Acquire common sense knowledge between events via weakly supervised approach

Wenlin Yao

TEXAS A&M UNIVERSITY

The Relationship Between Events

Voters go to the polls

Election ceremony

Voters go to the polls *before* Election ceremony

Rescue residents

Hurricane struck city

Rescue residents *after* Hurricane struck city

The Relationship Between Events

Typhoon Haiyan struck the eastern Philippines on Friday,

BEFORE

CAUSE

which killed thousands of people.

Temporal Relations: create event timelines, document summarization

Causal Relations: predict future events, risk analysis

Motivation

- Extract common sense relations between events (*happens_before* and *happens_after* relations)

Observation

- Regular event pairs: Event pairs that tend to show the same temporal relation despite of various contexts
- attacks *before* PEOPLE be arrested
    - Under pressure following suicide attacks, police arrested scores of activists on Monday.
    - Two men were arrested on suspicion of carrying out the Mumbai attacks.
    - Carlos was arrested in Sudan in August in connection with two bomb attacks in France in 1982.

System Overview

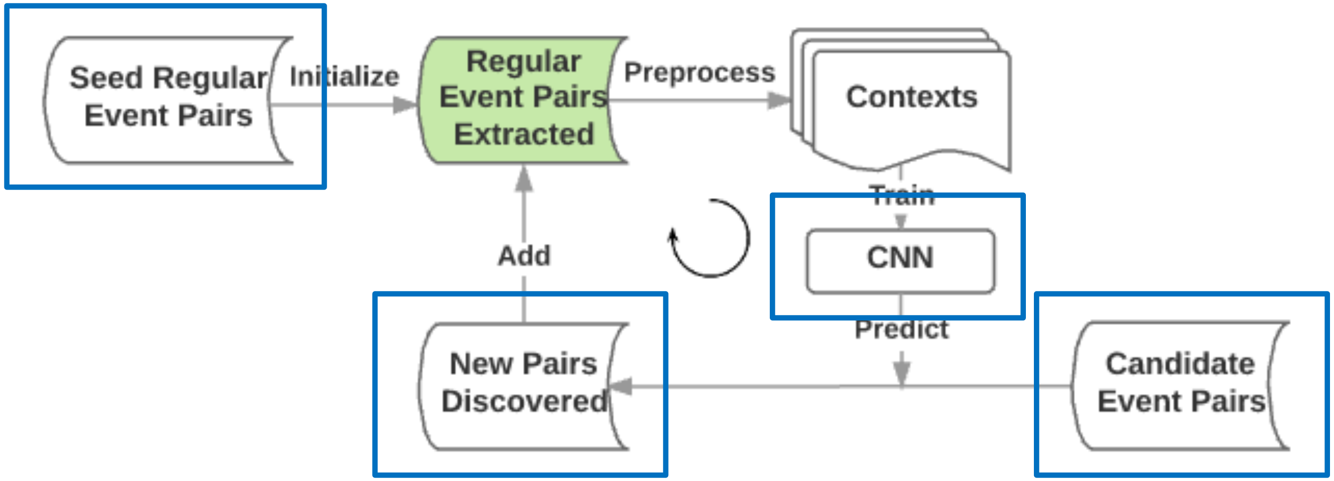


Figure: Overview of the bootstrapping system

- Data: English Gigaword (Napoles et al., 2012)
  - 10 million documents from seven news outlets (e.g., New York Times, Washington Post, etc.)
  - Stanford CoreNLP tools to tokenize, POS tag, parse, etc.

Event Representation

- Goal: Make individual events expressive and self-contained
- Verb events (use Stanford dependency relation)
  - Transitive verb: include the direct object (e.g., win lottery)
  - Intransitive verb: include the agent (e.g., water evaporates)
- Noun events (e.g., attack, election, etc.)
- Use named entity types (NER) to replace specific name
  - visit to Location ← visit to New York, visit to Houston

Seed/Candidate Event Pairs

- Governor and dependent word of pattern *after* and *before*
  - He worked in a company in New York *after* graduation.
- Regular event pairs show a temporal order most of the time (80%), more than 10 times
- Prepare event pairs that are likely to have temporal relations (narrow down the search space)
- Clue: Two event phrases co-occur many times within one sentence (40278 pairs)

Temporal Relation Classifier

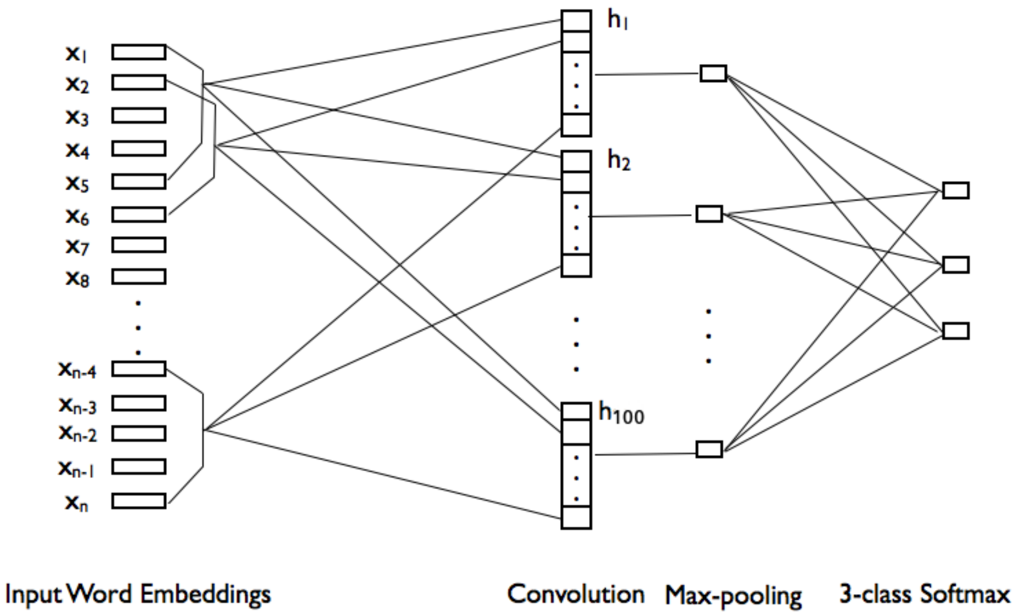
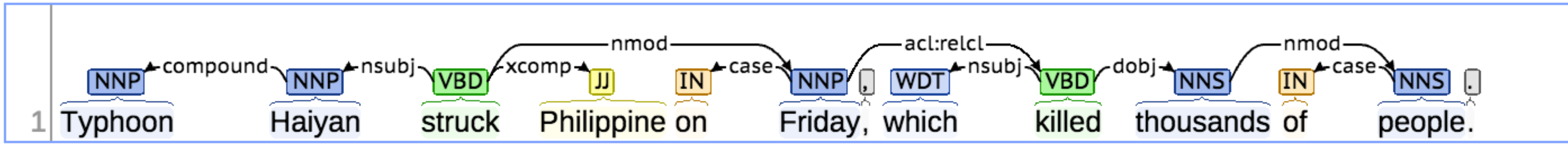


Figure: Temporal Relation Classifier

- CNN has been shown successful in sentiment analysis, sequence labeling, etc.
- Capture compositional meaning
- Pre-trained 300-dimension word2vec (Mikolov et al., 2013)
- 3 classes - after, before, other (no relation)

Sentence Contexts



1. Local window in the sentence
   - 5 words before the first event
   - All words in between
   - 5 words after the second event
2. Dependency path in the sentence
   - Consider dependency tree as an undirected graph
   - Extract the sequence of words connecting two events in the graph

Discover New Pairs

- Regular event pairs tend to show a particular temporal relation despite various contexts
- Selection criteria
  - Candidate event pairs are selected if 60% of context were labeled as *after* or *before* by CNN
  - Absolute difference between *after* and *before* labels > 40%
- Stop criteria
  - New pairs discovered is less than 100

Experiments

Common Sense	PERSON **worked** ← **graduation** **career** → **announced** retirement **wash** hands → **eating** PERSON **returned** ← **visit**
Politics	government be **formed** ← **elections** **fled** mainland ← **losing** war **imposed** sanctions ← **invasion** of LOCATION LOCATION **split** ← **war**
Business	**reached** agreement ← **negotiations** **hosted** banquet ← **meeting** **trading** → stock **closed**
Health	**cause** of death ← **cancer** PERSON be **hospitalized** ← **suffering** stroke PERSON **died** ← **admitted** to hospital
Sports	**games** → **ended** season PERSON be **sidelined** ← **undergoing** surgery PERSON be **suspended** ← **testing** for cocaine PERSON **returned** ← **recovering** from injury
Crime	**shooting** → PERSON be **arrested** **spending** in jail → PERSON be **released** PERSON be **arrested** ← **bombings** driver **fled** ← **accident**

Table: Regular event pairs extracted

Experiments

Systems	0 (Seeds)	1	2	3	4	5	Total
Basic System	1057	213	102	48	–	–	1420
+ Arg Generalization	2110	638	323	81	–	–	3152
+ Dependency Path Contexts (Full System)	2110	1230	555	288	156	62	4401

Table: Number of New Pairs Generated after Each Bootstrapping Iteration

<table>
<tr><th>Systems</th><th>Seed Pairs</th><th>New Pairs</th></tr>
<tr><td>Basic System</td><td>0.73</td><td>0.55</td></tr>
<tr><td>+ Arg Generalization</td><td rowspan="2">0.71</td><td>0.63</td></tr>
<tr><td>+ Dependency Path Contexts</td><td>0.67</td></tr>
</table>

Table: Accuracy of 100 Randomly Selected Event Pairs