Event Extraction

Xiachong Feng
Outline

1. Basic Conception
2. Dataset
3. Metric
4. Paper Counts
5. Approach And Challenge
6. Major Team
7. Future Work
1. Basic Conception
Two models of events

• TimeML model
  • An event is a word that points to a node in a network of temporal relations.
  • Every event is annotated.
  • Time is an important information, used to index events.

  It’s <EVENT class="OCCURRENCE">turning</EVENT>
  out to be another <EVENT class="STATE">bad</EVENT>
  financial week.

• ACE model
  • An event is a complex structure.
  • Only “interesting” events (events that fall into one of 34 predefined categories) are annotated.
Task Definition

• **Event Extraction (EE)** \textit{ACE05 task definition}
  - Event is represented as a structure comprising an \textit{event} trigger and a set of arguments.

• **Two core subtasks**
  - **Event Detection (ED):**
    - Identifying event triggers
    - Categorizing
  - **Argument Extraction (AE):**
    - Argument identification
    - Role classification

From “Automatically Labeled Data Generation for Large Scale Event Extraction” ACL17
“Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms” ACL17
Terminology

• **Event Trigger**
  - The main word that most clearly expresses the occurrence of an event (An ACE event trigger is typically a verb or a noun).

• **Event Attribute**
  - Type, Subtype, Modality (模态), Polairty（倾向性）, Genericity（普遍性）, Tense（时态）, 8 types and 33 subtypes. (34 = 33 + None)
Terminology

• Argument Role
  • The relationship between an argument to the event in which it participates.
  • All 35 argument roles:

<table>
<thead>
<tr>
<th>Plaintiff</th>
<th>Person</th>
<th>Place</th>
<th>Beneficiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyer</td>
<td>Seller</td>
<td>Price</td>
<td>Artifact</td>
</tr>
<tr>
<td>Origin</td>
<td>Destination</td>
<td>Giver</td>
<td>Recipient</td>
</tr>
<tr>
<td>Money</td>
<td>Org</td>
<td>Agent</td>
<td>Victim</td>
</tr>
<tr>
<td>Instrument</td>
<td>Entity</td>
<td>Target</td>
<td>Defendant</td>
</tr>
<tr>
<td>Adjudicator</td>
<td>Attacker</td>
<td>Prosecutor</td>
<td>Crime</td>
</tr>
<tr>
<td>Position</td>
<td>Sentence</td>
<td>Vehicle</td>
<td>time-after</td>
</tr>
<tr>
<td>time-before</td>
<td>time-at-end</td>
<td>time-starting</td>
<td>time-at-beginning</td>
</tr>
<tr>
<td>time-ending</td>
<td>time-holds</td>
<td>time-within</td>
<td></td>
</tr>
</tbody>
</table>

• Event Mention
  • A phrase or sentence within which an event is described, including a trigger and arguments.

From “RESEARCH ON CHINESE EVENT EXTRACTION” Hongye Tan doctoral thesis
Example

Example text from the diagram:

- Event Attribute:
  - Type: Life
  - Subtype: Be-Born
  - Person: 毛泽东
  - Time: 1893年
  - Place: 湖南湘潭

- Event Trigger

- Event Mention

- Argument role

From “REPRESENTATION LEARNING BASED INFORMATION EXTRACTION” Xiaocheng Feng doctoral thesis
2. Dataset
ACE 2005

• Contains 599 documents, which include about 6,000 labeled events.

• Annotated with single-token event triggers

• 8 event types and 33 event subtypes that, along with the “non-event” class, constitutes a 34-class classification problem.
Dataset Drawback

- Nearly 70% of event types in ACE 2005 have less than 100 labeled samples
- There are even 3 event types which have less than 10 labeled samples.

From “Event Detection via Gated Multilingual Attention Mechanism” AAAI18
3. Metric
Precision & Recall & F-score

**Precision**
\[
\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}
\]

**Recall**
\[
\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}
\]

**F-score**
\[
F_1 = \frac{2PR}{P+R}
\]
4. Paper Counts
ACL & EMNLP & AAAI & COLING & IJCAI
5. Approach And Challenge
Overview
Prior Method
Rule-based & Pattern based

• **Advantage**
  • Rules are interpretable and suitable for rapid development and domain transfer
  • Humans and machines can contribute to the same model

• **Disadvantage**
  • Not a “standard way to express rules”

• **Example**

```plaintext
1 - name: Phosphorylation_1
2  priority: 2
3  label: [Phosphorylation, Event]
4  pattern: |
5    trigger = [lemma="phosphorylation"]
6    theme:PhysicalEntity = prep_of
7      (nn|conj|cc)*
8    cause:PhysicalEntity? = prep_by
9      (nn|conj|cc)*
```

*From “A Domain-independent Rule-based Framework for Event Extraction” ACL15*
Rule & Pattern based Papers

• A Domain-independent Rule-based Framework for Event Extraction **ACL15**

• RBPB: Regularization-Based Pattern Balancing Method for Event Extraction **ACL16**
Clustering

• Open Domain: Twitter

• Challenge
  • Noisy
  • Wide Variety
  • Redundancy

• Method
  • Latent Event & Category Model (LECM): automatically grouping events into categories organized by event types.
  • Each event category is assigned with an event type label without manual intervention.
Clustering Papers

• An Unsupervised Framework of Exploring Events on Twitter: Filtering, Extraction and Categorization
  AAAI15

• Liberal Event Extraction and Event Schema Induction
  ACL16
Deep Learning
Basic Deep Learning

• Challenge
  • Same event might appear in the form of various trigger expressions
  • An expression might represent different events in different contexts

• CNN or LSTM(Multi-Class Classification Task)

From “Event Detection and Domain Adaptation with Convolutional Neural Networks” ACL15
“Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks” ACL15
New Technique

- Graph Convolutional Networks with Argument-Aware Pooling for Event Detection **AAAI18**
- Nugget Proposal Networks for Chinese Event Detection **ACL18**
- Self-regulation: Employing a Generative Adversarial Network to Improve Event Detection **ACL18**
Deep Learning Papers

• **Basic DL**
  
  • Event Detection and Domain Adaptation with Convolutional Neural Networks *ACL15*
  
  • Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks *ACL15*
  
  • A Language-Independent Neural Network for Event Detection *ACL16*
  
  • Event Nugget Detection with Forward-Backward Recurrent Neural Networks *ACL16*
  
  • Modeling Skip-Grams for Event Detection with Convolutional Neural Networks *EMNLP16*
  
  • Bidirectional RNN for Medical Event Detection in Electronic Health Records *NAACL16*

• **New Technique**
  
  • Graph Convolutional Networks with Argument-Aware Pooling for Event Detection *AAAI18*
  
  • Nugget Proposal Networks for Chinese Event Detection *ACL18*
  
  • Self-regulation: Employing a Generative Adversarial Network to Improve Event Detection *ACL18*
Joint Model
Joint Model

- **Two main approaches to EE**
  - **The joint approach** that predicts event triggers and arguments for sentences simultaneously as a structured prediction problem.
  - **The pipelined approach** that first performs trigger prediction and then identifies arguments in separate stages.

- **Joint framework**
  - Mitigating the error propagation problem of the pipelined approach.
  - Exploiting the inter-dependencies between event triggers and argument roles via discrete structures.

From “Joint Event Extraction via Recurrent Neural Networks” NAACL16
Joint Model Papers

• Joint Event Trigger Identification and Event Coreference Resolution with Structured Perceptron EMNLP15

• Event Detection and Co-reference with Minimal Supervision EMNLP16

• Joint Extraction of Events and Entities within a Document Context NAACL16

• Joint Learning for Event Coreference Resolution ACL17

• A Neural Model for Joint Event Detection and Summarization IJCAI17
External Knowledge

- External Knowledge
  - Auto Generate data
  - Cross-Lingual

- Prior Method
  - Rule-based
  - Cluster
  - Pattern-based

- Deep Learning
  - Basic DL
  - New Technique
  - ACE Dataset Drawback

- Joint
  - ED &

- Others
  - Modal Analysis
  - Full-use database
  - Transfer Learning
  - Document level
  - Joint models favor to AE task

- Cluster more use

Sentence-level

Sentence-level

Sentence-level

Sentence-level
Auto Generate Data

**Challenge**
- expensive to produce
- in low coverage of event types
- limited in size

**Method**
- World knowledge (Freebase)
- Linguistic knowledge (FrameNet)
- **Soft Distant Supervision (SDS)**

_from “Automatically Labeled Data Generation for Large Scale Event Extraction” ACL17_
Cross Lingual

• Challenge
  • Data scarcity
  • Monolingual ambiguity

• Model
  • Monolingual context attention
  • Gated cross-lingual attention

- Limited bilingual dictionaries
- Aligned multilingual word embeddings

From “Event Detection via Gated Multilingual Attention Mechanism” AAAI18

From “Leveraging Multilingual Training for Limited Resource Event Extraction” COLING16
External Knowledge Papers

• **Auto data generation**
  • Leveraging FrameNet to Improve Automatic Event Detection **ACL16**
  • Automatically Labeled Data Generation for Large Scale Event Extraction **ACL17**
  • Scale Up Event Extraction Learning via Automatic Training Data Generation **AAAI18**
  • Semi-Supervised Event Extraction with Paraphrase Clusters **NAACL18**

• **Cross-lingual**
  • Leveraging Multilingual Training for Limited Resource Event Extraction **COLING16**
  • Event Detection via Gated Multilingual Attention Mechanism **AAAI18**
Full Use Dataset

- Joint Models favor to Argument Extraction Task
  - Training corpus contains much more annotated arguments than triggers (about 9800 arguments and 5300 triggers in ACE 2005 dataset).
  - Pre-predicting potential triggers does not leverage any argument information.

From “Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms” ACL17
Document-Level

• Challenge
  • Lack of data
  • Document level data

• Method
  • Distant Supervision for generate data
  • Sequence tagging model for sentence-level events
  • Key-detection model and argument-filling strategy for document-level events

Other Papers

• Incremental Global Event Extraction **COLING16**
• Disease Event Detection based on Deep Modality Analysis **ACL15**
• Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms **ACL17**
• Zero-Shot Transfer Learning for Event Extraction **ACL18**
• DCFEE: A Document-level Chinese Financial Event Extraction System based on Automatically Labeled Training Data **ACL18**
• Document Embedding Enhanced Event Detection with Hierarchical and Supervised Attention **ACL18**
6. Major Team
Institute of Automation

• Team
  • National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

• People
  • Jun Zhao, Kang Liu, Yubo Chen......

• Papers
  • Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks ACL15
  • Leveraging FrameNet to Improve Automatic Event Detection ACL16
  • A Probabilistic Soft Logic Based Approach to Exploiting Latent and Global Information in Event Classification AAAI16
  • Automatically Labeled Data Generation for Large Scale Event Extraction ACL17
  • Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms ACL17
  • Event Detection via Gated Multilingual Attention Mechanism AAAI18
  • DCFEE: A Document-level Chinese Financial Event Extraction System based on Automatically Labeled Training Data ACL18

Kang Liu Google scholar: https://scholar.google.com/citations?user=DtZCfl0AAAAJ&hl=zh-CN&oi=sra
Institute of Automation
7. Future Work
Future Work

• Based on ACE05, do some high-level tasks, like domain specific event graph.
• Do some document-level tasks.
• Combine event graph with inference.
• *To Be Finished.*
Thank You!