Natural Language Processing (CS-472) Spring-2023

Muhammad Naseer Bajwa

Assistant Professor,
Department of Computing, SEECS
Co-Principal Investigator,
Deep Learning Lab, NCAI
NUST, Islamabad
naseer.bajwa@seecs.edu.pk

Overview of this week's lecture

Coreference Resolution

- Introduction to Coreference Resolution
- Coreference resolution approaches
  - Using mention-pair
  - Using mention-pair ranking
  - Using clustering
- Coreference resolution evaluation

Coreference resolution means resolving coreference

- Given a text, identify all mentions that refer to a real word entity.

Coreference resolution means resolving coreference

- Given a text, identify all mentions that refer to a real word entity.
    - First identify all mentions.

Elon Musk turns Twitter into 'hotel' for staff. An ex-worker said new Twitter boss Elon Musk has been staying at the headquarters since he bought the firm. California state senator Scott Wiener told the BBC on Wednesday: "He's now making them [workers] sleep at the company.

NUST
Defining futures
School of Electrical Engineering
& Computer Science

Coreference resolution means resolving coreference

- Given a text, identify all mentions that refer to a real word entity.
  - First identify all mentions.

Elon Musk turns **Twitter** into 'hotel' for **staff**. An ex-worker said new **Twitter boss Elon Musk** has been staying at the **headquarters** since **he** bought the **firm**. **California state senator Scott Wiener** told the **BBC** on Wednesday: "**He**'s now making **them [workers]** sleep at **the company**.

Coreference resolution means resolving coreference

- Given a text, identify all mentions that refer to a real word entity.
  - First identify all mentions.

Elon Musk turns **Twitter** into 'hotel' for **staff**. An ex-worker said new **Twitter boss Elon Musk** has been staying at the **headquarters** since **he** bought the **firm**. **California state senator Scott Wiener** told the **BBC** on Wednesday: "**He**'s now making **them [workers]** sleep at **the company**.

Coreference resolution can be tricky even for humans

The Islamabad High Court on Thursday barred FIA) from arresting Suleman Shehbaz, the son of Prime Minister Shehbaz Sharif, who is expected to return to Pakistan this week after four years of self-exile in London.

Suleman is living in London with his family. He had filed a plea before the IHC a day ago, seeking protective bail that would enable him to surrender before a trial court.
The court took up the plea for hearing today. Suleman’s counsel Amjad Pervez appeared before the court.
During the proceedings, IHC Chief Justice Aamer Farooq issued directives for the applicant, asking him to surrender before the court by Dec 13 and barred the FIA from arresting him until then. Justice Farooq said it was necessary for an applicant to be present in the court while seeking protective bail.

Pervez told the court that his client was returning to the country on Sunday. “He wants to appear before the relevant authorities,” he added. The counsel said the court had been granting similar protective bails in the past.

Following the arguments, the court granted them relief.

Coreference resolution can be tricky even for humans

The Islamabad High Court on Thursday barred FIA) from arresting Suleman Shehbaz, the son of Prime Minister Shehbaz Sharif, who is expected to return to Pakistan this week after four years of self-exile in London.

Suleman is living in London with his family. He had filed a plea before the IHC a day ago, seeking protective bail that would enable him to surrender before a trial court.
The court took up the plea for hearing today. Suleman's counsel Amjad Pervez appeared before the court.
During the proceedings, IHC Chief Justice Aamer Farooq issued directives for the applicant, asking him to surrender before the court by Dec 13 and barred the FIA from arresting him until then. Justice Farooq said it was necessary for an applicant to be present in the court while seeking protective bail.

Pervez told the court that his client was returning to the country on Sunday. "He wants to appear before the relevant authorities," he added. The counsel said the court had been granting similar protective bails in the past.

Following the arguments, the court granted them relief.

Split Antecedent

Why is coreference resolution important?

Why is coreference resolution important?

- To fully understand natural language.
    - Information extraction.

Salman’s advocate said that his client was living in London for the last four years and now he wanted to come back to Pakistan.

Who was living in London for the last four years?

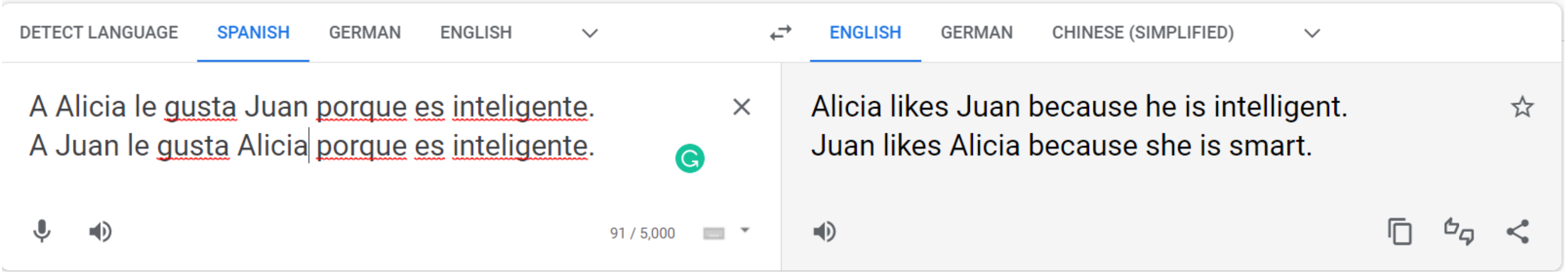
Why is coreference resolution important?

- To fully understand natural language.
  - Information extraction.

Salman's advocate said that his client was living in London for the last four years and now he wanted to come back to Pakistan.

Who was living in London for the last four years?

  - Machine translation



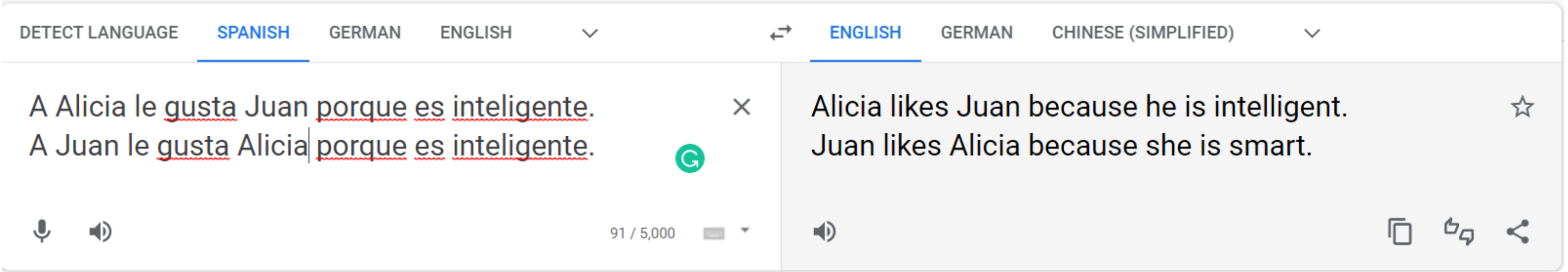
Why is coreference resolution important?

- To fully understand natural language.
  - Information extraction.

Salman's advocate said that his client was living in London for the last four years and now he wanted to come back to Pakistan.

Who was living in London for the last four years?

  - Machine translation



  - Dialogue Generation

Human: Book tickets to see James Bond.
Bot: Spectre is playing near you at 02:00 and 05:00. How many tickets would you like?
Human: Two, for showing at five.

Coreference resolution is performed in two steps

- Detect the mentions.

“ I voted for John Doe because he was most aligned with my values ”, she said.

Coreference resolution is performed in two steps

- Detect the mentions.
  - Mentions can be nested.

“[I] voted for [John Doe] because [he] was most aligned with [[my] values]”, [she] said.

Coreference resolution is performed in two steps

- Detect the mentions.
  - Mentions can be nested.
- Cluster the mentions.
  - Which mention is coreferent to what other mention.

"[I] voted for [John Doe] because [he] was most aligned with [[my] values]", [she] said.

"[I] voted for [John Doe] because [he] was most aligned with [[my] values]", [she] said.

How to do mention detection?

- A mention is a span of text referring to some entity.

How to do mention detection?

- A mention is a span of text referring to some entity.
- There are three types of mentions.
  - Pronouns: I, you, it, she, him, etc.
  - Named Entities People, Places, Organisations, etc.
  - Noun Phrases “A dog”, “the big fluffy cat struck in the tree”

How to do mention detection?

- A mention is a span of text referring to some entity.

- There are three types of mentions.

		How to detect them?
- Pronouns:	I, you, it, she, him, etc.	- Use Part-of-Speech (POS) taggers
- Named Entities	People, Places, Organisations, etc.	- Use NER systems
- Noun Phrases	“A dog”, “the big fluffy cat struck in the tree”	- Use a constituency parser

How to do mention detection?

- A mention is a span of text referring to some entity.

- There are three types of mentions.

Type	Examples	How to detect them?
Pronouns:	I, you, it, she, him, etc.	Use Part-of-Speech (POS) taggers
Named Entities	People, Places, Organisations, etc.	Use NER systems
Noun Phrases	“A dog”, “the big fluffy cat struck in the tree”	Use a constituency parser

- Detecting different types of mentions using separate methods serves as a pre-processing step.

NUST Defining futures
School of Electrical Engineering & Computer Science

How to do mention detection?

- A mention is a span of text referring to some entity.

- There are three types of mentions.

Type	Examples	How to detect them?
Pronouns:	I, you, it, she, him, etc.	Use Part-of-Speech (POS) taggers
Named Entities	People, Places, Organisations, etc.	Use NER systems
Noun Phrases	“A dog”, “the big fluffy cat struck in the tree”	Use a constituency parser

- Detecting different types of mentions using separate methods serves as a pre-processing step.

- You can also train a classifier that runs on the text and detects all types of mentions.

How to do mention detection?

- A mention is a span of text referring to some entity.
- There are three types of mentions.

Type	Examples	How to detect them?
Pronouns:	I, you, it, she, him, etc.	Use Part-of-Speech (POS) taggers
Named Entities	People, Places, Organisations, etc.	Use NER systems
Noun Phrases	"A dog", "the big fluffy cat struck in the tree"	Use a constituency parser

- Detecting different types of mentions using separate methods serves as a pre-processing step.
- You can also train a classifier that runs on the text and detects all types of mentions.
- You can also perform mention detection and coreference resolution simultaneously in end-to-end manner.
    - More on this in coming slides.

Mention detection can be tricky due to ambiguity in text

- Sometimes in order to capture all mentions, the system may over-generate mentions.

Mention detection can be tricky due to ambiguity in text

- Sometimes in order to capture all mentions, the system may over-generate mentions.

- The word ‘it’ is a mention when used as a pronoun. But is it a mention in the following sentence?

    - It is raining.

Mention detection can be tricky due to ambiguity in text

- Sometimes in order to capture all mentions, the system may over-generate mentions.

- The word ‘it’ is a mention when used as a pronoun. But is it a mention in the following sentence?

    - It is raining.

- Similarly, are the following words/phrases mention or not mention?

    - *No student* got an A Grade. (Is *no student* any reference?)
    - *Every student* got an A Grade. (Is *every student* a clear and concrete reference?)
    - The best teacher in the word. (Unclear. May or may not be)
    - Thousand miles (Not a mention)

Mention detection can be tricky due to ambiguity in text

- Sometimes in order to capture all mentions, the system may over-generate mentions.
- The word ‘it’ is a mention when used as a pronoun. But is it a mention in the following sentence?
    - It is raining.
- Similarly, are the following words/phrases mention or not mention?
    - *No student* got an A Grade. (Is *no student* any reference?)
    - *Every student* got an A Grade. (Is *every student* a clear and concrete reference?)
    - The best teacher in the word. (Unclear. May or may not be)
    - Thousand miles (Not a mention)
- How to deal with bad mentions?
    - Train a classifier to filter out such spurious mentions.

Mention detection can be tricky due to ambiguity in text

- Sometimes in order to capture all mentions, the system may over-generate mentions.

- The word ‘it’ is a mention when used as a pronoun. But is it a mention in the following sentence?

  - It is raining.

- Similarly, are the following words/phrases mention or not mention?

  - *No student* got an A Grade. (Is *no student* any reference?)
  - *Every student* got an A Grade. (Is *every student* a clear and concrete reference?)
  - The best teacher in the word. (Unclear. May or may not be)
  - Thousand miles (Not a mention)

- How to deal with bad mentions?

  - Train a classifier to filter out such spurious mentions.

  - Keep all mentions as candidate mentions and after clustering remove singletons.

Let’s learn just a tiny bit of linguistics

- Coreference is when two mentions refer to the same entity in the world.
    - “***Donald Trump’****s twitter account reinstated.* ***Trump*** *was banned from twitter in January 2021 after January 6 capitol riots*”.
        - Both Donald Trump and Trump refer to the same entity but both mentions can be understood and interpreted independently of each other.

Let’s learn just a tiny bit of linguistics

- Coreference is when two mentions refer to the same entity in the world.
  - “***Donald Trump’****s twitter account reinstated.* ***Trump*** *was banned from twitter in January 2021 after January 6 capitol riots*”.
    - Both Donald Trump and Trump refer to the same entity but both mentions can be understood and interpreted independently of each other.
- Anaphora is a related linguistic concept where a term (anaphor) refers to another term (antecedent) for interpretation.
  - **Trump** reacted to reactivation of **his** twitter account by saying **he** wouldn’t be quickly returning to twitter.

  Antecedent Anaphor Anaphor

Let’s learn just a tiny bit of linguistics

- Coreference is when two mentions refer to the same entity in the world.
    - “***Donald Trump’***s *twitter account reinstated.* ***Trump*** *was banned from twitter in January 2021 after January 6 capitol riots*”.
        - Both Donald Trump and Trump refer to the same entity but both mentions can be understood and interpreted independently of each other.
- Anaphora is a related linguistic concept where a term (anaphor) refers to another term (antecedent) for interpretation.
    - **Trump** reacted to reactivation of **his** twitter account by saying **he** wouldn’t be quickly returning to twitter.

  Antecedent Anaphor Anaphor
- The concept of anaphors (and cataphors, coming soon) is not widely used in NLP.

Let’s learn just a tiny bit of linguistics (continued)

- Barak Obama visited Florida. During the visit, Obama said his government will increase budget for health.

Text

- Barak Obama
- Obama

Real World

Let’s learn just a tiny bit of linguistics (continued)

- Barak Obama visited Florida. During the visit, Obama said his government will increase budget for health.

Text

- Barak Obama
- Obama

Real World

- Barak Obama visited Florida. He said his government will increase budget for health.

Text

- Barak Obama
- He

Real World

Let’s learn just a tiny bit of linguistics (continued)

- Barak Obama visited Florida. During the visit, Obama said his government will increase budget for health.

انہوں نے کہا،

Text

- Barak Obama
- Obama

Real World

- Barak Obama visited Florida. He said his government will increase budget for health.

Text

- Barak Obama
- He

Real World

Let's learn just a tiny bit of linguistics (continued)

- Barak Obama visited Florida. During the visit, Obama said his government will increase budget for health.

Text

- Barak Obama
- Obama

Real World

انہوں نے کہا، میں مانگنے نہیں آیا لیکن مجبوری ہے.

- Barak Obama visited Florida. He said his government will increase budget for health.

Text

- Barak Obama
- He

Real World

- Barak Obama visited Florida. During the visit, Obama said his government will increase budget for health.

Text

- Barak Obama
- Obama

Real World

انہوں نے کہا، میں مانگنے نہیں آیا لیکن مجبوری ہے۔

- Barak Obama visited Florida. He said his government will increase budget for health.

Text

- Barak Obama
- He

Real World

?

Real World

Not all anaphoric relations are coreferential

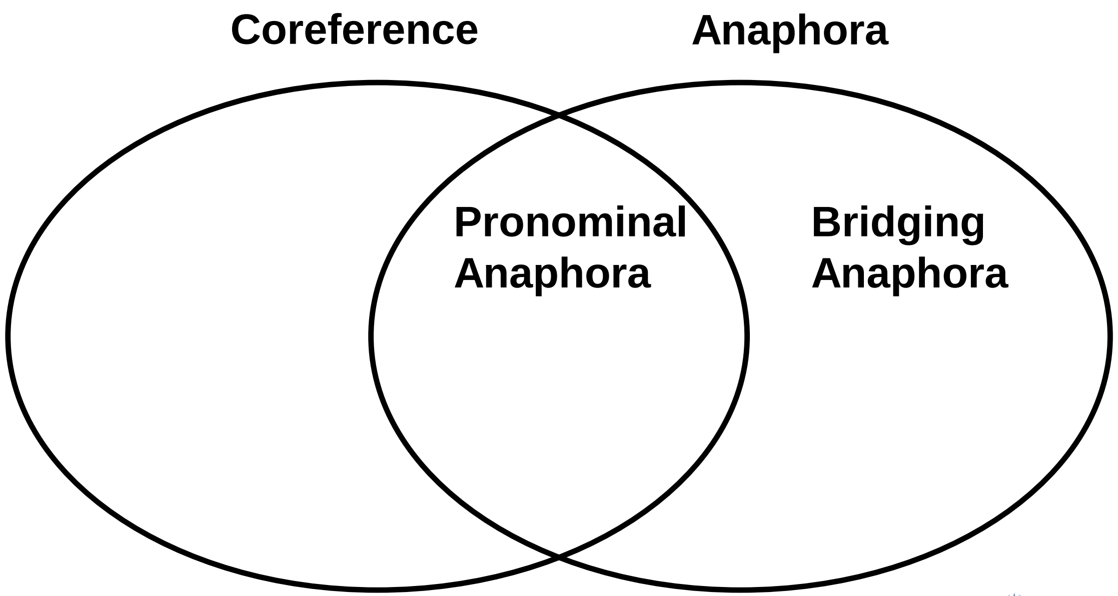
- Not every noun phrase has reference.
  - No dancer twisted her knee.
  - Every dancer twisted her knee.

Not all anaphoric relations are coreferential

- Not every noun phrase has reference.
    - No dancer twisted her knee.
    - Every dancer twisted her knee.
- We went to see the match. The tickets were already sold.

Not all anaphoric relations are coreferential

- Not every noun phrase has reference.
  - No dancer twisted her knee.
  - Every dancer twisted her knee.
- We went to see the match. The tickets were already sold.
  - This is a bridging anaphora.



Anaphora vs Cataphora

- Position of antecedent with reference to the dependent term is important, linguistically.

  - The king, in the first speech after his coronation, freed the prisoners.

    Antecedent **Anaphora**

Anaphora vs Cataphora

- Position of antecedent with reference to the dependent term is important, linguistically.

  - The king, in the first speech after his coronation, freed the prisoners.

    Antecedent **Anaphora**

  - In the first speech after his coronation, the king freed the prisoners.

    Cataphora **Antecedent**

Anaphora vs Cataphora

- Position of antecedent with reference to the dependent term is important, linguistically.

  - The king, in the first speech after his coronation, freed the prisoners.
    Antecedent **Anaphora**

  - In the first speech after his coronation, the king freed the prisoners.
    Cataphora **Antecedent**

- In modern linguistics, the term cataphora is completely disused.

NUST Defining futures
School of Electrical Engineering & Computer Science

Anaphora vs Cataphora

- Position of antecedent with reference to the dependent term is important, linguistically.

  - The king, in the first speech after his coronation, freed the prisoners.

    Antecedent **Anaphora**

  - In the first speech after his coronation, the king freed the prisoners.

    Cataphora **Antecedent**

- In modern linguistics, the term cataphora is completely disused.

- In NLP also, the position of antecedent is not paid much attention to.

  - NLP systems only look back to find the textual reference (antecedent).

Sometimes knowledge of the world in important to resolve coreference

- She poured water from the pitcher into the cup until it was full.

- She poured water from the pitcher into the cup until it was empty.

Sometimes knowledge of the world in important to resolve coreference

- She poured water from the pitcher into the cup until it was full.
- She poured water from the pitcher into the cup until it was empty.
- The city council refused the protesters a permit because they feared violence.
- The city council refused the protesters a permit because they wanted violence.

Sometimes knowledge of the world in important to resolve coreference

- She poured water from the pitcher into the cup until it was full.
- She poured water from the pitcher into the cup until it was empty.
- The city council refused the protesters a permit because they feared violence.
- The city council refused the protesters a permit because they wanted violence.

Winograd Schema
An alternative to Turing test

Sometimes knowledge of the world in important to resolve coreference

- She poured water from the pitcher into the cup until it was full.
- She poured water from the pitcher into the cup until it was empty.
- The city council refused the protesters a permit because they feared violence.
- The city council refused the protesters a permit because they wanted violence.

Winograd Schema
An alternative to Turing test

https://www.youtube.com/watch?v=fKk9KhGRBdI

Sometimes knowledge of the world in important to resolve coreference

- She poured water from the pitcher into the cup until it was full.
- She poured water from the pitcher into the cup until it was empty.
- The city council refused the protesters a permit because they feared violence.
- The city council refused the protesters a permit because they wanted violence.

Winograd Schema
An alternative to Turing test

"If you have fully solved coreference, arguably you have fully solved AI."

Hector J. Levesque

https://www.youtube.com/watch?v=fKk9KhGRBdI

Levesque, Hector J. "On our best behaviour." Artificial Intelligence 212 (2014): 27-35.

- Take pairs of mentions and train a binary classifier to classify whether the pair is coreferent or not.

"I voted for John Doe because he was most aligned with my values", she said.

I | **John Doe** | **He** | **My** | **She**

Cluster 1
Cluster 2

Coreference Model: Mention Pair

- Take pairs of mentions and train a binary classifier to classify whether the pair is coreferent or not.
  - The classifier assigns every pair of mentions a probability $p(m_i, m_j)$.
  - Read the text from left to right.
  - When a mention is encountered, compute its probability will all preceding mentions.
    - For positive examples, the probability should be high.

"I voted for John Doe because he was most aligned with my values", she said.

I | **John Doe** | **He** | **My** | **She**

Cluster 1
Cluster 2

Coreference Model: Mention Pair

- Take pairs of mentions and train a binary classifier to classify whether the pair is coreferent or not.
  - The classifier assigns every pair of mentions a probability $p(m_i, m_j)$.
  - Read the text from left to right.
  - When a mention is encountered, compute its probability will all preceding mentions.
    - For positive examples, the probability should be high.

"I voted for John Doe because he was most aligned with my values", she said.

I | John Doe | He | My | She

Cluster 1
Cluster 2

$$J = -\sum_{i=2}^{N}\sum_{j=1}^{i-1} y_{ij} \log p(m_j, m_i)$$

Coreference Model: Mention Pair

- Take pairs of mentions and train a binary classifier to classify whether the pair is coreferent or not.
    - The classifier assigns every pair of mentions a probability $p(m_i, m_j)$.
    - Read the text from left to right.
    - When a mention is encountered, compute its probability will all preceding mentions.
        - For positive examples, the probability should be high.

"I voted for John Doe because he was most aligned with my values", she said.

I	John Doe	He	My	She

Cluster 1
Cluster 2

$$J = -\sum_{i=2}^{N}\sum_{j=1}^{i-1} y_{ij} \log p(m_j, m_i)$$

Iterate through mentions

Iterate through previous mentions (candidate antecedents)

Coreference Model: Mention Pair

- Take pairs of mentions and train a binary classifier to classify whether the pair is coreferent or not.
  - The classifier assigns every pair of mentions a probability $p(m_i, m_j)$.
  - Read the text from left to right.
  - When a mention is encountered, compute its probability will all preceding mentions.
    - For positive examples, the probability should be high.

“I voted for John Doe because he was most aligned with my values”, she said.

I | **John Doe** | **He** | **My** | **She**

Cluster 1
Cluster 2

What about clustering?

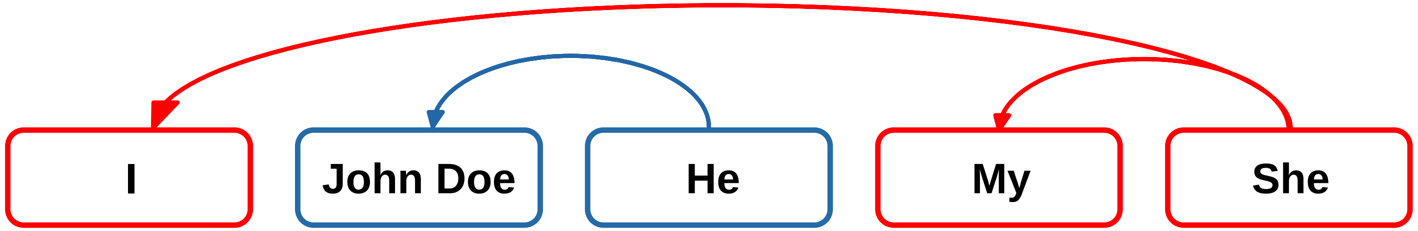
$$J = -\sum_{i=2}^{N}\sum_{j=1}^{i-1} y_{ij} \log p(m_j, m_i)$$

Iterate through mentions

Iterate through previous mentions (candidate antecedents)

- Pick a threshold, say 0.5, and add reference links between mentions where $p(m_i, m_j)$ is higher.

“I voted for John Doe because he was most aligned with my values”, she said.



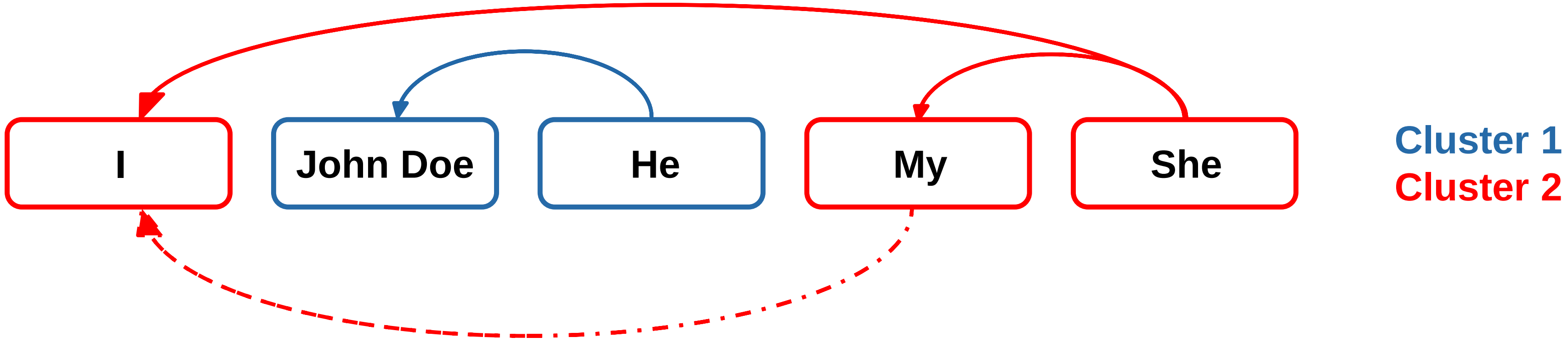
Cluster 1
Cluster 2

NUST
Defining futures
School of Electrical Engineering & Computer Science

Coreference Model: Mention Pair Test Time

- Pick a threshold, say 0.5, and add reference links between mentions where $p(m_i, m_j)$ is higher.
    - My is connected to I by transitive closure.

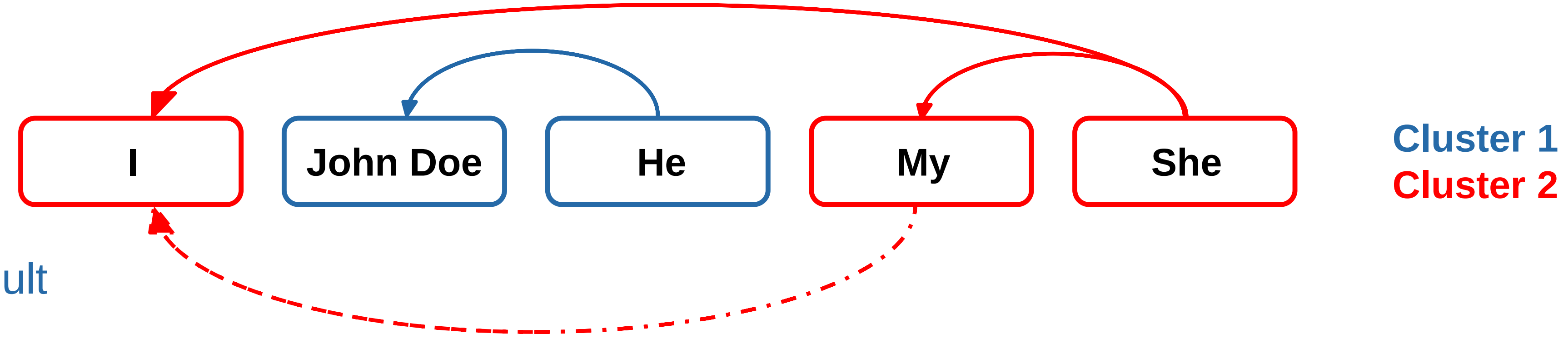
“I voted for John Doe because he was most aligned with my values”, she said.



Coreference Model: Mention Pair Test Time

- Pick a threshold, say 0.5, and add reference links between mentions where $p(m_i, m_j)$ is higher.

“I voted for John Doe because he was most aligned with my values”, she said.

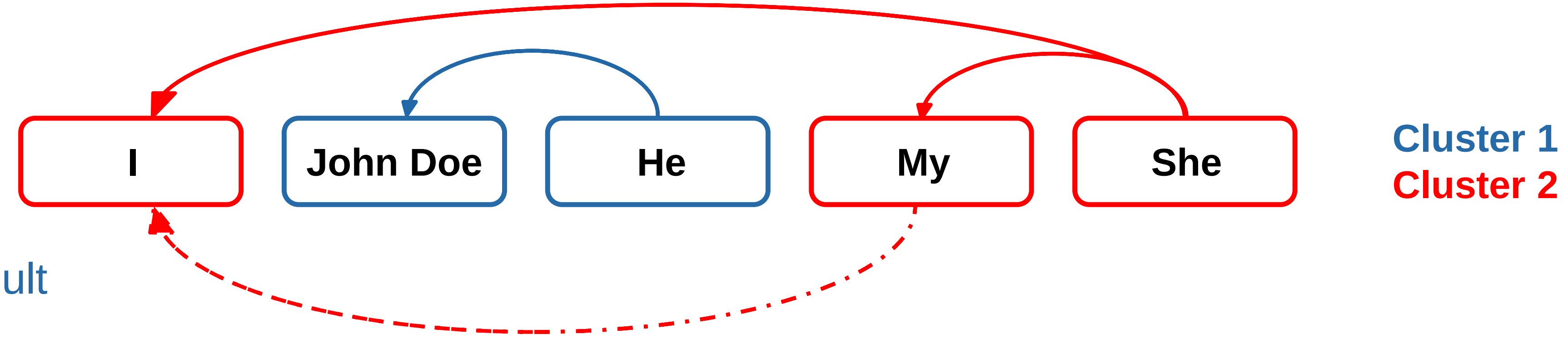


  - My is connected to I by transitive closure.
  - Linking mentions using transitive closure may result into over clustering.
  - Some mentions may be coreferent to nothing.

Coreference Model: Mention Pair Test Time

- Pick a threshold, say 0.5, and add reference links between mentions where $p(m_i, m_j)$ is higher.

“I voted for John Doe because he was most aligned with my values”, she said.

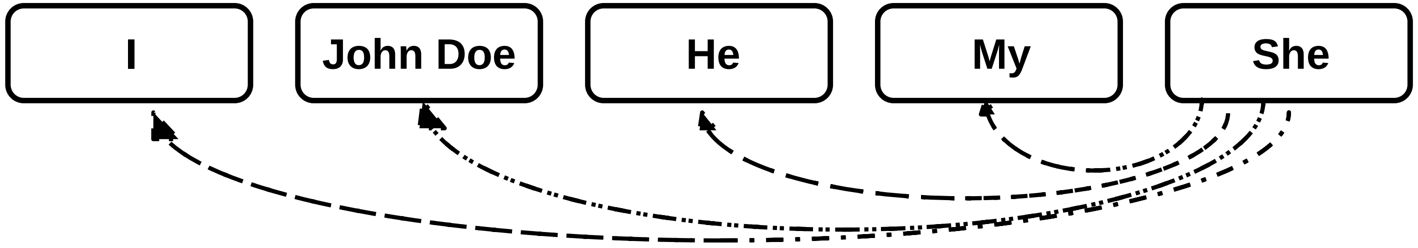


  - My is connected to I by transitive closure.
  - Linking mentions using transitive closure may result into over clustering.
  - Some mentions may be coreferent to nothing.
- Disadvantage: The mention-pair model does not work well for longer documents.
  - Many mentions only have one clear antecedent but mention-pair model tries to identify all of them.
  - Solution: Train the model to predict only one antecedent per mention.

Coreference Model: Mention Ranking

- Assign each mention its highest ranking candidate antecedent.
  - What's the best antecedent for She?

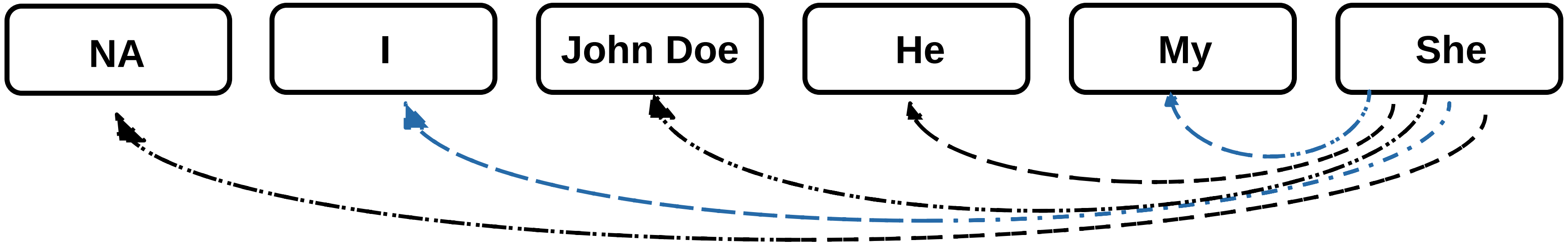
"I voted for John Doe because he was most aligned with my values", she said.



Coreference Model: Mention Ranking

- Assign each mention its highest ranking candidate antecedent.
  - What's the best antecedent for She?
- What about singletons?
  - Use dummy antecedent.

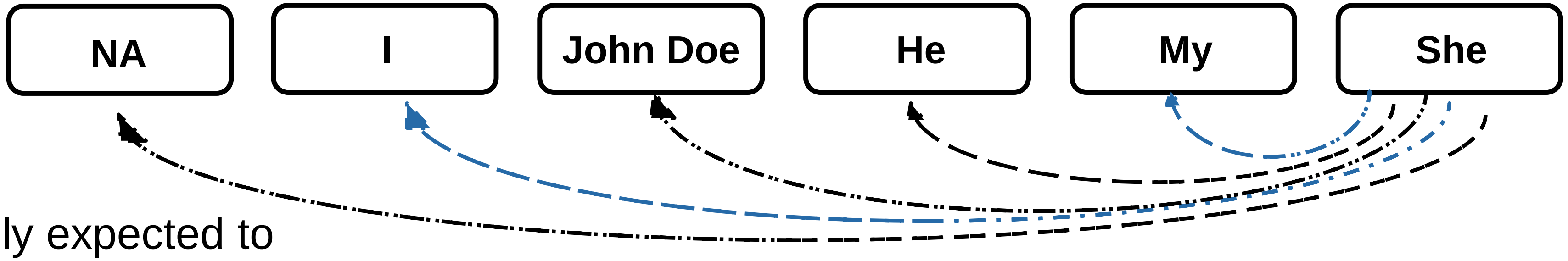
“I voted for John Doe because he was most aligned with my values”, she said.



Coreference Model: Mention Ranking

- Assign each mention its highest ranking candidate antecedent.
  - What's the best antecedent for She?
- What about singletons?
  - Use dummy antecedent.
- During training, the model is only expected to assign high probability to any one of the antecedents (positive examples).

"I voted for John Doe because he was most aligned with my values", she said.



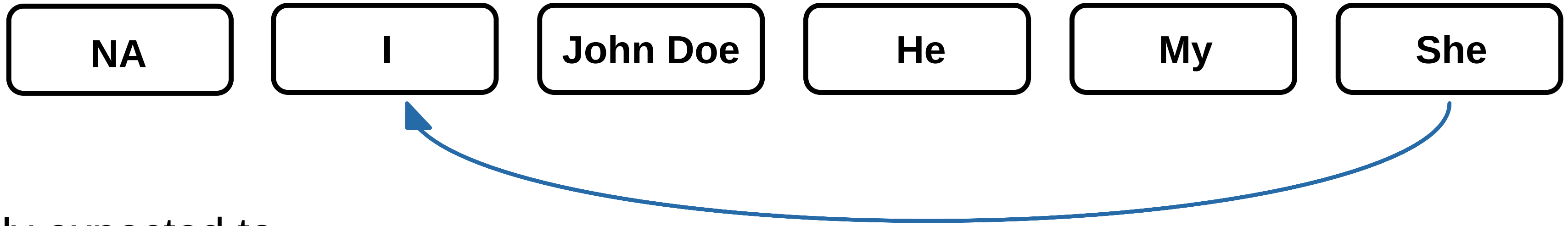
$$p(NA, she) = 0.1$$
$$p(I, she) = 0.5$$
$$p(John\ Doe, she) = 0.1$$
$$p(he, she) = 0.1$$
$$p(my, she) = 0.2$$

Coreference Model: Mention Ranking

- Assign each mention its highest ranking candidate antecedent.
  - What's the best antecedent for She?

“I voted for John Doe because he was most aligned with my values”, she said.

- What about singletons?
  - Use dummy antecedent.



- During training, the model is only expected to assign high probability to any one of the antecedents (positive examples).
- Add only the highest scoring coreference link.

$$p(NA, she) = 0.1$$
$$p(I, she) = 0.5$$
$$p(John\ Doe, she) = 0.1$$
$$p(he, she) = 0.1$$
$$p(my, she) = 0.2$$

Coreference Model: Training Mention Ranking

- Current mention m_j should be linked to any one of the candidate antecedents it's coreferent with.

Coreference Model: Training Mention Ranking

- Current mention m_j should be linked to any one of the candidate antecedents it's coreferent with.

- The loss function will be

Maximise this probability

$$J = \sum_{i=2}^{N} -\log\left(\overbrace{\sum_{j=1}^{i-1} \mathbb{1}(y_{ij} = 1)p(m_j, m_i)}\right)$$

Iterate through candidate antecedents

for ones that are coreferent to m_j

and assign them high probability

  - Producing high probability with any one of candidate antecedents will result in overall high probability.

Coreference Model: Training Mention Ranking

- Current mention m_j should be linked to any one of the candidate antecedents it's coreferent with.

- The loss function will be

$$J = \sum_{i=2}^{N} - \log \left(\overbrace{\sum_{j=1}^{i-1} \mathbb{1}(y_{ij} = 1) p(m_j, m_i)}^{\text{Maximise this probability}} \right)$$

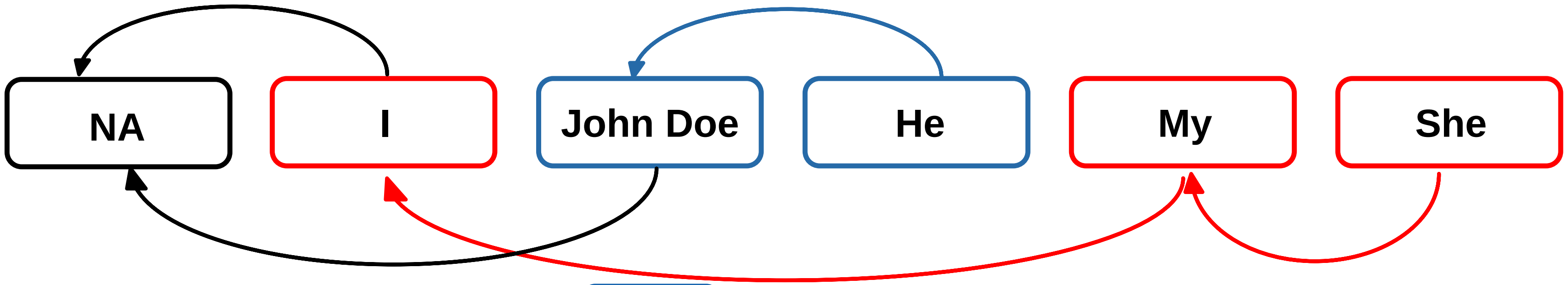
Iterate through candidate antecedents

for ones that are coreferent to m_j

and assign them high probability

  - Producing high probability with any one of candidate antecedents will result in overall high probability.

- During Test



Coreference Model: Clustering Based

- Use a clustering algorithm from the beginning to clusters all detected mentions.
    - Each cluster represents an entity.

Coreference Model: Clustering Based

- Use a clustering algorithm from the beginning to clusters all detected mentions.
  - Each cluster represents an entity.
- Bottom-up agglomerative hierarchical clustering is the most suitable approach.
  - Start with each mentions in its own singleton cluster.

Coreference Model: Clustering Based

- Use a clustering algorithm from the beginning to clusters all detected mentions.
    - Each cluster represents an entity.
- Bottom-up agglomerative hierarchical clustering is the most suitable approach.
    - Start with each mentions in its own singleton cluster.
- Merge a pair of clusters to form bigger clusters.
    - Merge on the basis of what criterion?

Coreference Model: Clustering Based

- Use a clustering algorithm from the beginning to clusters all detected mentions.
  - Each cluster represents an entity.
- Bottom-up agglomerative hierarchical clustering is the most suitable approach.
  - Start with each mentions in its own singleton cluster.
- Merge a pair of clusters to form bigger clusters.
  - Merge on the basis of what criterion?
    - Euclidian Distance
    - Cosine Similarity
    - Neural Networks

Clark, Kevin, and Christopher D. Manning. "Improving coreference resolution by learning entity-level distributed representations." arXiv preprint arXiv:1606.01323 (2016).

Clustering-based coreference models make easy decisions first

Google recently, …, **the company** announced **Google Plus**, …, **the product** features …

Clustering-based coreference models make easy decisions first

Google recently, …, **the company** announced **Google Plus**, …, **the product** features …

Google

The company

Google Plus

The product

Clustering-based coreference models make easy decisions first

Google recently, …, **the company** announced **Google Plus**, …, **the product** features …

Google

The company

Google Plus

The product

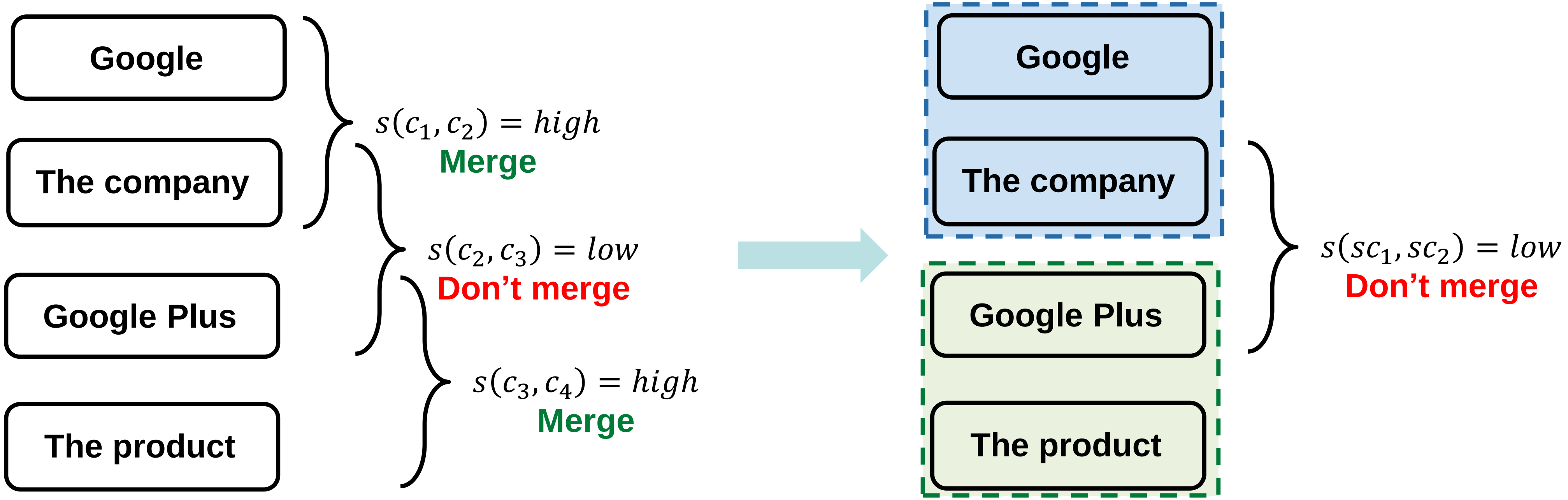
$s(c_1, c_2) = high$
Merge

$s(c_2, c_3) = low$
Don't merge

$s(c_3, c_4) = high$
Merge

Clustering-based coreference models make easy decisions first

Google recently, …, **the company** announced **Google Plus**, …, **the product** features …



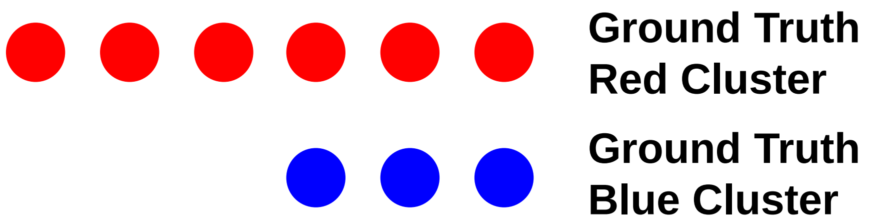
Many evaluation metrics exist for coreference resolution.

- MUC, CEAF, LEA, B-CUBED, BLANC are some of the metrics.

- Often, average over some of the metrics is reported.

Many evaluation metrics exist for coreference resolution.

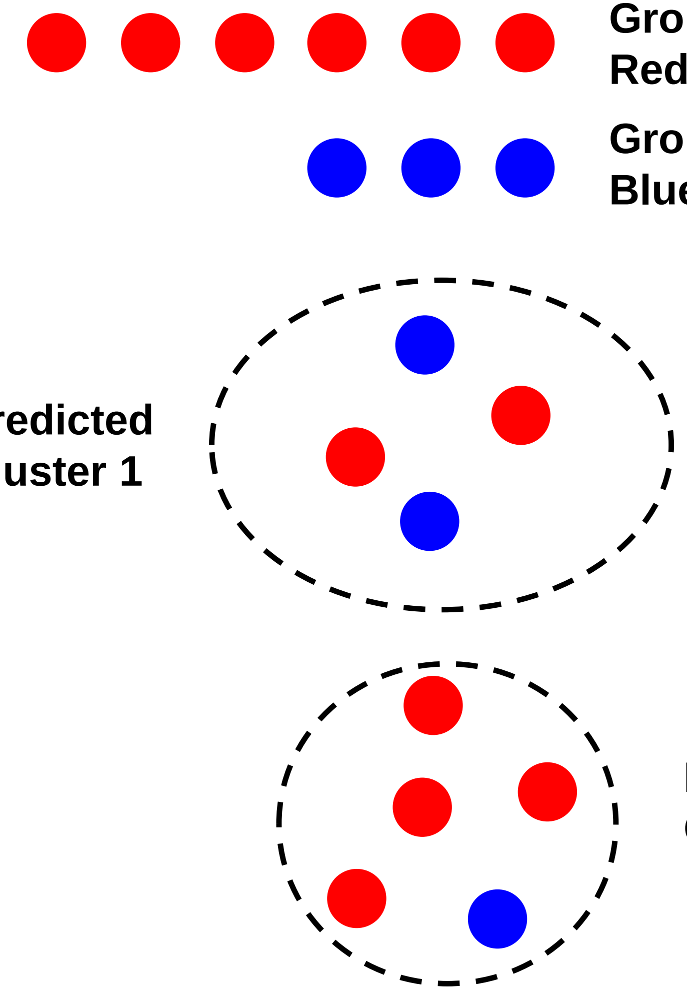
- MUC, CEAF, LEA, B-CUBED, BLANC are some of the metrics.
- Often, average over some of the metrics is reported.
- B-CUBED algorithms calculates weighted average precision and recall.



Bagga, Amit, and Breck Baldwin. "Entity-based cross-document coreferencing using the vector space model." COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics. 1998.

Many evaluation metrics exist for coreference resolution.

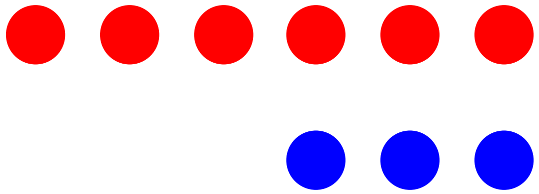
- MUC, CEAF, LEA, B-CUBED, BLANC are some of the metrics.
- Often, average over some of the metrics is reported.
- B-CUBED algorithms calculates weighted average precision and recall.



Bagga, Amit, and Breck Baldwin. "Entity-based cross-document coreferencing using the vector space model." COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics. 1998.

Many evaluation metrics exist for coreference resolution.

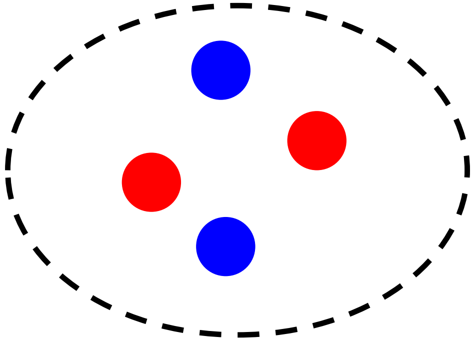
- MUC, CEAF, LEA, B-CUBED, BLANC are some of the metrics.
- Often, average over some of the metrics is reported.
- B-CUBED algorithms calculates weighted average precision and recall.



Ground Truth Red Cluster

Ground Truth Blue Cluster

Predicted Cluster 1



$$Prec_{Red} = \frac{2}{4}$$

$$Recall_{Red} = \frac{2}{6}$$

$$Prec_{Blue} = \frac{2}{4}$$

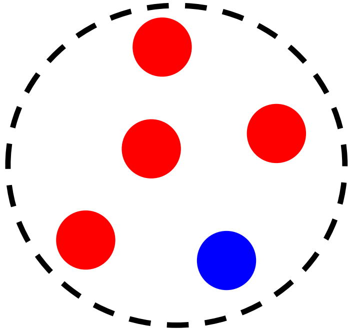
$$Recall_{Blue} = \frac{2}{3}$$

$$Prec_{Red} = \frac{4}{5}$$

$$Recall_{Red} = \frac{4}{6}$$

$$Prec_{Blue} = \frac{1}{5}$$

$$Recall_{Blue} = \frac{1}{3}$$



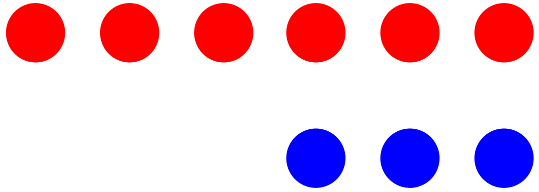
Predicted Cluster 2

Bagga, Amit, and Breck Baldwin. "Entity-based cross-document coreferencing using the vector space model." COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics. 1998.

Many evaluation metrics exist for coreference resolution.

- MUC, CEAF, LEA, B-CUBED, BLANC are some of the metrics.
- Often, average over some of the metrics is reported.
- B-CUBED algorithms calculates weighted average precision and recall.

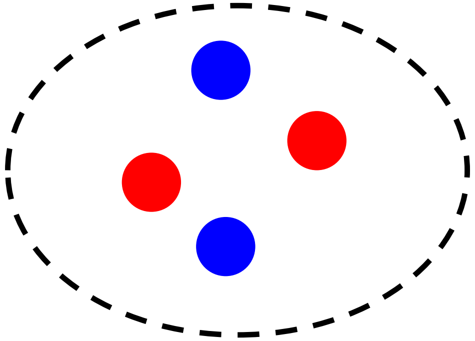
$$Weighted\ Precision = \frac{\left[4\frac{4}{5} + 1\frac{1}{5} + 2\frac{2}{4} + 2\frac{2}{4}\right]}{9} = 0.6$$



Ground Truth Red Cluster

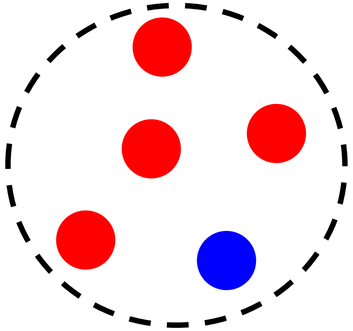
Ground Truth Blue Cluster

Predicted Cluster 1



$$Prec_{Red} = \frac{2}{4}$$
$$Recall_{Red} = \frac{2}{6}$$
$$Prec_{Blue} = \frac{2}{4}$$
$$Recall_{Blue} = \frac{2}{3}$$

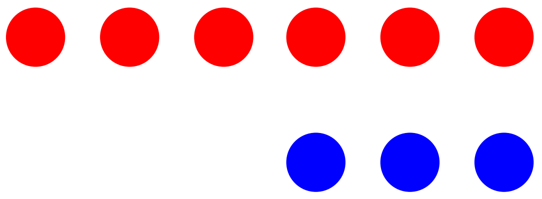
$$Prec_{Red} = \frac{4}{5}$$
$$Recall_{Red} = \frac{4}{6}$$
$$Prec_{Blue} = \frac{1}{5}$$
$$Recall_{Blue} = \frac{1}{3}$$



Predicted Cluster 2

Bagga, Amit, and Breck Baldwin. "Entity-based cross-document coreferencing using the vector space model." COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics. 1998.

Ideally a predicted cluster should be complete and homogenous

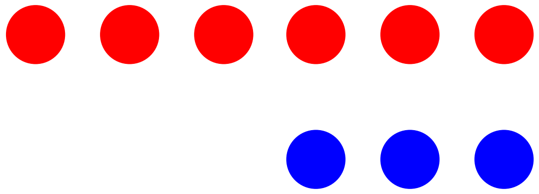


Ground Truth Red Cluster

Ground Truth Blue Cluster

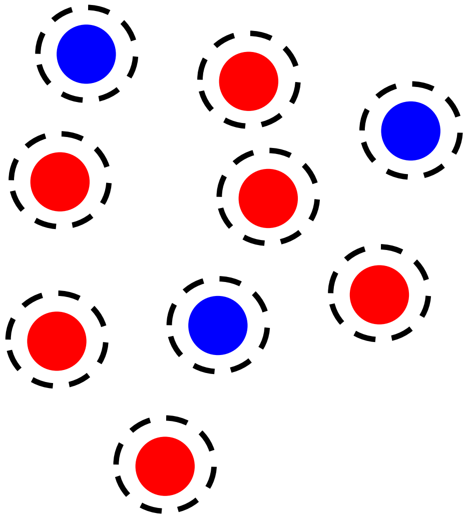
Ideally a predicted cluster should be complete and homogenous

- **Homogeneity:** A predicted cluster is homogenous if it all of its members belong to one ground truth cluster.



Ground Truth Red Cluster

Ground Truth Blue Cluster



$$Prec_{Blue} = \frac{1}{1}$$

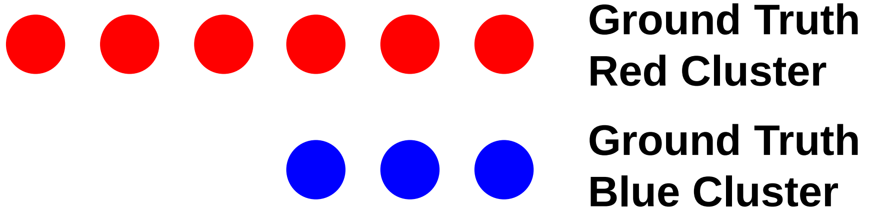
$$Recall_{Blue} = \frac{1}{3}$$

School of Electrical Engineering & Computer Science

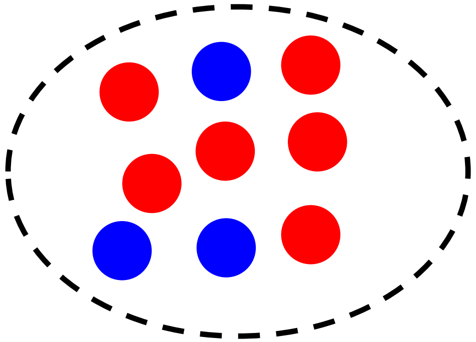
Ideally a predicted cluster should be complete and homogenous

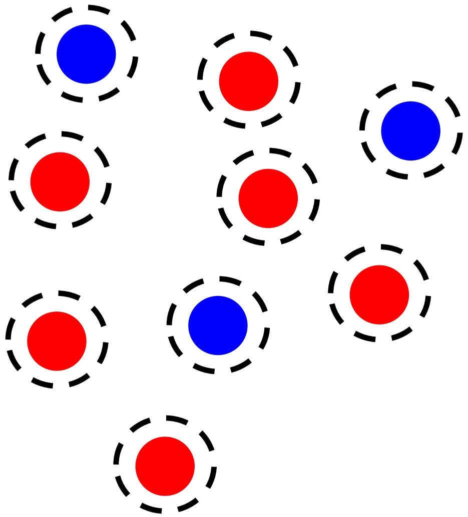
- **Homogeneity:** A predicted cluster is homogenous if it all of its members belong to one ground truth cluster.

- **Completeness:** A predicted cluster is complete if it contains all members of a ground truth cluster.



$$Prec_{Blue} = \frac{3}{9}$$
$$Recall_{Blue} = \frac{3}{3}$$





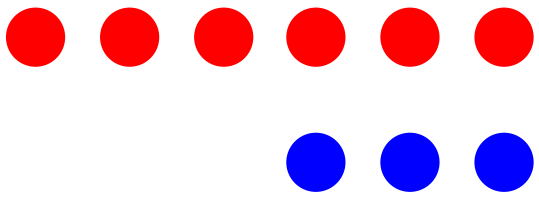
$$Prec_{Blue} = \frac{1}{1}$$
$$Recall_{Blue} = \frac{1}{3}$$

Ideally a predicted cluster should be complete and homogenous

- **Homogeneity:** A predicted cluster is homogenous if it all of its members belong to one ground truth cluster.

- **Completeness:** A predicted cluster is complete if it contains all members of a ground truth cluster.

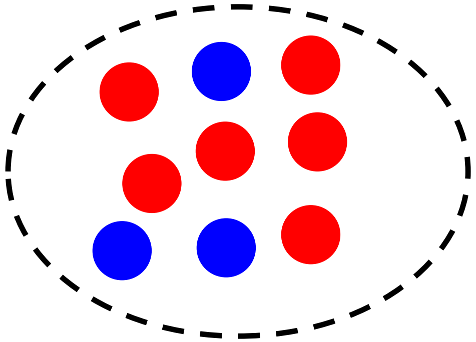
- Balance between homogeneity and completeness of predicted clusters should be maintained.

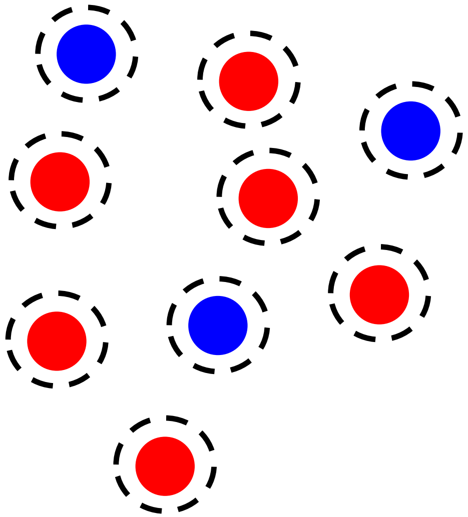


Ground Truth Red Cluster

Ground Truth Blue Cluster

$$Prec_{Blue} = \frac{3}{9}$$
$$Recall_{Blue} = \frac{3}{3}$$





$$Prec_{Blue} = \frac{1}{1}$$
$$Recall_{Blue} = \frac{1}{3}$$

Do you have any problem?

Some material (images, tables, text etc.) in this presentation has been borrowed from different books, lecture notes, and the web. The original contents solely belong to their owners, and are used in this presentation only for clarifying various educational concepts. Any copyright infringement is ***not at all*** intended.