A Survey of Open Domain Event Extraction

Presented by: Zecheng Zhang, Yuncheng Wu, Zesheng Wang
Outline

Introduction
Types Of Event Extraction
Data And Ontologies
Contemporary Approaches Evaluation
Related Techniques In Our Approach
Conclusions
Introduction
Introduction

- Text data increases rapidly.
- Most text data is unstructured in machine view, hard to mine knowledge automatically.
- Many systems such as information retrieval systems were created.
- Event extraction is one of the most useful techniques in IR/IE systems.

Hogenboom, Frederik, et al. “An overview of event extraction from text”.
A Simple Example on Event Extraction

- The purpose of current event extraction is to detect type of events and extract arguments with different roles from human readable text.

```
Destroyer, :ARG0

Students in Siebel Center destroyed a computer last month.

Target, :ARG1

Location, :ARG2

Time, :time
```
Types Of Event Extraction
Types Of Event Extraction - By Approaches

- Data-driven approach: statistical models and global-based patterns of event schema.
- Semantic-driven approach: lexical knowledge and relative local semantic pattern.
- Hybrid-driven approach: combines data-driven approach and semantic-driven approach.

Hogenboom, Frederik, et al. “An overview of event extraction from text”.
Types Of Event Extraction - By Objectives

- Many event extractions are applied on some specific realms such as biology, finance and medicine.
- Open domain event extraction.
Data And Ontologies
FreeBase

- A semantic knowledge base
- Uses Compound Value Type (CVT) to combine multiple values into single value

Remedy Corp was sold to BMC Software as the Service Management Business Unit in 2004.

<table>
<thead>
<tr>
<th>id</th>
<th>company_acquired</th>
<th>acquiring_company</th>
<th>date</th>
<th>divisions_formed</th>
</tr>
</thead>
<tbody>
<tr>
<td>m.07bh4j7</td>
<td>Remedy Corp</td>
<td>BMC Software</td>
<td>2004</td>
<td>Service Management Business Unit</td>
</tr>
</tbody>
</table>
Data And Ontologies - Data Resources

FrameNet:

- A linguistic resource storing information about lexical and predicate argument semantics.
- Each frame of FrameNet has a set of lemmas with POS tags that can evoke the frame which are called Lexical Units.
- Example

Michelle baked her mother a cake for her birthday.

- Lexical Unit: Bake(v.)
- Frame: Cooking_creation;
- Can be mapped to Cooking_creation in FreeBase
Data And Ontologies - Data Resources

Wikipedia

- One of the largest, semi-structured and wide-coverage knowledge base.
- A multilingual based knowledge base that has rapid increasing number of articles.
- Helpful for entity linking and entity extraction
- Application: Wikification
Motivation

● Open domain extraction usually need large amount of labeled data.
● Disadvantages of small hand-labeled data
  ○ Expensive to produce
  ○ Low coverage of event types
  ○ Limited in size

Approaches

● Generate data by identifying trigger words and key arguments
● Generate data by identifying key arguments only
Generate data by identifying trigger words and key arguments - Intuition

Yesterday, some demonstrators threw stones at soldiers in Israeli.

- Threw: trigger word of Attack event
- Underlined words: arguments of Attack event
Generate data by identifying trigger words and key arguments - General Idea

- Key argument extraction: extract arguments that serves as vital clues to distinguish events
  - Role Saliency: Saliency of an argument to represent a specific event instance given event type
  - Event Relevanace: Ability that an argument can be used to discriminate different event types
- Trigger word detection: extract triggers for each event type
  - Verb tends to express occurrence of an event in a sentence
  - Idea similar to TF-IDF
- Trigger words filtering and expansion: include words other than verbs
  - Example: *Marriage* triggers *wedding* event
- Automatic labeled data generation
  - Find trigger word frame in FrameNet and map it to FreeBase

Chen, et al. Automatically labeled data generation for large scale event extraction
Generate data by only identifying key arguments - Intuition

- Several key arguments together imply the event type
- Argument roles and values define \textit{business.acquisition} comprehensively.

\textbf{Remedy Corp} was sold to \textbf{BMC Software} as the \textbf{Service Management Business} Unit in \textbf{2004}.

\textit{company acquired} \hspace{1cm} \textit{acquiring company} \hspace{1cm} \textit{divisions formed} \hspace{1cm} \textit{date}
Generate data by only identifying key arguments - General idea

- Extract key arguments
  - Include arguments with high importance score.
  - Include time-related argument

- Generate training data
  - Take existing structured tables or lists like FreeBase CVT tables to label events.
  - Use alias information to match arguments with different surface names. (Microsoft vs. MS)
Contemporary Approaches
Evaluation
Contemporary Approaches Evaluation

Event Schema Extraction By Trigger Clustering

Zero-Shot Learning
Libera Event Extraction By Trigger Clustering

- Identify candidate triggers and arguments
- Disambiguate senses and link to OntoNotes senses.
- Represent trigger and argument in distributional vectors
- Compose and represent event structure
- Cluster with joint constraints

Huang, et al. Liberal event extraction and event schema induction
Zero-shot relation extraction via reading comprehension

- Relation extraction can be reduced to answering simple reading comprehension questions, by designing one or more natural-language questions for each argument slots.
- Recent reading comprehension techniques (RNNs, BiDAF Model) can be extended to learning relation models.
- For given relation models, distant supervision can be combined with designed crowd-source questions to generate large effective training set.
- Zero-shot learning approach can be applied to handle new relation types that are only specified at test-time, whose labeled training examples are not provided.

Zero-Shot Transfer Learning for Event Extraction

- Many traditional supervised methods usually do not have ability to handle new event types since new event types mean we cannot use annotations from old events.
- Event types and event mentions can be represented by some forms of structure.
- Zero-Shot Learning: solves a classification problem when there is not enough training labels available for some labels.

Related Techniques In Our Approach
Local Approach:

each mention is disambiguated individually, utilizing common features such as textual similarity between the given document and each candidates Wikipedia disambiguation page.

Global Approach:

based on the assumption such that, if all mentions inside the same document are disambiguated correctly, those disambiguations will tend to form a set of related concepts with certain level of coherence.
Wikification - cont.

**GLOW:** Two-stage optimization framework that utilizes both local and global features. Generalize both local and global features, NGD, PMI and average/maximum of the above features (in longer document, which may cover more sub-topics, maximum might be more informative than simple average).

**Ranker:** First stage, ranker is trained to generate top non-null disambiguation candidates for a entity mention.

**Linker:** Second stage, linker is used to determine which candidate title should the entity mention disambiguated to, or null. The null possibility is handled specifically to account for one of the following three cases: 1) the queried mention does not have a corresponding Wikipedia page, 2) the queried mention does have a corresponding Wikipedia page, but is not included in the candidate list and 3) the ranker made a completely wrong decision that chose an incorrect disambiguation.

Ratinov, Lev, et al. Local and global algorithms for disambiguation to wikipedia.
Coreference Resolution - Intuition

- Definition: Recognize which mentions in text refer to the same real-world entity.
- Motivation: Capture more semantic information in sentences.
- Example

  Tom is eating an apple, but he likes pear the best.

  With coreference resolution, machine will treat Tom and he as the same entity.

  Machine will know Tom likes pear the best.
Coreference Resolution - General Idea

- **Mention pair encoder**
  - Capture features of a mention m and a candidate antecedent a, and encode the mention pair into a vector.

- **Cluster pair encoder**
  - Capture entity level information instead of mention level information.
  - Pooling from mention pair encoder results.

- **Mention ranking model**
  - Learns effective weights for mention-pair encoder. (pretraining of cluster ranking model)
  - Scores can be used to determine a cluster decision is right or wrong.

- **Cluster ranking model**
  - Consider global information between paris of mentions in entity level.

Clark, et al. Improving coreference resolution by learning entity-level distributed representations
Conclusions
Conclusions

- Objective: Open domain event extraction.
- Event extraction types
- Data: Useful knowledge bases and data expansion techniques.
- Contemporary approaches
  - Trigger clustering
  - Deep Learning
- Related techniques in our approach
  - Wikification
  - Coreference resolution
Thank you !