Chinese Event Coreference Resolution: Understanding the State of the Art

Chen Chen and Vincent Ng

Human Language Technology Research Institute
The University of Texas at Dallas
Event Coreference

• Goal: Determine which event mentions in a text refer to the same real-world event
Event Coreference

• Goal: Determine which event mentions in a text refer to the same real-world event

Since there is few work on event coreference, our understanding of this task is fairly limited!
Goal

• Understand how a state-of-the-art end-to-end event coreference resolver can be improved
Goal

• Understand how a state-of-the-art end-to-end event coreference resolver can be improved
  – Question 1: To what extent is the noise inherent in the output of each of its upstream components limiting the performance?
Goal

• Understand how a state-of-the-art end-to-end event coreference resolver can be improved
  – Question 1: To what extent is the noise inherent in the output of each of its upstream components limiting the performance?
  – Question 2: What are the major types of errors that are attributable to the resolution algorithm?
Why Chinese Event Coreference Resolution

• Lack of publicly available results on Chinese event coreference resolution
  – Most of recent work on event coreference are for English
    • Humphreys et al.(1997), Chen et al.(2009), Bejan and Harabagiu(2010), Chen et al.(2011), Lee et al.(2012)
Plan for the Talk

• ACE Event Coreference
• Six Upstream Components in the Pipeline
• Results and Analysis – Answer to Question 1
• Conclusion
Plan for the Talk

• ACE Event Coreference
• Six Upstream Components in the Pipeline
• Results and Analysis – Answer to Question 1
• Conclusion
(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.
(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.

- This example contains three event mentions, each of which has a type and subtype and is associated with arguments
(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.

- This example contains three event mentions, each of which has a type and subtype and is associated with arguments
  - [injured]: Type LIFE; SubType INJURE; arguments: (Zhang Jiarong), (the road) and (yesterday evening)
(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.

- This example contains three event mentions, each of which has a type and subtype and is associated with arguments
  - [injured]: Type LIFE; SubType INJURE; arguments: (Zhang Jiarong), (the road) and (yesterday evening)
  - [stabbed]: Type CONFLICT; SubType: ATTACK; arguments: (two men), (him) and (a knife)
ACE Event Coreference

(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.

- This example contains three event mentions, each of which has a type and subtype and is associated with arguments
  - [injured]: Type LIFE; SubType INJURE; arguments: (Zhang Jiarong), (the road) and (yesterday evening)
  - [stabbed]: Type CONFLICT; SubType: ATTACK; arguments: (two men), (him) and (a knife)
  - [criminal]: Type CONFLICT; SubType: ATTACK; arguments: (The thugs) and (Zhang Jiarong)
ACE Event Coreference

(Zhang Jiarong) was cycling on (the road) (yesterday evening) and was [injured] when (two men) [stabbed] (him) with (a knife). (The thugs)’ [criminal] motivation may have something to do with (Zhang Jiarong)’s testimony in a criminal case.

- This example contains three event mentions, each of which has a type and subtype and is associated with arguments
- [stabbed] and [criminal] are coreferent because they refer to the same real-world event
Plan for the Talk

• ACE Event Coreference
• Six Upstream Components in the Pipeline
• Results and Analysis – Answer to Question 1
• Conclusion
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary

Event Mention Attributes

Diagram shows the relationships between these components, illustrating how they interconnect to support event coreference resolution.
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary
Six Upstream Components for Event Coreference Resolution

- Event Coreference
  - Event Mention Attributes
  - Event Mention Boundary & Subtyping
  - Event Argument & Role
  - Entity Coreference
    - Entity Typing & SubTyping
    - Entity Mention Boundary
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary

Event Mention Attributes
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary &Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing &SubTyping
- Entity Mention Boundary

Event Mention Attributes
Six Upstream Components for Event Coreference Resolution

Event Coreference

- Event Mention Attributes
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary
Six Upstream Components for Event Coreference Resolution

- Event Coreference
  - Event Mention Boundary & Subtyping
  - Event Argument & Role
  - Entity Coreference
  - Entity Typing & SubTyping
  - Entity Mention Boundary
  - Event Mention Attributes
Event Mention
Boundary Identification & SubTyping

• Goals
  – Provide the event mentions for event coreference
Event Mention
Boundary Identification & SubTyping

• Goals
  – Provide the event mentions for event coreference
  – Label each event mention with its event subtype
Event Mention
Boundary Identification & SubTyping

• Goals
  – Provide the event mentions for event coreference
  – Label each event mention with its event subtype,

• Why is this component useful for event coreference?
Event Mention
Boundary Identification & SubTyping

• Goals
  – Provide the event mentions for event coreference
  – Label each event mention with its event subtype,

• Why is this component useful for event coreference?
  – Two event mentions with different type or subtype cannot be coreferent (useful feature for event coreference)
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary
- Event Mention Attributes
Event Mention
Attribute Value Computation

• Goal
  – Assign each event mention with 4 attributes:
    POLARITY, MODALITY, GENERICITY and TENSE
Event Mention
Attribute Value Computation

• Goal
  – Assign each event mention with 4 attributes: POLARITY, MODALITY, GENERICITY and TENSE

• Why is this component useful for event coreference?
Event Mention
Attribute Value Computation

• Goal
  – Assign each event mention with 4 attributes: POLARITY, MODALITY, GENERICITY and TENSE

• Why is this component useful for event coreference?
  – Two events differ in any of four attributes cannot be coreferent (useful feature for event coreference)
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary

Event Mention Attributes
Event Argument & Role Classification

• Goals
  – Identify arguments for an event mention (e.g., the participants, time, place)
Event Argument & Role Classification

• Goals
  – Identify arguments for an event mention (e.g., the participants, time, place)
  – Assign a role (e.g. VICTIM, PLACE, TIME-WITHIN) to each argument
Event Argument & Role Classification

• Goals
  – Identify arguments for an event mention (e.g., the participants, time, place)
  – Assign a role (e.g. VICTIM, PLACE, TIME-WITHIN) to each argument

• Why is this component useful for event coreference?
Event Argument & Role Classification

• Goals
  – Identify arguments for an event mention (e.g., the participants, time, place)
  – Assign a role (e.g. VICTIM, PLACE, TIME-WITHIN) to each argument

• Why is this component useful for event coreference?
  – Useful features for event coreference
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary

- Event Mention Attributes
Entity Coreference Resolution

• Goal
  – Create entity coreference clusters
Entity Coreference Resolution

• Goal
  – Create entity coreference clusters

• Why is this component useful for event coreference?
Entity Coreference Resolution

- **Goal**
  - Create entity coreference clusters

- **Why is this component useful for event coreference?**
  - Two event mentions having coreferent arguments are likely to be coreferent (useful feature for event coreference)
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Entity Typing & SubTyping
- Entity Mention Boundary
Entity Mention
Boundary Identification

• Goal
  – Provide candidate arguments and entity mentions needed by the aforementioned components
Entity Mention
Boundary Identification

• Goal
  – Provide candidate arguments and entity mentions needed by the aforementioned components
    • Indirect influence on event coreference
Six Upstream Components for Event Coreference Resolution

- Event Coreference
- Event Mention Boundary & Subtyping
- Event Argument & Role
- Entity Coreference
- Event Typing & SubTyping
- Entity Mention Boundary
Entity Typing & SubTyping

• Goal
  – Determine the type and subtype of entity mention
Entity Typing & SubTyping

• Goal
  – Determine the type and subtype of entity mention
    (Indirect influence on event coreference)
Entity Typing & SubTyping

• Goal
  – Determine the type and subtype of entity mention
    (Indirect influence on event coreference)

• Why is this component useful for event coreference?
Entity Typing & SubTyping

• Goal
  – Determine the type and subtype of entity mention (Indirect influence on event coreference)

• Why is this component useful for event coreference?
  – Features for classifying the role of event arguments
Plan for the Talk

- ACE Event Coreference
- Six Upstream Components in the Pipeline
- Results and Analysis – Answer to Question 1
- Conclusion
Evaluation Methodology

• Start with an event coreference resolver that assumes all six upstream components are error free
• Replace each oracle component with its system (i.e., automatically computed) counterpart one by one
Replacement 1:
Using System Event Mention Attribute Values for Train & Test

```
Event Coreference

Event Mention Boundary &Subtyping

Event Argument & Role

Entity Coreference

Entity Typing &SubTyping

Entity Mention Boundary
```

Event Mention Attributes

---

---
Replacement 1: Using System Event Mention Attribute Values for Train & Test

• How to implement this component?
  – Following Chen et al. (2009), we train 4 classifiers to compute these attributes, with one classifier per attribute
Replacement 1:
Using System Event Mention Attribute Values for Train & Test

• How to implement this component?
  – Following Chen et al. (2009), we train 4 classifiers to compute these attributes, with one classifier per attribute

• Each of four attribute classifiers is only marginally better than a simple majority baseline
Replacement 1:
Using System Event Mention Attribute Values for Train & Test

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th>B³</th>
<th></th>
<th>CEAF&lt;sub&gt;e&lt;/sub&gt;</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>80.4</td>
<td>70.0</td>
<td>74.8</td>
<td>88.4</td>
<td>79.7</td>
<td>57.3</td>
</tr>
<tr>
<td>After</td>
<td>72.5</td>
<td>64.5</td>
<td>68.3</td>
<td>83.8</td>
<td>77.4</td>
<td>80.5</td>
</tr>
</tbody>
</table>

- The average F decreases by 5.1%
Replacement 1:
Using System Event Mention Attribute Values for Train & Test

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>80.4</td>
<td>70.0</td>
<td>74.8</td>
<td>88.4</td>
<td>79.7</td>
<td>83.8</td>
<td>1.7</td>
</tr>
<tr>
<td>After</td>
<td>72.5</td>
<td>64.5</td>
<td>68.3</td>
<td>83.8</td>
<td>77.4</td>
<td>80.5</td>
<td>56.3</td>
</tr>
</tbody>
</table>

• Conclusion 1:
  – Improving the four event attribute classifiers could significantly improve event coreference
Replacement 2:
Using System Event Argument & Role

Event Coreference

Event Mention Boundary & Subtyping

Event Argument & Role

Entity Coreference

Entity Typing & SubTyping

Entity Mention Boundary
Replacement 2: Using System Event Argument & Role

• How to implement this component?
  – Implemented as part of our Chinese event extraction system (Chen and Ng, 2012c)
Replacement 2: Using System Event Argument & Role

- How to implement this component?
  - Implemented as part of our Chinese event extraction system (Chen and Ng, 2012c)

- Given gold event mention boundary and subtyping, the F-score of event argument and role classification are 76.9% and 68.2%
Replacement 2: Using System Event Argument & Role

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th>B³</th>
<th></th>
<th>CEAFᵋ</th>
<th></th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>72.5</td>
<td>64.5</td>
<td>68.3</td>
<td>83.8</td>
<td>77.4</td>
<td>80.5</td>
<td>53.1</td>
</tr>
<tr>
<td>After</td>
<td>71.2</td>
<td>61.2</td>
<td>65.8</td>
<td>83.9</td>
<td>74.9</td>
<td>79.1</td>
<td>49.9</td>
</tr>
</tbody>
</table>

• After replacing gold with system event argument and role, average F-score drops slightly, though significantly, by 2.1%
Replacement 2: Using System Event Argument & Role

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
<th>CEAF e</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>Before</td>
<td>72.5</td>
<td>64.5</td>
<td>68.3</td>
<td>83.8</td>
</tr>
<tr>
<td>After</td>
<td>71.2</td>
<td>61.2</td>
<td>65.8</td>
<td>83.9</td>
</tr>
</tbody>
</table>

• Conclusion 2:
  – Event argument and role classification have a small, but significant, impact on event coreference performance
Replacement 3: Using System Entity Coreference

- Event Coreference
  - Event Mention Attributes
  - Event Mention Boundary & Subtyping
  - Event Arguments & Role
  - Entity Typing & SubTyping
  - Entity Mention Boundary
  - Entity Coreference
Replacement 3: Using System Entity Coreference

• How to implement this component?
  – Provided by our Chinese entity coreference resolver (Chen and Ng, 2012b)
Replacement 3: Using System Entity Coreference

• How to implement this component?
  – Provided by our Chinese entity coreference resolver (Chen and Ng, 2012b)

• Given gold entity mentions, our system entity coreference resolver achieves a MUC F-score of 78.0%
Replacement 3: Using System Entity Coreference

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
<th>CEAF&lt;sub&gt;e&lt;/sub&gt;</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>Before</td>
<td>71.2</td>
<td>61.2</td>
<td>65.8</td>
<td>83.9</td>
</tr>
<tr>
<td>After</td>
<td>61.6</td>
<td>58.5</td>
<td>60.0</td>
<td>79.0</td>
</tr>
</tbody>
</table>

- Replacing gold with system entity coreference incurs a 3.7% drop
Replacement 3: Using System Entity Coreference

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>71.2</td>
<td>61.2</td>
<td>65.8</td>
<td>83.9</td>
<td>74.9</td>
<td>79.1</td>
</tr>
<tr>
<td>After</td>
<td>61.6</td>
<td>58.5</td>
<td>60.0</td>
<td>79.0</td>
<td>75.7</td>
<td>77.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>CEAF&lt;sub&gt;e&lt;/sub&gt;</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>P</td>
</tr>
<tr>
<td>Before</td>
<td>49.9</td>
<td>58.0</td>
<td>53.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After</td>
<td>49.1</td>
<td>51.5</td>
<td>50.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>53.6</td>
</tr>
<tr>
<td>After</td>
<td>62.5</td>
</tr>
</tbody>
</table>

- Conclusion 3: Improve entity coreference could significantly improve event coreference
Replacement 4: Using System Entity Typing & SubTyping
Replacement 4: Using System Entity Typing & SubTyping

• How to implement this component?
  – We determine entity types and subtypes by training two SVM classifiers
Replacement 4: Using System Entity Typing & SubTyping

• How to implement this component?
  – We determine entity types and subtypes by training two SVM classifiers

• Given gold entity mentions, system entity type and subtype classifiers achieve F-scores of 90.1% and 81.6%
Replacement 4: Using System Entity Typing & SubTyping

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AvgF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td></td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>61.6</td>
<td>58.5</td>
<td>60.0</td>
<td>79.0</td>
<td>75.7</td>
<td>77.3</td>
<td>49.1</td>
<td>51.5</td>
<td>50.3</td>
</tr>
<tr>
<td>After</td>
<td>62.2</td>
<td>57.9</td>
<td>60.0</td>
<td>79.4</td>
<td>75.2</td>
<td>77.2</td>
<td>49.0</td>
<td>52.3</td>
<td>50.6</td>
</tr>
</tbody>
</table>

- After replacing gold with system entity types and subtypes, event coreference performance does not drop
Replacement 4: Using System Entity Typing & SubTyping

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th></th>
<th>CEAFₑ</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
</tr>
<tr>
<td>Before</td>
<td>61.6</td>
<td>58.5</td>
<td>60.0</td>
<td>79.0</td>
<td>75.7</td>
<td>77.3</td>
<td>49.1</td>
<td>51.5</td>
</tr>
<tr>
<td>After</td>
<td>62.2</td>
<td>57.9</td>
<td>60.0</td>
<td>79.4</td>
<td>75.2</td>
<td>77.2</td>
<td>49.0</td>
<td>52.3</td>
</tr>
</tbody>
</table>

• Conclusion 4:
  – Improving entity typing & subTyping classification is unlikely to improve event coreference
Replacement 5:
Using System Entity Mention Boundary detection

Event Coreference

Event Mention Boundary & Subtyping

Event Argument & Role

Entity Coreference

Entity Typing & SubTyping

Entity Mention Boundary
Replacement 5:
Using System Entity Mention Boundary detection

• How to implement this component?
  – We train CRF classifiers to extract entity mentions, time expressions and value expressions
Replacement 5: Using System Entity Mention Boundary detection

• How to implement this component?
  – We train CRF classifiers to extract entity mentions, time expressions and value expressions

• System entity mention boundary detection component achieves an F-score of 84.7%
Replacement 5:
Using System Entity Mention Boundary detection

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th>CEAF&lt;sub&gt;e&lt;/sub&gt;</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>62.2</td>
<td>57.9</td>
<td>60.0</td>
<td>79.4</td>
<td>75.2</td>
<td>77.2</td>
<td>49.0</td>
</tr>
<tr>
<td>After</td>
<td>63.3</td>
<td>57.4</td>
<td>60.2</td>
<td>80.2</td>
<td>74.4</td>
<td>77.2</td>
<td>48.2</td>
</tr>
</tbody>
</table>

- Replacing gold with system entity mention boundary detection does not alter event coreference performance
Replacement 5: Using System Entity Mention Boundary detection

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B$^3$</th>
<th>CEAF$_e$</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>Before</td>
<td>62.2</td>
<td>57.9</td>
<td>60.0</td>
<td>79.4</td>
</tr>
<tr>
<td>After</td>
<td>63.3</td>
<td>57.4</td>
<td>60.2</td>
<td>80.2</td>
</tr>
</tbody>
</table>

• Conclusion 5:
  – Improving entity mention boundary detection may not improve event coreference
Replacement 6: Using System Event Mention Boundary Identification & subtyping
Replacement 6: Using System Event Mention Boundary Identification & subtyping

• How to implement this component?
  – Implemented as part of our Chinese event extraction system (Chen and Ng, 2012c)
Replacement 6: Using System Event Mention Boundary Identification & subtyping

• How to implement this component?
  – Implemented as part of our Chinese event extraction system (Chen and Ng, 2012c)

• System event mention boundary identifier achieves an F-score of 65.1%

• System event subtype classifier achieves an F-score of 61.3%
Replacement 6: Using System Event Mention Boundary Identification & subtyping

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th>B³</th>
<th></th>
<th>CEAFₑ</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
<td>P</td>
<td>F</td>
</tr>
<tr>
<td>Before</td>
<td>63.3</td>
<td>57.4</td>
<td>60.2</td>
<td>80.2</td>
<td>74.4</td>
<td>77.2</td>
</tr>
<tr>
<td>After</td>
<td>37.4</td>
<td>36.7</td>
<td>37.1</td>
<td>72.8</td>
<td>71.1</td>
<td>71.9</td>
</tr>
</tbody>
</table>

- Replacing gold with system event mention boundary and subtyping causes average F-score to drop by 12.7%
Replacement 6: Using System Event Mention Boundary Identification & subtyping

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
<th>CEAF&lt;sub&gt;e&lt;/sub&gt;</th>
<th>AvgF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>Before</td>
<td>63.3</td>
<td>57.4</td>
<td>60.2</td>
<td>80.2</td>
</tr>
<tr>
<td>After</td>
<td>37.4</td>
<td>36.7</td>
<td>37.1</td>
<td>72.8</td>
</tr>
</tbody>
</table>

**Conclusion 6:**

- Event mention boundary identification and subtyping is the upstream component that has the largest impact.
Answer to Question 1

• Components whose noise have an impact on event coreference performance (in decreasing order of impact):
  – Event Mention Boundary Identification & SubTyping
  – Event Mention Attribute Value Computation
  – Entity Coreference Resolution
  – Event Argument & Role Classification

• Components whose noise do not have an impact:
  – Entity Mention Boundary Identification
  – Entity Typing & SubTyping
Plan for the Talk

• ACE Event Coreference
• Six Upstream Components
• Results and Analysis – Answer to Question 1
• Error Analysis – Answer to Question 2
• Conclusion
Precision Errors

• Lack of event timestamping
  – Only events occurring exactly at the same time can be coreferent
  – TENSE is just a very rough approximation
Lack of event timestamping

E1: In last March, Yang Guangnan was [arrested] in Shanghai for the first time.

E2: Yang Guangnan was [arrested] again in Shanghai.
Precision Errors

• Incompatible triggers
  – Two events containing coreferent arguments, but triggers are semantically incompatible
E1: On the 28\textsuperscript{th}, Sam Nujoma arrived in Pyongyang by plane for an official goodwill [visit] to the DPRK.

E2: Namibian President Sam Nujoma [arrived] in Pyongyang by plane on the 28\textsuperscript{th}.
Precision Errors

• Incompatible important arguments
  – Two events containing strong hint to be coreferent, while some of their important arguments are incompatible
Incompatible important arguments


E2: During their [visit] in Denmark, the Chinese Christian delegation held a press conference.
Recall Errors

• Coreferent mentions with synonymous triggers
  – Event mentions that have synonymous but lexically different trigger words
Coreferent mentions with synonymous triggers

E1: Jewish [violence] against the Arabs.

E2: Two parties of [conflict].
Recall Errors

• Coreferent mentions with compatible arguments
  – Though two arguments are not coreferent, while they are compatible
E1: Yugoslavia’s head of state [visited] Bosnia-Herzegovina for the first time.

E2: Kostunica [visited] Sarajevo, the capital of Bosnia-Herzegovina.
Plan for the Talk

• ACE Event Coreference
• Six Upstream Components in the Pipeline
• Results and Analysis – Answer to Question 1
• Conclusion
Conclusion

• We analyzed an ACE-style Chinese event coreference system by investigating:
  – The extent to which its performance is affected by the errors made by its upstream components
  – Types of errors made by the resolution algorithm