Knowledge Questions from Knowledge Graphs

Dominic Seyler (dseyler2@illinois.edu) University of Illinois at Urbana-Champaign

Mohamed Yahya (myahya6@bloomberg.net) Bloomberg LP, London

Klaus Berberich (kberberi@mpi-inf.mpg.de) Max Planck Institute for Informatics and HTW Saar
Which Italian Renaissance painter and inventor created Mona Lisa?

**Topic:** Painters  
**Answer:** Leonardo da Vinci  
**Difficulty:** Easy  
**Distractor Hard:** Michelangelo  
**Distractor Easy:** Vincent van Gogh
Question Generation Applications

- **Education** (e.g., learning assessment tests for students)
- **Professional training** (e.g., questionnaires about products for new employees)
- **Leisure** (e.g., quiz games)
- **Human Computing / Crowdsourcing** (e.g., generate test questions as honey pots)
Benefits of Question Generation Automation

• Saves human resources
• Enables to generate questions on large scale
• Automatic answer evaluation through multiple-choice
• Evaluate user expertise by inferring question difficulty automatically
"Bottom-up" Question Generation

SELECT ?x WHERE {
?x created Mona_Lisa .
?x type inventor .
?x It._Renaissance_painters
}

Which Italian Renaissance painter and inventor created Mona Lisa?

Answer Entity

Triple-Pattern Query

Natural Language Question
Question Generation Pipeline

\[ T \rightarrow \text{Query Generation} \rightarrow \text{Difficulty Estimation} \rightarrow \text{Query Verbalization} \rightarrow \text{Distractor Generation} \rightarrow QC \]

\( T = \text{Topic} \) (a set of entities related to \( T \))
\( Q = \text{Question} \) (question and correct answer)
\( MCQ = \text{Multiple Choice Question} \) (Q with incorrect answer options “distractors”)
SELECT ?x WHERE {
  ?x created Mona Lisa .
  ?x type inventors .
  ?x type It._Rennaissance_painters
}
Leonardo da Vinci

**easy**
- ?x type painter.
- ?x created Mona_Lisa.
- ?x created Vitruvian_Man.
- ?x created The_Last_Supper

**hard**
- ?x type scientist.
- ?x type engineer.
- ?x influences Victor_Bregeda.
- ?x created Portrait_of_a_Musician
Question Difficulty Estimation

- Ground Truth: Jeopardy! question – difficulty pairs
  - $200 Question -> Easy
  - $1000 Question -> Hard

- Annotation of entities with AIDA[1]

- Training and evaluation of logistic regression classifier

- Features based on:
  - Entity salience
  - Coherence of entity pairs
  - Entity types

Query Verbalization

• Verbalize using pattern:

Which verbalize(type₁), ..., and verbalize(typeₘ)
verbalize(p₁,o₁), ..., and verbalize(pₙ,oₙ)?

SELECT ?x WHERE {
?x created Mona_Lisa .
?x type inventors .
?x type It._Ren._painters
}

Which Italian Renaissance painter and inventor created Mona Lisa?
Distractor Generation

• Relax Query

```
SELECT ?x WHERE {
  ?x type It._Ren._painters
}
```

• All but one retrieved entities will be incorrect answers to target query
• Measure “confusability” between answer \((e_a)\) and distractor entity \((e_{dist})\):

\[
\text{conf}(Q, e_a, e_{dist}) = 1 - |P(\text{diff}(Q, e_a) = \text{easy}) - P(\text{diff}(Q, e_{dist}) = \text{easy})|
\]
Evaluation: Question Difficulty

1. Evaluation on held-out data with ten-fold cross validation

2. User study to evaluate difficulty ranking of questions - Kendall’s $\tau : 0.593$, indicating moderate agreement

<table>
<thead>
<tr>
<th>SAL</th>
<th>COH</th>
<th>TYPE</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>66.4%</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>65.8%</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>62.6%</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>62.2%</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>60.0%</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>57.8%</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>52.4%</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>no</td>
<td>50.0%</td>
</tr>
</tbody>
</table>
Evaluation: Distractor Confusability

• Crowdsourcing Experiment

• 400 Questions, each evaluated by five judges

• Evaluate whether judges agree with confusability estimate

• 76% of confusability estimates correct

• Cohen’s κ of 0.521, indicating moderate agreement
Additional Resources


Graph Characteristics

Facts (3)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;Dan_Crow_(computer_scientist)&gt;</td>
<td>&lt;worksAt&gt;</td>
<td>&lt;Google&gt;</td>
</tr>
<tr>
<td>&lt;Google&gt;</td>
<td>&lt;created&gt;</td>
<td>&lt;Postini&gt;</td>
</tr>
<tr>
<td>&lt;Google&gt;</td>
<td>rdf:type</td>
<td>&lt;wikicat_Companies_listed_on_NASDAQ&gt;</td>
</tr>
</tbody>
</table>

Verbalization

This company list on NASDAQ created Postini and has employee Dan Crow (computer scientist).
Summary

• Question generation applications: Education, Training, Leisure, etc.
• Can be generated on large scale and reduces human workload
• Generate question starting at the answer and retrieve question content from knowledge graph
• Represent question as query over knowledge graph and check for uniqueness of answer
• Train difficulty classifier using entity salience, coherence and type information from Jeopardy! ground truth
• Verbalize query using template
• Retrieve distractor answers by relaxing the question query and measure confusability
Knowledge Questions from Knowledge Graphs

Dominic Seyler  
(dseyler2@illinois.edu)  
Mohamed Yahya  
(myahya6@bloomberg.net)  
Klaus Berberich  
(kberberi@mpi-inf.mpg.de)