Constructing Information Networks Using One Single Model

Qi Li, Heng Ji, Yu Hong, Sujian Li

Rensselaer Polytechnic Institute
Peking University
Information Extraction Hill Climbing in the Past 20 Years

- Our Previous Solutions
  - Joint Extraction of Entity Mentions and Relations (Li and Ji, ACL14)
  - Joint Extraction of Event Triggers and Event Arguments (Li et al., 2013)
What’s New in This Paper

• Interactions between entity mentions and events
• Interactions between relations and events
  o events are dynamic relations
  o relations are often initial/ending states of events

  “Asif Mohammed Hanif detonated explosives in Tel Aviv.”

  The attack event triggered by “detonated” indicates the Physical relation between "Asif Mohammed Hanif" and "Tel Aviv"

  o in ACE’05, 15% relations overlap with events
• Compared standard perceptron with K-best MIRA
• Fill in knowledge gap with FrameNet
Information Networks Construction

Traditional Approach

Task 1: Entity Mention

\[ E_1 \quad E_2 \quad ... \quad E_n \]

Task 2: Relation

\[ E_1 \xrightarrow{R_{12}} E_2 \]

Task 3: Event

\[ E_1 \xleftarrow{arg_1} T_i \xrightarrow{arg_2} E_n \]

Our Approach

- Entity

- Relation

- Event

\[ E_1 \quad E_2 \quad ... \quad E_n \]

\[ E_1 \xrightarrow{R_{12}} T_i \xrightarrow{arg_2} E_n \]

\[ \text{Traditional Approach} \quad \text{Our Approach} \]
Joint Search Algorithm

• Node-step (search for entity mentions and event triggers)
  o propose various segments at the current token
  o append to previous assignments
  o evaluate and rank new assignments

...  

| O  | ORG | PER |
|----|-----|-----|
| Asif Mohammed Hanif detonated explosives in Tel Aviv |
Joint Search Algorithm

- Node-step (search for entity mentions and event triggers)
  - propose various segments at the current token
  - append to previous assignments
  - evaluate and rank new assignments

Asif Mohammed Hanif detonated explosives in Tel Aviv
Joint Search Algorithm

• Node-step (search for entity mentions and event triggers)
  o propose various segments at the current token
  o append to previous assignments
  o evaluate and rank new assignments

\[ ... \]

Asif Mohammed Hanif detonated explosives in Tel Aviv
Joint Search Algorithm

- **Edge-step**
  - At each sub-step, connect each new mention with a previous one by a relation/event-arg link, or NIL.

Asif Mohammed Hanif detonated explosives in Tel Aviv.
Structured Perceptron with Beam Search

Update Weights:

• Standard Perceptron:
  \[ w \leftarrow w + f(x, y_{1:|z|}) - f(x, z) \]

• K-best MIRA (McDonald et al., 2005)
  \[
  \min ||w^{new} - w|| \\
  \text{s.t. } w^{new}f(x, y) - w^{new}f(x, z) \geq L(y, z) \\
  \forall z \in \text{best}_k(x, w)
  \]

Loss Function

Top \( k \) candidates
Loss Functions

- F-measure loss
  \[ L_1(y, z) = 1 - \frac{2 \cdot |y \cap z|}{|y| + |z|} \]

- 0-1 loss
  \[ L_2(y, z) = \begin{cases} 
  1 & y \neq z \\
  0 & y = z 
\end{cases} \]

- Similar to F-measure, but sensitive to the size of structures
  \[ L_3(y, z) = |y| + |z| - 2 \cdot |y \cap z| \]

Asif Mohammed Hanif detonated explosives in Tel Aviv.
Fill in Knowledge Gap with FrameNet

• Event triggers: knowledge sparsity
  o 21.5% triggers appear < twice in ACE’05
  o E.g., “Others were mutilated beyond recognition.”
    • “mutilated” is out-of-vocabulary
    • But we know “stab”, “smash” etc.

  o FrameNet (Baker and Sato, 2003) can help generalize trigger words and bridge the knowledge gap
FrameNet Features

- Use lexical units and their POS tags as features
- Many frames are relevant to ACE events
  - e.g., “cause_harm” is closely related with “injure” event

FrameNet can reduce long-tail problem of trigger words
# Experiment

- **12.2% F-score gain over state-of-the-art**

| Approach                        | Entity Mention | Relation Trigger | Event Trigger | Event Argument |
|---------------------------------|----------------|------------------|---------------|----------------|
| **Previous Work**               |                |                  |               |                |
| Pipelined Baseline              | 79.5           | 51.6             | 64.4          | 35.7           |
| Pipeline + Li et al. (2013)     | 64.5           |                  |               |                |
| Li and Ji (2014)                | 80.8           | 52.1             |               |                |
| **This Work**                   |                |                  |               |                |
| Joint w/ Avg. Perceptron        | 81.0           | 52.0             | 65.3          | 45.6           |
| Joint w/ MIRA w/ F1 Loss        | 79.0           | 49.2             | 61.5          | 47.4           |
| Joint w/ MIRA w/ 0-1 Loss       | 80.0           | 51.0             | 63.2          | **47.9**       |
| Joint w/ MIRA w/ Loss 3         | 80.7           | **52.8**         | 65.2          | 46.8           |
### Experiment

- **Top 10 features about events**

| Rank | Feature                        | Weight |
|------|--------------------------------|--------|
| 1    | Frame=Killing                  | 0.80   |
| 2    | Frame=Travel                   | 0.61   |
| 3    | Physical(Artifact, Destination)| 0.60   |
| 4    | $w_1$="home"                  | 0.59   |
| 5    | Frame=Arriving                 | 0.54   |
| 6    | ORG-AFF(Person, Entity)        | 0.48   |
| 7    | Lemma=charge                   | 0.45   |
| 8    | Lemma=birth                    | 0.44   |
| 9    | Physical(Artifact, Origin)     | 0.44   |
| 10   | Frame=Cause_harm               | 0.43   |

#### FrameNet features
- Play important role in classifying event triggers
- Provided 2.2% F measure gain in total

#### Joint Relation-Event features
- Connect cooccurrent relations and events
- Provided 0.8% F measure gain in total
Remaining Challenges

• Entity Mentions
  o Identification: “Asian Pulp and Paper Joint Stock Company, Lt. of Singapore”
  o Classification: "FAW has also utilized the capital market to directly finance,... (FAW = First Automotive Works)

• Non-verb and multi-word expression as event triggers
  o It’s important that people all over the world know that we don’t believe in the war. Nobody questions whether this is right or not.

• Knowledge scarcity is not completely solved
  o Today I was let go from my job after working there for 4 1/2 years.

• Global context and World Knowledge
  o I didn’t want to hurt him. I miss him to death.
  o I threw a stone out of the window. vs. I threw him out of the window.
  o Ellison to spend $10.3 billion to get his company.
  o We believe that the likelihood of them using those weapons goes up.
  o Fifteen people were killed and more than 30 wounded Wednesday as a suicide bomber blew himself up on a student bus in Haifa.
Two children and one adult were killed in three separate traffic accidents on Thursday and Friday of last week, according to residents and news media.

**Mention:** Two children and one adult were killed

**Event_type:** Die

**Trigger:** killed

**Arguments:**
- **Victim:** one adult
- **Victim:** Two children

**Entity Extraction**

- **Entity_type:** PER
- **Entity_mention:** Two children
- **Entity_head:** children

- **Entity_type:** PER
- **Entity_mention:** one adult
- **Entity_head:** adult

- **Entity_type:** PER
- **Entity_mention:** residents
- **Entity_head:** residents

- **Entity_type:** ORG
- **Entity_mention:** news media
- **Entity_head:** media
Thank you
Parameter Estimation

• Structured Perceptron with Beam Search

For Example:

\[ \mathcal{Z} \]

\[ y_{1:|z|} \]

\[ \text{Asif Mohammed Hanif detonated explosives in Tel Aviv} \]

\[ \text{Victim red means error} \]

\[ \text{Asif Mohammed Hanif detonated explosives in Tel Aviv} \]

\[ \text{Early update (Collins et. al., 2004, Huang et. al. 2012)} \]
Conclusions

• We formulate the core tasks of information extraction: entity mention, relation, and event extractions as Constructing Information Networks.

• This new formulation can be resolved by a single joint model, allowing full exploration of inter-task and long-distance dependencies.

• We demonstrate that FrameNet can help reduce sparsity problem in trigger words.

• Future work includes applying this framework to other similar Information Extraction tasks.
Construct Information Networks

Traditional Approach

Our Approach

A better way of enjoying fruits
Problem Statement

Joint Extraction of Event Triggers and Arguments
Li et. al., ACL 2013

Joint Extraction of Entity Mentions and Relations
Li et. al., ACL 2013

Jointly Construct Information Networks

Joint search algorithm (Beam-Search)

\[
\arg\max \ f(x, y') \cdot w \\
y' \in \mathcal{Y}(x)
\]

Kiichiro Toyoda founded the automaker
Jointly Construct Information Networks

Joint Extraction of Event Triggers and Arguments
Li et. al., ACL 2013

Joint Extraction of Entity Mentions and Relations
Li et. al., ACL 2013

Joint search algorithm (Beam-Search)

$\arg \max f(x, y') \cdot w$

$y' \in \mathcal{Y}(x)$

Kiichiro Toyoda founded the automaker

Kiichiro Toyoda $\textit{founded}$ the automaker

$\text{ORG-AFF}$

Agent

Start-Org

Org

PER

ORG
..., dozens of Israeli tanks advanced into the northern Gaza Strip backed by helicopters which fired at least three rockets in the Jabaliya area, Palestinian security sources said. 

AFP 2003/03/05
Joint Search Algorithm

• Return the candidate with the highest model score as the final prediction

Asif Mohammed Hanif detonated explosives in Tel Aviv
Parameter Estimation

• K-best MIRA

\[
\begin{align*}
\min & \ |w^{new} - w| \\
\text{s.t.} & \ w^{new}f(x, y) - w^{new}f(x, z) \geq L(y, z) \\
\forall z & \in best_k(x, w)
\end{align*}
\]

Top k candidates

Use (arbitrary) loss function \( L(y, z) \) and information from top \( k \) candidate structures to guide the update.