Constraint-Based Question Answering via Knowledge Graph

Ahmad Aghaebrahimian, Filip Jurčiček

Institute of Formal and Applied Linguistics
Faculty of Mathematics and Physics
Charles University
Content

• Preliminaries
• Introduction
• Algorithm
• Model
• Result
Constraint-Based Question Answering via Knowledge Graph

• Open factual questions
  – Open however limited to the domains in KG
  – Factual not open-ended
• Questions’ answer are available in KG
• Questions with one property and one entity

Like:
× How are you?
× Where did Beethoven and Mozart born?
✓ What are some works of John Milton? Paradise lost, Paradise regained ...
✓ What is the time zone in Dublin? Western European Time Zone

Simple QA is neither trivial nor solved!
What is the time zone in Dublin?
Preliminaries (Entity recognition)

- FB entity detection
  - Looking for every possible entity with every size which has a valid FB mid

**What is the time zone in Dublin** /m/02cft, /m/9jh, /m/kjh, ...

*Solution: Type enforcement (type constraints); Enforce expected type on the type of answers*

- FB entity disambiguation
  Maximizing similarity

  Using id and alias node to get the most similar entity

  - Dublin -> /m/9jh (/en/ohio_dublin)
  - Dublin -> /m/02cft (/en/dublin)

*Solution: Similarity enforcement (similarity constraints); Enforce to have maximal similarity with id and name nodes*
**Introduction**

**Training data**

SimpleQuestions (108K questions with KG assertions) (Bordes et al, 2015)

Question : Assertion

What is my time zone in Dublin? {(/m/02cft), (/…/time_zones), (/m/dfrg)}

---

**Input / Output _ training step**

**Input:**
- Training questions
- Assertions

**Output:**
- Trained classifier for properties

---

**Input / Output _ test step**

**Input:**
- Testing questions
- Knowledge graph
- Trained classifier

**Output:**
- Test assertions
**Algorithm**

- Train a classifier for property detection on training questions

- Test
  - Given a question
    - Property detection (using the classifier)
    - Entity recognition (via KG validation)
    - Logical form per property/ID
    - Return the assertion (entity, property, answer)

- Evaluation (based on Boredes et al., 2015)
  - Path accuracy:
    A predicted assertion is correct when both the property and the entity mid are correct.
Model (Property detection – LR Model)

\[ N - \text{best property}(q) = P(p \mid q) \]
\[ P(p \mid q) = \frac{\exp(\omega_p^T \varphi(q))}{\sum_{p_i} \exp(\omega_{p_i}^T \varphi(q))} \]

\( q: \) question

\( \varphi(q) \in \mathbb{R}^{20 \times 128}: \) q features in vector space

\( p: \) properties of the dataset (~1800 properties)
  - /location/location/time_zone
  - /people/person/place_of_birth
  - ...

\( \omega: \) vector of properties’ weights (optimized in LR model)
Model (Entity recognition)

\[
\text{best match entity}(q) = \text{maximize } \alpha^T s(p_{q\cdot e_q})
\]

\(q: \text{ test question}\)

\(\alpha^T: \text{ vector of indicator variables}\)

\([p_{q\cdot e_q}] \in P_q \times E_q\)

\(P_q: \text{ space of N-best properties for } q \hspace{1em} (\sim 30 \text{ - } 100)\)

\(E_q: \text{ space of valid entities in } q \hspace{1em} (\sim 1000 \text{ - } 5000)\)

\(s(p_{q\cdot e_q}): \text{ vector of } p_q \text{ probabilities}\)
Model (Entity recognition) : example

best match entity(q) = maximize α^T s(p_q-e_q)

q: what is the time zone in Dublin?  

E_q : zone, Dublin(Ireland), Dublin(Ohio)

P_q : /people/person/place_of_birth: 0.70 (/location/location)  
     /location/location/time_zones: 0.60 (/time/time_zone)  
     /music/artist/genre: 0.50 (/music/genre)

\[ (p_q-e_q) \in P_q \times E_q \]
\[ s(p_q-e_q) : \text{vector of } p_q \text{ probabilities} \]

- 0 ≤ α_i ≤ 1 and integer
- \( \sum \alpha \leq 1 \)
- C1 for \( p_q-e_q \) is true if expected type(p_q) ∈ types(ans(p_q,e_q))
- C2 for \( p_q-e_i \) is true if sim(name(e_i), id(e_i)) = max_{e_q} (sim(name(e_q), id(e_q)))
• State of the art result

<table>
<thead>
<tr>
<th></th>
<th>Trained on</th>
<th>Accuracy (prop./overall)</th>
<th>Knowledge graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bordes et al.</td>
<td>SQ</td>
<td>* / 61.6%</td>
<td>FB5M</td>
</tr>
<tr>
<td>Constraint-based(LR)</td>
<td>SQ</td>
<td>70% / 61.20%</td>
<td>Full Freebase</td>
</tr>
<tr>
<td>Constraint-based(NN)</td>
<td>SQ</td>
<td>74% / 63.89%</td>
<td>Full Freebase</td>
</tr>
<tr>
<td>Constraint-based(CNN_1 layer)</td>
<td>SQ</td>
<td>78% / 65.16%</td>
<td>Full Freebase</td>
</tr>
<tr>
<td>Constraint-based(CNN_2 layers)</td>
<td>SQ</td>
<td>78.8% / 65.19%</td>
<td>Full Freebase</td>
</tr>
</tbody>
</table>
Thank you!