An Informativeness Approach to Open Information Extraction Evaluation

William Léchelle, Philippe Langlais

RALI — University of Montreal

April 8, 2016
1. Open Information Extraction

2. Evaluation,

3. an Informativeness Approach
Open Information Extraction

Evaluation,

an Informativeness Approach
Open IE

- Information Extraction without pre-specified relations
- pattern-writing / training data is expensive
Open IE

- Information Extraction without pre-specified relations
  - pattern-writing / training data is expensive
- KnowItAll project at University of Washington
- TextRunner (Banko 2007)
Basic

- **Reverb** (Fader 2011)
- POS tags patterns, looking for **NP VP NP**
• **Reverb** (Fader 2011)
• POS tags patterns, looking for NP VP NP
• Redundancy $\rightarrow$ confidence
  • unsupervised
  • requires lots of text
Basic

- **Reverb** (Fader 2011)
- POS tags patterns, looking for NP VP NP
- Redundancy → confidence
  - unsupervised
  - requires lots of text
- fast
• **Ollie**: learning dependency patterns bootstrapping over Reverb
• **ClausIE**: analysis of clauses
1 Open Information Extraction

2 Evaluation,

3 an Informativeness Approach
Current methods

- Precision / yield / AUC
- Sampling + manual correctness judgement
Problem

- Sampling + manual correctness judgement
Problem

- Sampling + manual correctness judgement
- costly!
Problem

- Sampling + manual correctness judgement
  - costly!
  - Arguable decisions. Criteria?

Sentence: That’s a lot of maybes in a sport where the right thing seldom happens, but [given X, Y should Z if he wants T].

Fact: (That; is; a lot of maybes) for sure...
Arguable decisions (cont.)

Sentence: In response, a group of Amherst College students held a patriotism rally in October, reciting the Pledge of Allegiance.

Fact: (a group of Amherst College students; held; a patriotism rally) why? when?

Sentence: The scandal has now forced resignations at Japan’s fourth-largest bank and three of Japan’s Big Four brokerages.

Fact: (Japan; has; fourth-largest bank) informative?
Past guidelines

Ollie:

- “Two annotators tagged the extractions as correct if the sentence asserted or implied that the relation was true.”

ClausIE:

“We also asked labelers to be liberal with respect to coreference or entity resolution; e.g., a proposition such as (‘he’, ‘has’, ‘office’), or any unlemmatized version thereof, is treated as correct.”
Our goals

- evaluation script
Our goals

- evaluation script
- less subjective (if possible)
Our goals

- evaluation script
- less subjective (if possible)
- oriented towards informativeness / usefulness
An Informativeness Approach to Open Information Extraction Evaluation

1. Open Information Extraction

2. Evaluation,

3. an Informativeness Approach
Simplistic QA

- Ask questions about the text.
- Answer them with IE.
Simplistic QA

- Ask questions about the text.
- Answer them with IE.

- Given some input text, annotate all factoid questions that can be answered by information contained in that text, and the answers. Run the extractors on the input text. The evaluation metric is the number of questions that can be answered using only the output of each system.
Annotation process

• Annotate all factoid questions that can be answered with that text.
Annotation process

- Annotate all factoid questions that can be answered with that text.
- ... that make sense! (and are useful)
Annotation process

- Annotate all factoid questions that can be answered with that text.
- ... that make sense! (and are useful)
- Annotator can see OIE output
Example

S: Robert Barnard (born 23 November 1936) is an English crime writer, critic and lecturer.

Q: Who is Robert Barnard?
A: an English crime writer

Q: When is Robert Barnard born?
A: 23 November 1936
**Example**

_S_: Mrs. Yogeswaran was shot five times with a pistol near her Jaffna home on May 17, 1998.

_Q_: What was Mrs. Yogeswaran shot with?
_A_: a pistol

_Q_: How many times was Mrs. Yogeswaran shot?
_A_: five
Difficulties

- The hardest part is dealing with ambiguity
  - Some facts only make sense in context
  - What do entities refer to?
  - Pronouns
S: While the Arab world is a rich prize in itself, Europe has been and remains the primary objective.
E: (Europe ; remains ; the primary objective)
E: (the Arab world ; is ; a rich prize)
S: While the Arab world is a rich prize in itself, Europe has been and remains the primary objective.
E: (Europe ; remains ; the primary objective)
E: (the Arab world ; is ; a rich prize)

- Context is missing
Ambiguity

S: Wu worked as a reporter for United Press International from 1973 until 1978 when she joined WGBH-TV, Boston’s public television station, as the Massachusetts State House reporter until 1983.
E: (Wu; worked as a reporter for; United Press International)

S: Daughter of the actor Ismael Sanchez Abellan and actress and writer Ana Maria Bueno, Gabriel was born in San Fernando, Cadiz ...
E: (Gabriel; was born in; San Fernando)
Ambiguity

**S:** Wu worked as a reporter for United Press International from 1973 until 1978 when she joined WGBH-TV, Boston’s public television station, as the Massachusetts State House reporter until 1983.

**E:** (Wu; worked as a reporter for; United Press International)

**S:** Daughter of the actor Ismael Sanchez Abellan and actress and writer Ana Maria Bueno, Gabriel was born in San Fernando, Cadiz ...

**E:** (Gabriel; was born in; San Fernando)

- What can we ask about?
Ambiguity

S: Wu worked as a reporter for United Press International from 1973 until 1978 when she joined WGBH-TV, Boston's public television station, as the Massachusetts State House reporter until 1983.
E: (Wu; worked as a reporter for; United Press International)

S: Daughter of the actor Ismael Sanchez Abellan and actress and writer Ana Maria Bueno, Gabriel was born in San Fernando, Cadiz ...
E: (Gabriel; was born in; San Fernando)

- What can we ask about?
  - Einstein - Wu - Gabriel - he
  - We drew a line in the sand
Coreference

S: He was born in New York and died at Livonia, Michigan.
E: (He; was born in; New York)
S: Had I known then what I know now, I might have argued for a different arrangement.
Coreference

S: He was born in New York and died at Livonia, Michigan.
E: (He; was born in; New York)
S: Had I known then what I know now, I might have argued for a different arrangement.

- Coreference is key
An Informativeness Approach to Open Information Extraction Evaluation

William Léchelle, Philippe Langlais

Dataset

- Public dataset already in use, distributed by ClausIE
- 900 random sentences from Web, Wikipedia, NYT
- We annotated 100 sentences, with 100 Q/A pairs.
  - 3 Q/A per sentence, but 2/3 sentences are empty.

*S: His teammates indeed loved the show.*
Evaluation metric

- Automatic script to answer questions using extracted information
- Basic string-matching
- Goal: be simple and fair, not sophisticated and powerful
Q: What was Mrs. Yogeswaran shot with?

Match words: mrs. shot what yogeswaran

Facts

| Esaka; will quit; to take responsibility for 67.28 million yen in payoffs | 0.0 |

| Mrs. Yogeswaran; was shot; five times with a pistol near her Jaffna home on May 17 1998 | 1.0 | Correct |

| Mrs. Yogeswaran; was shot with; a pistol | 1.0 | Correct |

| Ashcroft; said; through Mr. Hilton that he had made the point that there would be no peace between him and the Governor until ... | 0.75 | Wrong |

A: a pistol
## Results

<table>
<thead>
<tr>
<th></th>
<th>Answered</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReVerb</td>
<td>35%</td>
<td>19%</td>
<td>6%</td>
</tr>
<tr>
<td>Ollie</td>
<td>39%</td>
<td>43%</td>
<td>17%</td>
</tr>
<tr>
<td>ClausIE</td>
<td>68%</td>
<td>42%</td>
<td>29%</td>
</tr>
<tr>
<td>All</td>
<td>71%</td>
<td>53%</td>
<td>38%</td>
</tr>
</tbody>
</table>

- ClausIE > Ollie > ReVerb
Future Work

- More data!
  - Application focus?
- Annotate full documents and quantify inter-sentential information
Summary

- Automatic evaluation system based on some simple question/answer pairs
- Small corpus of 100 Q/A pairs
- 40% recall for clear cut facts, without intersentential coref, using currently available technology
- Evaluation script + annotated data: http://www.CICLing.org/2016/data/92
Questions?