Source: <https://arxiv.org/pdf/1809.08267.pdf>



Ryan

Data Scientist

Previous Post

Day 127: NLP
Papers Summary -
Neural Approaches
to Conversational
AI - Introduction

Next Post

Day 129: NLP
Papers Summary -
Neural Approaches
to Conversational
AI - KB-QA (Neural
Methods)





[Data Science](#) [Natural Language Processing](#) [NLP Papers Summary](#)

Day 365: NLP Papers Summary – A Survey on Knowledge Graph Embedding



Ryan

30th December 2020





[Data Science](#) [Implementation](#) [Natural Language Processing](#)

Day 364: Ryan's PhD Journey – OpenKE-PyTorch Library Analysis + code snippets for 11 KE models



Ryan

29th December 2020

