Open Relation Extraction for Support Passage Retrieval: Merit and Open Issues
Amina Kadry and Laura Dietz

ABSTRACT
Our goal is to complement an entity ranking with human-readable explanations of how those retrieved entities are connected to the information need. Relation extraction technology should aid in finding such support passages, especially in combination with entities and query terms. This work explores how, the ClausIE system, a current state of the art method for unsupervised relation extraction (OpenIE) contributes to a solution for the task, assessing potential, limitations, and avenues for further investigation.

1 TASK & IDEA
Task:

Why is \( \xi \) relevant for \( Q \)?

\( \xi \) is an ingredient for \( Q \).

\( \xi \) states that \( Q \) is married to \( \xi \).

\( \xi \) is married to \( \xi \).

=> Good explanation
why \( \xi \) is relevant for \( Q \)

2 FOUNDATION: CLAUSEIE

Phase 1. Clause types.
Complementary clause \( SV(C(set_{s}, Varepsilon, Srules_{i}, A7, of_{f}) \)
Adverbial clause \( SVA(V playa_{d}, Ssport_{16}, Sby_{13}) \)

Phase 2. Propositions of relation tuples.
The rules of golf are a standard set of regulations
The rules of golf are a standard set of procedures
The rules of golf should be played by a standard set of regulations
The sport of golf should be played by a standard set of procedures

Assumption: Whenever entity and query terms are contained in the same proposition, this sentence is likely to explain the connection between query and entity.

3 FEATURES FOR RELEVANCE EXPLANATIONS

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>1. sentence length measured in number of words</td>
</tr>
<tr>
<td></td>
<td>2. sentence position measured as a fraction of the document</td>
</tr>
<tr>
<td></td>
<td>3. fraction words that are stop words</td>
</tr>
<tr>
<td></td>
<td>4. fraction of query terms covered by sentence</td>
</tr>
<tr>
<td></td>
<td>5. sum of IDF of query terms (IDF is inverse document frequency)</td>
</tr>
<tr>
<td></td>
<td>6. average of IDF of query terms</td>
</tr>
<tr>
<td></td>
<td>7. sum of TF-IDF of query terms</td>
</tr>
<tr>
<td></td>
<td>8. number of entities mentioned</td>
</tr>
<tr>
<td>MLP</td>
<td>9 - 12 for nouns/verbs/adjectives/adverbs: fraction of words with POS tag</td>
</tr>
<tr>
<td></td>
<td>13 whether sentence contains a named entity</td>
</tr>
<tr>
<td></td>
<td>14 - 16 for NER types PERSON/GPE: whether NER type is contained</td>
</tr>
<tr>
<td>DP</td>
<td>17 number of edges on the path between two entities in dependency tree</td>
</tr>
<tr>
<td></td>
<td>18 indicate whether path goes through root node</td>
</tr>
<tr>
<td></td>
<td>19 indicate whether path goes through query term</td>
</tr>
</tbody>
</table>

ClausIE
20 whether ClausIE generated an extraction from this sentence
21 - 27 for all seven clause types: whether clause of this type is extracted
28 proposition length measured in tokens
29 maximum constituent length (size of dependency tree) in proposition
30 - 33 for subject/object/bid: if another entity is in subject and/or object position of the proposition
34 for subject/object/bid: if given entity is in position of proposition
35 - 34 for subject/object position: if entity is in position of proposition
36 - 34 for subject/object position: if an entity link in is in position of prop.
37 - 34 for subject/object position: if a query term (ignoring stopword) is in position of proposition
42 - 43 for subject/object position: if a nested entity (NER) is in position of proposition

4 EXPERIMENTAL EVALUATION
Test collection, available: www.cs.unh.edu/~dietz/appendix/openie4r
- ten 2013/2014 TREC Web track queries
- and (up to) 10 relevant entities per query from RWSQ gold standard
- resulting in 75 support passage rankings (one per query and entity).
- 22,731 annotations for candidate sentences from entity's Wiki page
- 2,906 are marked as relevant for AQ1

Assessors imagine writing knowledge article on the topic \( Q \). Which information to include about the given entity \( e_i \)?

AQ1) Explanation: Does the sentence explain the relevance of entity \( e_i \)?
AQ2) Relation: Does the sentence mention any relationship involving \( e_i \)?
AQ3) Rel rel: Is this relationship relevant for the explanation?
AQ4) ClausIE: Does ClausIE extract a valid relationship from sentence?
AQ5) ClausIE rel: Is ClausIE's extraction relevant for the explanation?

We study these annotations in combination with two heuristics:
Qterm: Does the sentence include query terms (stopwords ignored)?
Name: Does the sentence include the entity's name?

4.1 Experiment 1: Any hope?

Table 1: Performance of AQ1–5 as predictors for explanations.

<table>
<thead>
<tr>
<th>Relation</th>
<th>Rel rel</th>
<th>ClausIE</th>
<th>ClausIE rel</th>
<th>Qterm (+)</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prec (+)</td>
<td>0.46 ± 0.06</td>
<td>0.52 ± 0.05</td>
<td>0.45 ± 0.05</td>
<td>0.45 ± 0.05</td>
<td>0.38 ± 0.04</td>
</tr>
<tr>
<td>Recall</td>
<td>0.38 ± 0.08</td>
<td>0.21 ± 0.11</td>
<td>0.20 ± 0.03</td>
<td>0.14 ± 0.01</td>
<td>0.49 ± 0.04</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.27</td>
<td>0.52</td>
<td>0.53</td>
<td>0.49</td>
<td>0.47</td>
</tr>
<tr>
<td>Count</td>
<td>11767 (0)</td>
<td>1955 (0)</td>
<td>11172 (0)</td>
<td>636 (3)</td>
<td>6476 (398)</td>
</tr>
</tbody>
</table>

4.2 Experiment 2: Automatic Method with LTR

Figure 1: Ranking of sentences explaining entity relevance.

4.3 Experiment 3: Open Issues
- Among all sentences that express a relation, ClausIE is missing this relation in 32% of the cases.
- Only half of the sentences with relation expressions actually actually contain a relation that is relevant for the query-entity pair.
- Only 636 sentences with relevant ClausIE extractions (3%) of all 22,731 annotated sentences.
- In contrast, our data set contains 2906 sentences (13%) with explanations of relevance.
- An ideal ranking would obtain a MAP value of 0.41 (theoretical upper bound). Equals ClausIE feature set alone, MAP 0.41, cf. Figure 1.
- We conclude that our approach obtains an optimal ranking under limitations imposed by the off-the-shelf OpenIE system.
- Improving coverage of OpenIE systems is likely to translate to immediate quality improvements for text-ranking tasks.

http://mschulme.github.io/rewoq/