Annotation Scheme for Social Network Extraction

Apoorv Agarwal
Owen Rambow
Rebecca J. Passoneau
Columbia University

August 4, 2010
Extract Social Networks out of texts

Figure: System that extracts Social Networks out of texts
Social Networks?

- Nodes: People or group of people
Social Networks?

- **Nodes**: People or group of people
- **Edges**: Social events that we define in this annotation effort
Social Networks?

- Nodes: People or group of people
- Edges: Social events that we define in this annotation effort

This annotation effort is about annotating Social Events
Social Networks?

- Nodes: People or group of people
- Edges: Social events that we define in this annotation effort

This annotation effort is about annotating **Social Events**

- Interaction (INR): talking, meeting
Social Networks?

- Nodes: People or group of people
- Edges: Social events that we define in this annotation effort

This annotation effort is about annotating **Social Events**

- **Interaction (INR):** talking, meeting
- **Physical Proximity (PPR):** seeing, suicide-bombing
Social Networks?

- **Nodes**: People or group of people
- **Edges**: Social events that we define in this annotation effort

This annotation effort is about annotating **Social Events**

- **Interaction (INR)**: talking, meeting
- **Physical Proximity (PPR)**: seeing, suicide-bombing
- **Perception (PCR)**: seeing on TV, magazine
Social Networks?

- Nodes: People or group of people
- Edges: Social events that we define in this annotation effort

This annotation effort is about annotating **Social Events**

- **Interaction (INR)**: talking, meeting
- **Physical Proximity (PPR)**: seeing, suicide-bombing
- **Perception (PCR)**: seeing on TV, magazine
- **Cognition (COG)**: thinking, loving, missing
Example

Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder on Saturday still urging for a central role for the UN in a post-war revival of Iraq.
Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder on Saturday still urging for a central role for the UN in a post-war revival of Iraq.
Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder on Saturday still urging for a central role for the UN in a post-war revival of Iraq

- We do not consider transitive or inferred events (No link between Chirac and Schroeder)
Motivation
Motivation

Other Annotations - ACE
Motivation
Other Annotations - ACE
Details of Social Events
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
- Conclusion and Future Work
How are characters clustered according to their structural properties?

Figure: Block-modelling: Social positions and roles
Motivation

Other Annotations - ACE

Details of Social Events

Comparison with ACE Events

Inter-annotator Agreement

Conclusion and Future Work
Automatic Content Extraction (ACE)

Annotate newswire, broadcast conversation, and weblogs with Entities, Relations and Events.
Automatic Content Extraction (ACE)

Annotate newswire, broadcast conversation, and weblogs with Entities, Relations and Events.

Entity: An Entity is an object or set of objects in the world.
Automatic Content Extraction (ACE)

Annotate newswire, broadcast conversation, and weblogs with Entities, Relations and Events.

Entity: An Entity is an object or set of objects in the world.

Types of Entities:

- **Person**: Individual human being or a group of human beings
- **Organization**: Agencies, Corp. (Organizational structures)
- **Geo-political**: Geo. regions defined by political/social groups
- **Location**: Geographical land-masses, bodies of water etc.
- **Facility**: Buildings and other man-made structures
Automatic Content Extraction (ACE)

Annotate newswire, broadcast conversation, and weblogs with Entities, Relations and Events.

Entity: An Entity is an object or set of objects in the world.

Types of Entities:

- **Person**: Individual human being or a group of human beings
- **Organization**: Agencies, Corp. (Organizational structures)
- **Geo-political**: Geo. regions defined by political/social groups
- **Location**: Geographical land-masses, bodies of water etc.
- **Facility**: Buildings and other man-made structures
Automatic Content Extraction (ACE)

Annotate newswire, broadcast conversation, and weblogs with Entities, Relations and Events.

Entity: An Entity is an object or set of objects in the world.

Types of Entities:

- **Person:** Individual human being or a group of human beings
- **Organization:** Agencies, Corp. (Organizational structures)
- **Geo-political:** Geo. regions defined by political/social groups
- **Location:** Geographical land-masses, bodies of water etc.
- **Facility:** Buildings and other man-made structures

We are not concerned with ACE Relations

We will see shortly what are the ACE Events and how our annotations are different
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
- Conclusion and Future Work
**Definition:** Two entities have an interaction relation if both entities are mutually aware of their *presence* and the *event* taking place between the entities.
**Definition:** Two entities have an interaction relation if both entities are mutually aware of their *presence* and the *event* taking place between the entities.
**Definition:** Two entities have an interaction relation if both entities are mutually aware of their *presence* and the *event* taking place between the entities.
Examples of INR

Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder.

INR. Verbal-Near
Examples of INR

Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder. Putin had secretly invited Tony Blair. . .
Examples of INR

Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder.

Putin had secretly invited Tony Blair . . .

Within weeks he was arrested and charged with sodomizing an official driver . . .
Examples of INR

Putin concluded his two days of talks in St. Petersburg with Jacques Chirac of France and Chancellor Gerhard Schroeder.

Putin had secretly invited Tony Blair.

Within weeks he was arrested and charged with sodomizing an official driver.

The writer will retain the rights to his books and films, although he has agreed to split a raft of possessions with Anne Marie.
Physical Proximity Relation (PPR)

**Definition:** Entities are said to have a physical proximity relation if:

1. only one entity is aware of the other entity AND
2. one entity can see/observe the other live
**Physical Proximity Relation (PPR)**

**Definition:** Entities are said to have a physical proximity relation if:

1. only one entity is aware of the other entity AND
2. one entity can see/observe the other live
Physical Proximity Relation (PPR)

**Definition:** Entities are said to have a physical proximity relation if:

1. only one entity is aware of the other entity AND
2. one entity can see/observe the other live

Directed relationship

A. Agarwal
Examples of PPR

Alice saw the Rabbit running by her

PPR
Examples of PPR

Alice saw the Rabbit running by her

W. Rabbit ran by Alice

PPR

No Annotation
Examples of PPR

Alice saw the Rabbit running by her

W. Rabbit ran by Alice

Fifteen people were killed and more than 30 wounded as a suicide bomber blew himself up ...
Definition: Entities are said to have a perception relation if:

1. only one entity is aware of the other entity AND
2. one entity can observe the other through a broadcast medium
**Definition:** Entities are said to have a perception relation if:

1. only one entity is aware of the other entity AND
2. one entity can observe the other through a broadcast medium
**Perception Relation (PCR)**

**Definition**: Entities are said to have a perception relation if:

1. only one entity is aware of the other entity AND
2. one entity can observe the other through a broadcast medium

![Diagram of entities A and B with a directed relationship via TV Magazine medium]
**Definition:** Entities are said to have a perception relation if:

1. only one entity is aware of the other entity AND
2. one entity can observe the other through a broadcast medium
The pair flew to Singapore last year after hearing of the successful surgery on Nepalese babies Ganga and Jamuna Shrestha... PCR
**Definition**: Entities are said to have a cognition relation if:

1. an entity is aware of the other entity AND
2. relation is NOT PPR or PCR
Definition: Entities are said to have a cognition relation if:

1. an entity is aware of the other entity AND
2. relation is NOT PPR or PCR
Examples of COG

The police say the bomber blew himself up outside this cafe
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
- Conclusion and Future Work
Comparison with ACE Events

A. INR
B. Conflict

A. PPR
B. Contact

A. PCR
B. Justice

A. COG
B. Life

A. A ⊂ B (No need for our annotations)
B. Type \( i \) ⊂ B (No need for part of our annotations)

In fact, neither of these scenarios happens \( \Rightarrow \) our annotations are different from ACE.
Comparison with ACE Events

\[ A \subset B \] (No need for our annotations)
Comparison with ACE Events

- $A \subset B$ (No need for our annotations)
- $A_{Type_i} \subset B$ (No need for part of our annotations)
Comparison with ACE Events

- $A \subset B$ (No need for our annotations)
- $A_{Type_i} \subset B$ (No need for part of our annotations)

In fact, neither of these scenarios happens $\implies$ our annotations are different from ACE
Comparison with ACE Events

• However, there is some overlap; mainly INR ↔ Contact
Comparison with ACE Events

- However, there is some overlap; mainly INR ↔ Contact
- Specifically, ACE captures “meeting” events that are INR for us
Comparison with ACE Events

- However, there is some overlap; mainly INR ↔ Contact
- Specifically, ACE captures “meeting” events that are INR for us
- But our definition of INR is broader:
  
  *Toujan Faisal, 54, said she was informed of the refusal by an Interior Ministry committee overseeing election preparations.*

INR.Verbal-Far
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
- Conclusion and Future Work
Inter-annotator Agreement

- Two annotators, 62 documents, ACE-2005 corpus
Inter-annotator Agreement

- Two annotators, 62 documents, ACE-2005 corpus

Figure: Decision points an annotator goes through for selecting type and sub-type of social events.

A. Agarwal
Inter-annotator Agreement

- We report Confusion matrices, Kappa, F-measure for each decision point
We report Confusion matrices, Kappa, F-measure for each decision point

<table>
<thead>
<tr>
<th></th>
<th>$Yes_{Ann_1}$</th>
<th>$No_{Ann_1}$</th>
<th>$Yes_{Ann_2}$</th>
<th>$No_{Ann_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Yes_{Ann_1}$</td>
<td>$tp$</td>
<td>$fp$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$No_{Ann_1}$</td>
<td>$fn$</td>
<td>$tn$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Confusion Matrix
Inter-annotator Agreement

- We report Confusion matrices, Kappa, F-measure for each decision point

<table>
<thead>
<tr>
<th></th>
<th>Yes$_{Ann_1}$</th>
<th>No$_{Ann_1}$</th>
<th>Yes$_{Ann_2}$</th>
<th>No$_{Ann_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes$_{Ann_1}$</td>
<td>$tp$</td>
<td>$fp$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No$_{Ann_1}$</td>
<td>$fn$</td>
<td>$tn$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Confusion Matrix

- $Kappa = \frac{P(a)-P(e)}{1-P(e)}$
We report Confusion matrices, Kappa, F-measure for each decision point

<table>
<thead>
<tr>
<th></th>
<th>$Yes_{Ann_1}$</th>
<th>$No_{Ann_1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Yes_{Ann_2}$</td>
<td>$tp$</td>
<td>$fp$</td>
</tr>
<tr>
<td>$No_{Ann_2}$</td>
<td>$fn$</td>
<td>$tn$</td>
</tr>
</tbody>
</table>

Table: Confusion Matrix

- **Kappa** = \( \frac{P(a)-P(e)}{1-P(e)} \)
- **F - measure** = \( \frac{2PR}{P+R} \); \( P = \frac{tp}{tp+fp} \); \( R = \frac{tp}{tp+fn} \)
## Inter-annotator Agreement

<table>
<thead>
<tr>
<th>Decision Point</th>
<th>Confusion Matrix</th>
<th>Kappa</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tp</td>
<td>fp</td>
<td>fn</td>
</tr>
<tr>
<td>1 (+/- Relation)</td>
<td>133</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>2 (1 or 2 way)</td>
<td>51</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>3 (Verbal/NonV)</td>
<td>40</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4 (NonV-Near/Far)</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5 (Verbal-Near/Far)</td>
<td>30</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6 (+/- PPR)</td>
<td>71</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7 (PCR/COG)</td>
<td>69</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Outline

- Motivation
- Other Annotations - ACE
- Details of Social Events
- Comparison with ACE Events
- Inter-annotator Agreement
- Conclusion and Future Work
Conclusion

- Presented an annotation scheme for social network extraction from text where social events are based on cognitive states of people involved
- Showed our annotations are novel (compared with ACE)
- Reported IAA measures that show our annotations are reliable
Future Work

- Extend manual for entity annotation: *Mouse was talking to Dodo, Duck etc. He said to them...*
- Extend manual for social event annotation: *The Mouse said to Alice, “Oh! Mary Ann”*
- Annotate sentiment on already annotated social events
Acknowledgements

- David Elson for helping us with the annotation manual and reviewing the paper
- Natural Language Processing Group at Columbia University for useful comments on our annotation scheme

Thanks!
It said that the EU was seriously concerned about Iran's nuclear activities and urged it to cooperate fully with the IAEA.

"This would be a significant step in demonstrating Iran's stated peaceful intentions with regard to its nuclear program," the seven-point statement said.

George Papandreou, foreign minister of Greece, the current holder of the EU presidency, said he was in personal contact with IAEA Director-General Mohamed ElBaradei and would follow his lead on this matter.

EU foreign ministers reiterated that improved trade links with Iran
### Backup Slides

<table>
<thead>
<tr>
<th>62 Documents</th>
<th>Conflict (5)</th>
<th>Contact (32)</th>
<th>Justice-* (13)</th>
<th>Life (7)</th>
<th>Transaction (2)</th>
<th>Not Found</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Attack</td>
<td>Meet</td>
<td>Phone-Write</td>
<td>Die</td>
<td>Divorce</td>
<td>Injure</td>
</tr>
<tr>
<td>INR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Near (66)</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Far (17)</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NonVerbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Near (14)</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Far (3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG (109)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PPR (2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PCR (1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Errors</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

A. Agarwal
### Email Content

<table>
<thead>
<tr>
<th>Sender → Receiver</th>
<th>Email Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kate → Sam</strong></td>
<td>[Jacob], the City attorney had a couple of questions which [I] will {attempt to relay} without having a copy of the documents.</td>
</tr>
<tr>
<td><strong>Sam → Kate, Mary</strong></td>
<td>Can you obtain the name of Glendale’s bond counsel (lawyer’s name, phone number, email, etc.)?</td>
</tr>
<tr>
<td><strong>Kate → Sam</strong></td>
<td>Glendale’s City Attorney is Jacob. Please let [me] {know} if [you] need anything else.</td>
</tr>
<tr>
<td><strong>Mary → Sam</strong></td>
<td>I do not see a copy of an opinion in the file nor have we received one since [I] {sent} the execution copies of the ISDA to [Jacob].</td>
</tr>
<tr>
<td><strong>Kate → Jacob</strong></td>
<td>Jacob, could you provide the name, phone number, etc. of your bond council for our attorney, Sam?</td>
</tr>
<tr>
<td><strong>Kate → Sam</strong></td>
<td>[I] will {work on this for} [you] - and will be in touch.</td>
</tr>
</tbody>
</table>