Event Coreference Resolution using Convolutional Neural Networks

Nirbhay Modhe Parag Bansal

[1]Department of Computer Science
Indian Institute of Technology Kanpur

Outline

Outline

What is Coreference Resolution?

Parag broke his arm when he fell down the stairs. It was Nirbhay who pushed him.

- Two or more expressions in a text are **coreferent** if they refer to the same entity.

What is Coreference Resolution?

Parag broke his arm when he fell down the stairs. It was Nirbhay who pushed him.

- Two or more expressions in a text are **coreferent** if they refer to the same entity.
- The task of identifying coreferent expressions is called **Entity Coreference Resolution**

What is an Event?

- An event is any occurrence or happening, typically associated with a trigger word or phrase called an **event trigger**.

  In Baghdad, a cameraman died when an American tank fired on the Palestine Hotel.

What is an Event?

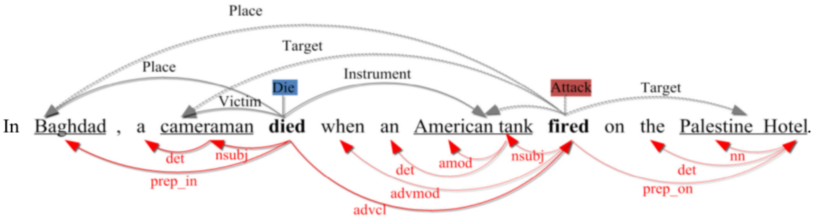
- An event is any occurrence or happening, typically associated with a trigger word or phrase called an **event trigger**.

  In Baghdad, a cameraman <u>died</u> when an American tank <u>fired</u> on the Palestine Hotel.

- What does the event **Die** consist of?
  - Trigger Word(s) - "**died**"
  - Victim - "**cameraman**"
  - Instrument - "**American Tank**"
  - Place - "**Baghdad**"

Outline

Identifying an Event



- Upper side: Event mentions **Die** and **Attack**
- Lower size: A subset of dependency parse tags

Source: Chen et. al. in "*Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks*"

Inherent Ambiguity in Events

- Multiple events in a sentences, overlap in arguments

Inherent Ambiguity in Events

- Multiple events in a sentences, overlap in arguments
- Same trigger words, different event types

  E1: *Donald Trump beats Clinton.*
  E2: *Muhammad Ali beats his opponent.*

Inherent Ambiguity in Events

- Multiple events in a sentences, overlap in arguments
- Same trigger words, different event types

  E1: *Donald Trump <u>beats</u> Clinton.*
  E2: *Muhammad Ali <u>beats</u> his opponent.*

- Trigger words are not discriminatory enough!

Event Extraction

A Machine Learning Approach

- The aim is learn how to differentiate between true and false event triggers.
- What discriminatory information can be used to automatically identify true event triggers?

Event Extraction

A Machine Learning Approach

- The aim is learn how to differentiate between true and false event triggers.
- What discriminatory information can be used to automatically identify true event triggers?
  - Syntactic features such as POS, dependency tags.

Event Extraction

A Machine Learning Approach

- The aim is learn how to differentiate between true and false event triggers.
- What discriminatory information can be used to automatically identify true event triggers?
  - Syntactic features such as POS, dependency tags.
  - Semantic information, estimated using word vectors.

Event Extraction

A Machine Learning Approach

- The aim is learn how to differentiate between true and false event triggers.
- What discriminatory information can be used to automatically identify true event triggers?
  - Syntactic features such as POS, dependency tags.
  - Semantic information, estimated using word vectors.
- Learning features automatically using deep learning.

Related Work

Event Extraction using CNNs

- Recently Chen et. al. proposed a convolutional neural network approach for event extraction.
- Incorporated multiple event sentences by using a dynamic multi-pooling framework.

Chen's CNN Architecture

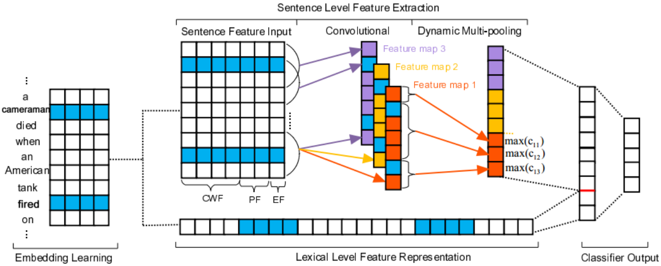


Figure 2: The architecture for the stage of argument classification in the event extraction. It illustrates the processing of one instance with the predict trigger *fired* and the candidate argument *cameraman*.

Source: Chen et. al. in "*Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks*"

Outline

Coreference Resolution

- Classifying event pairs as coreferent or not - simple binary classification?

Coreference Resolution

- Classifying event pairs as coreferent or not - simple binary classification?
- Best performance (a coreference metric related to F-score) until 2015 was around 60 for the ACE 2005 dataset.

Coreference Resolution

- Classifying event pairs as coreferent or not - simple binary classification?
- Best performance (a coreference metric related to F-score) until 2015 was around 60 for the ACE 2005 dataset.
- We aim to empirically analyse the performance of a CNN approach for feature extraction.

Our CNN Architecture

- Features of the two events taken by the CNN are:
  - Word vectors of event trigger context window
  - Position feature embeddings of event triggers
- A single convolutional layer with multiple convolution filter sizes, followed by a dynamic max-pooling layer
- Three FC (fully connected) layers followed by softmax over the three predicted classes:
  - 1 - First event is true or false
  - 2 - Second event is true or false
  - 3 - The two events are coreferent or not

Conclusion

- Our joint CNN architecture is the first such approach for Coreference Resolution and Event Extraction
- We hope to test whether sequence learning methods (RNNs) can provide better features for coreference resolution.

Thank You