Coreference Resolution
Lilian Wanzare, Supervised by: Sandro Castronovo

Multimodal Ontology Based Dialogue Systems Seminar

Computational Linguistics Department, Saarland University

June 18, 2010
1 Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2 Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3 Conclusion
   - Summary
   - References
Outline

1. Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2. Approaches
   - The task of Co-reference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3. Conclusion
   - Summary
   - References
Coreference Resolution vs Anaphoric Resolution.

- Coreference chain - a set of co-referent referring expressions in a discourse
- Anaphora - co-reference of one referring expression with its antecedent
- Anaphor - a referring expression (often a pronoun) which refers back to something mentioned previously (e.g. she, this day, the cat . . . but not Peter etc.)
- Co-reference vs. anaphora
  - Co-reference Resolution: find the co-reference chains in a text i.e. the task of identifying noun phrases that refer to the same extralinguistic entity in a text.
  - Anaphora Resolution: find the antecedent of an anaphor.
Coreference Resolution vs Anaphoric Resolution.

- Coreference chain - a set of co-referent referring expressions in a discourse
- Anaphora - co-reference of one referring expression with its antecedent
- Anaphor - a referring expression (often a pronoun) which refers back to something mentioned previously (e.g. she, this day, the cat . . . but not Peter etc.)

co-reference vs. anaphora

- Co-reference Resolution: find the co-reference chains in a text i.e. the task of identifying noun phrases that refer to the same extralinguistic entity in a text.
- Anaphora Resolution: find the antecedent of an anaphor.
Coreference Resolution vs Anaphoric Resolution.

- Coreference chain - a set of co-referent referring expressions in a discourse
- Anaphora - co-reference of one referring expression with its antecedent
- Anaphor - a referring expression (often a pronoun) which refers back to something mentioned previously (e.g. she, this day, the cat . . . but not Peter etc.)

co-reference vs. anaphora

- Co-reference Resolution: find the co-reference chains in a text i.e. the task of identifying noun phrases that refer to the same extralinguistic entity in a text.
- Anaphora Resolution: find the antecedent of an anaphor.
Coreference Resolution vs Anaphoric Resolution.

- Coreference chain - a set of co-referent referring expressions in a discourse
- Anaphora - co-reference of one referring expression with its antecedent
- Anaphor - a referring expression (often a pronoun) which refers back to something mentioned previously (e.g. she, this day, the cat . . . but not Peter etc.)
- co-reference vs. anaphora
  - Co-reference Resolution: find the co-reference chains in a text i.e. the task of identifying noun phrases that refer to the same extralinguistic entity in a text.
  - Anaphora Resolution: find the antecedent of an anaphor.
Examples

- **Anaphors:**
  - Personal pronouns
    - E.g. he, she, it, they, them
  - Possessive pronouns
    - E.g. its, his, etc.
  - Demonstratives
    - This, that, these, those, here, there, now, then
  - Definite pronouns
    - E.g. the girl
  - Comparatives
    - The same as, Different, Another, The third, etc.
Example Sentence

David Beckham won’t be appearing in his fourth World Cup, though. The 35-year-old midfielder tore his left Achilles’ tendon while playing for AC Milan on March 14 and will miss the entire tournament. Still, he is with the Three Lions lending support and giving advice, which could come in handy when England plays the United States - and Beckham’s Los Angeles Galaxy teammates, Landon Donovan and Edson Buddle.\(^a\)

\(^a\)Adapted from http://sportsillustrated.cnn.com/2010/soccer/world-cup-2010/06/12/beckham.wc.ap/index.html
Outline

1 Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2 Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3 Conclusion
   - Summary
   - References
Linguistic constraints

- **Agreement**: Cliff saw the girl. She asked him for help. she = girl
- **Selection preferences**: The monkey ate the banana. It was hungry. it = cat
- **Salience**: Ruth told Lilian she was in danger. she = Ruth
- **Prosody**: Ruth told Lilian SHE was in danger. she = Lilian
- **Lexical semantics**: Ruth warned Lilian she was in danger. she = Lilian
- **World knowledge**: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Linguistic constraints

- **Agreement**: Cliff saw the girl. She asked him for help. she = girl
- **Selection preferences**: The monkey ate the banana. It was hungry. it = cat
- **Salience**: Ruth told Lilian she was in danger. she = Ruth
- **Prosody**: Ruth told Lilian SHE was in danger. she = Lilian
- **Lexical semantics**: Ruth warned Lilian she was in danger. she = Lilian
- **World knowledge**: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Linguistic constraints

- **Agreement**: Cliff saw the girl. She asked him for help. she = girl
- **Selection preferences**: The monkey ate the banana. It was hungry. it = cat
- **Salience**: Ruth told Lilian she was in danger. she = Ruth
- **Prosody**: Ruth told Lilian SHE was in danger. she = Lilian
- **Lexical semantics**: Ruth warned Lilian she was in danger. she = Lilian
- **World knowledge**: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Linguistic constraints

- **Agreement**: Cliff saw the girl. She asked him for help. she = girl
- **Selection preferences**: The monkey ate the banana. It was hungry. it = cat
- **Salience**: Ruth told Lilian she was in danger. she = Ruth
- **Prosody**: Ruth told Lilian SHE was in danger. she = Lilian
- **Lexical semantics**: Ruth warned Lilian she was in danger. she = Lilian
- **World knowledge**: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Linguistic constraints

- Agreement: Cliff saw the girl. She asked him for help. she = girl
- Selection preferences: The monkey ate the banana. It was hungry. it = cat
- Salience: Ruth told Lilian she was in danger. she = Ruth
- Prosody: Ruth told Lilian SHE was in danger. she = Lilian
- Lexical semantics: Ruth warned Lilian she was in danger. she = Lilian
- World knowledge: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Linguistic constraints

- **Agreement**: Cliff saw the girl. She asked him for help. she = girl
- **Selection preferences**: The monkey ate the banana. It was hungry. it = cat
- **Salience**: Ruth told Lilian she was in danger. she = Ruth
- **Prosody**: Ruth told Lilian SHE was in danger. she = Lilian
- **Lexical semantics**: Ruth warned Lilian she was in danger. she = Lilian
- **World knowledge**: President Obama met Nelson Mandela. The old man’s great granddaughter had been involved in an accident. the old man = Nelson Mandela
Motivation
Why is it useful?

Applications where Coreference plays a role:

- (Multimodal) Dialogue Systems

Mobile Shop Assistant

In car system
Example: In car system

System: [A list of available songs is displayed on the screen]
User: “Play the first song.”
System: [Plays the first song]
User: “And the second.”

- Information Extraction
- Questions Answering
- Summarization
- ...

Motivation...
Difficulties
Why is it difficult?

In Multimodal Dialogue systems:

- (personal and demonstrative) pronouns with non-NP-antecedents or no antecedents at all.
- pronouns that pick up different kinds of abstract objects from the previous discourse, e.g. events, states, concepts, propositions or facts

Example

A1: [There is alot of theft, alot of assult dealing with, uh, people trying to get money from drugs ]
B2: I think [that] is a national problem
A1 : [It]'s shocking how things have changed over the years...
In Multimodal Dialogue systems:

- (personal and demonstrative) pronouns with non-NP-antecedents or no antecedents at all.
- Pronouns that pick up different kinds of abstract objects from the previous discourse, e.g. events, states, concepts, propositions or facts.

Example:

A1: [There is a lot of theft, a lot of assault dealing with, uh, people trying to get money from drugs]i
B2: I think [that]i is a national problem
A1: [It]’s shocking how things have changed over the years...
Difficulties...

- Different forms $\not\Rightarrow$ different referents
  
  \textit{(David Beckham vs The 35-year-old midfielder vs he )}

- Same form $\not\Rightarrow$ same referent
  
  (Andreas the student vs Andreas the lecturer)

- Not all “anaphoric“ expressions are anaphora: \textbf{David Beckham} won’t be appearing in his fourth World Cup, though.

- Incorporates both linguistic and non-linguistic knowledge
Introduction

- Basic Idea
- Factors affecting co-reference resolution

Approaches

- The task of Coreference resolution
- Early Implementations: Rule based
- Machine Learning: Corpus based

Conclusion

- Summary
- References
What do we need to do?

Generally...

- Identify tokens (words) and sentences
- Tag text with POS
- Parsing
  - Identify Named Entities
  - Identify type of NP: E.g. the-np, pronoun
  - Identify some agreement features: PER, NUM, GEN ...
  - Internal structure of NEs: E.g. modifiers and heads
  - Identify anaphors / markables
  - Identify potential antecedents
  - Find coreferent for each anaphor / markable
What do we need to do?

Generally...

- Identify tokens (words) and sentences
- Tag text with POS
- Parsing
- Identify Named Entities
- Identify type of NP: E.g. the-np, pronoun
- Identify some agreement features: PER, NUM, GEN ...
- Internal structure of NEs: E.g. modifiers and heads
- Identify anaphors / markables
- Identify potential antecedents
- Find coreferent for each anaphor / markable
What do we need to do?

Generally...

- Identify tokens (words) and sentences
- Tag text with POS
- Parsing
- Identify Named Entities
- Identify type of NP: E.g. the-np, pronoun
- Identify some agreement features: PER, NUM, GEN ...
- Internal structure of NEs: E.g. modifiers and heads
- Identify anaphors / markables
- Identify potential antecedents
- Find coreferent for each anaphor / markable
What do we need to do?

For Multimodal Dialogue systems, one needs to take care of:

- contextual information
  - General knowledge.
  - Physical context: the physical surroundings of a dialogue.
  - Linguistic context: previous utterances or text.

Example

User: “What’s playing at the movies tonight?”
System: [Displays a list of movies] “Here [↗] you can see tonight’s movie theater program.”
User: “And what’s on TV?”
System: [Displays a list of broadcasts] “Here [↗] are tonight’s broadcasts.”
User: “Is there a [one] with Arnold Schwarzenegger?”
What do we need to do?

For Multimodal Dialogue systems, one needs to take care of:

- **contextual information**
  - General knowledge.
  - Physical context: the physical surroundings of a dialogue.
  - Linguistic context: previous utterances or text.

**Example**

User: “What’s playing at the movies tonight?”
System: [Displays a list of movies] “Here [↗] you can see tonight’s movie theater program.”
User: “And what’s on TV?”
System: [Displays a list of broadcasts] “Here [↗] are tonight’s broadcasts.”
User: “Is there a [one] with Arnold Schwarzenegger?”
What do we need to do?

For Multimodal Dialogue systems, one needs to take care of:

- Multimodal fusion
- Incrementally build up a discourse model.
- Store and update entities in it.
- Add conversational metadata e.g. speaker, turn, type of utterance
Outline

1 Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2 Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
     - Machine Learning: Corpus based

3 Conclusion
   - Summary
   - References
Resolution of Anaphora Procedure

Salience weights

<table>
<thead>
<tr>
<th>Salience weights</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence recency</td>
<td>100</td>
</tr>
<tr>
<td>Subject emphasis</td>
<td>80</td>
</tr>
<tr>
<td>Existential emphasis</td>
<td>70</td>
</tr>
<tr>
<td>Accusative emphasis</td>
<td>50</td>
</tr>
<tr>
<td>Indirect object and oblique complement emphasis</td>
<td>40</td>
</tr>
<tr>
<td>Non-adverbial emphasis</td>
<td>50</td>
</tr>
<tr>
<td>Head noun emphasis</td>
<td>80</td>
</tr>
</tbody>
</table>
RAP Algorithm

1. Collect the potential referents (up to four sentences back).
2. Remove potential referents that do not agree in number or gender with the pronoun.
3. Remove potential referents that do not pass intrasentential syntactic coreference constraints.
4. Compute the total salience value of the referent by adding any applicable values for role parallelism (+35) or cataphora (-175).
5. Select the referent with the highest salience value. In case of a tie, select the closest referent in terms of string position.
Example: Pronoun resolution

John Smith talks about the EU. He likes the family of nations.

Weights:

- **John Smith**: 98 (recency) + 78 (subj) + 78 (head noun) + 48 (non-adv) = 302
- **the EU**: 98 (recency) + 48 (acc) + 78 (head noun) + 48 (non-adv) = 272
- **the family of nations**: 100 (recency) + 50 (acc) + 80 (head noun) + 50 (non-adv) = 280
- **nations**: 100 (recency) + 50 (acc) + 50 (non-adv) = 200

Resolving “he”:

- “he“ = “John Smith“ by morpho-syntactic filter

---

1 Adapted from Slides by Caroline Sporleder CoLi, 2009/10
Example: Pronoun resolution

John Smith talks about the EU. He likes the family of nations.

Weights:

- John Smith: 98 (recency) + 78 (subj) + 78 (head noun) + 48 (non-adv) = 302
- the EU: 98 (recency) + 48 (acc) + 78 (head noun) + 48 (non-adv) = 272
- the family of nations: 100 (recency) + 50 (acc) + 80 (head noun) + 50 (non-adv) = 280
- nations: 100 (recency) + 50 (acc) + 50 (non-adv) = 200

Resolving “he”:

“he” = “John Smith” by morpho-syntactic filter

1Adapted from Slides by Caroline Sporleder CoLi, 2009/10
Example: Pronoun resolution

John Smith talks about the EU. He likes the family of nations.

Weights:

- John Smith: 98 (recency) + 78 (subj) + 78 (head noun) + 48 (non-adv) = 302
- the EU: 98 (recency) + 48 (acc) + 78 (head noun) + 48 (non-adv) = 272
- the family of nations: 100 (recency) + 50 (acc) + 80 (head noun) + 50 (non-adv) = 280
- nations: 100 (recency) + 50 (acc) + 50 (non-adv) = 200

Resolving “he”:

- “he” = “John Smith” by morpho-syntactic filter

---

1 Adapted from Slides by Caroline Sporleder CoLi, 2009/10
Example: Pronoun resolution

John Smith talks about the EU. He likes the family of nations.

Weights:

- John Smith: 98 (recency) + 78 (subj) + 78 (head noun) + 48 (non-adv) = 302
- the EU: 98 (recency) + 48 (acc) + 78 (head noun) + 48 (non-adv) = 272
- the family of nations: 100 (recency) + 50 (acc) + 80 (head noun) + 50 (non-adv) = 280
- nations: 100 (recency) + 50 (acc) + 50 (non-adv) = 200

Resolving “he”:

- “he” = “John Smith” by morpho-syntactic filter

---

1 Adapted from Slides by Caroline Sporleder CoLi, 2009/10
Example: Pronoun resolution

John Smith talks about the EU. He likes the family of nations.

Weights:

- John Smith: $98 \text{ (recency)} + 78 \text{ (subj)} + 78 \text{ (head noun)} + 48 \text{ (non-adv)} = 302$
- the EU: $98 \text{ (recency)} + 48 \text{ (acc)} + 78 \text{ (head noun)} + 48 \text{ (non-adv)} = 272$
- the family of nations: $100 \text{ (recency)} + 50 \text{ (acc)} + 80 \text{ (head noun)} + 50 \text{ (non-adv)} = 280$
- nations: $100 \text{ (recency)} + 50 \text{ (acc)} + 50 \text{ (non-adv)} = 200$

Resolving “he”:

- “he” = “John Smith” by morpho-syntactic filter

---

1Adapted from Slides by Caroline Sporleder CoLi, 2009/10
Outline

1. Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2. Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3. Conclusion
   - Summary
   - References
BART is a modular toolkit for coreference resolution proposed by Yannick et al.

System architecture

- Uses MMAX2 for preprocessing:
  - marking out noun chunks and named entities
  - POS tagging
  - Merging this information into markables

- Chunking pipeline:
  - Stanford POS tagger
  - YAMCha chunker
  - Stanford Named Entity recognizer
BART...

- Parsing pipeline:
  - Charniak and Johnson’s reranking parser

- Carafe pipeline: Uses Ace mentions tagger by MITRE, a specialized tagger then discards any base NP that was not detected to be an ACE mention

- Feature extraction: Each pair of anaphor and candidate is represented in a PairInstance object, which is enriched with classification features by feature extractors
Learning: The pairInstance is handed over to a machine-learning based classifier

- BART uses
  - WEKA machine learning toolkit
  - SVMLight
  - Maximum entropy classifier.

Training: The pairs that are to be used as training examples have to be selected in a process of sample selection,

Testing: It has to be decided which pairs are to be given to the decision function and how to group mentions into equivalence relations given the classifier decisions. Training and testing part of encoder/decoder.
BART System Configuration

Unannotated Text → Preprocessing → ACE Mention Tagger → Merger

Parser

Basic features

Syntactic features

Knowledge-based features

Mention (with basic properties):
- Number
- Gender
- Mention Type
- Modifiers

Mention Factory

Decoder

SVM Classifier

Coreference Chains
BART demo

Demo...
The task of Coreference resolution

Early Implementations: Rule based

Machine Learning: Corpus based

BART performance on ACE-2

<table>
<thead>
<tr>
<th></th>
<th>BNews</th>
<th></th>
<th>NPaper</th>
<th></th>
<th>NWire</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recl</td>
<td>Prec</td>
<td>F</td>
<td>Recl</td>
<td>Prec</td>
<td>F</td>
</tr>
<tr>
<td>basic feature set</td>
<td>0.594</td>
<td>0.522</td>
<td>0.556</td>
<td>0.663</td>
<td>0.526</td>
<td>0.586</td>
</tr>
<tr>
<td>extended feature set</td>
<td>0.607</td>
<td>0.654</td>
<td>0.630</td>
<td>0.641</td>
<td>0.677</td>
<td>0.658</td>
</tr>
<tr>
<td>Ng 2007*</td>
<td>0.561</td>
<td>0.763</td>
<td>0.647</td>
<td>0.544</td>
<td>0.797</td>
<td>0.646</td>
</tr>
</tbody>
</table>

*: “expanded feature set” in Ng 2007; Ng trains on the entire ACE training corpus.
Outline

1. Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2. Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3. Conclusion
   - Summary
   - References
Key points

- Coreference resolution vs Anaphora resolution
- Difficulties and motivation
- The tasks involved
- Approaches:
  - Rule based
    - RAP
  - Corpus based
    - BART
Outline

1 Introduction
   - Basic Idea
   - Factors affecting co-reference resolution

2 Approaches
   - The task of Coreference resolution
   - Early Implementations: Rule based
   - Machine Learning: Corpus based

3 Conclusion
   - Summary
   - References

Multimodal Ontology Based Dialogue Systems Seminar
Coreference Resolution
References

- Slides by Pascal Denis Alpage INRIA Paris-Rocquencourt, 2010. *Automatic Reference Resolution*  
References...