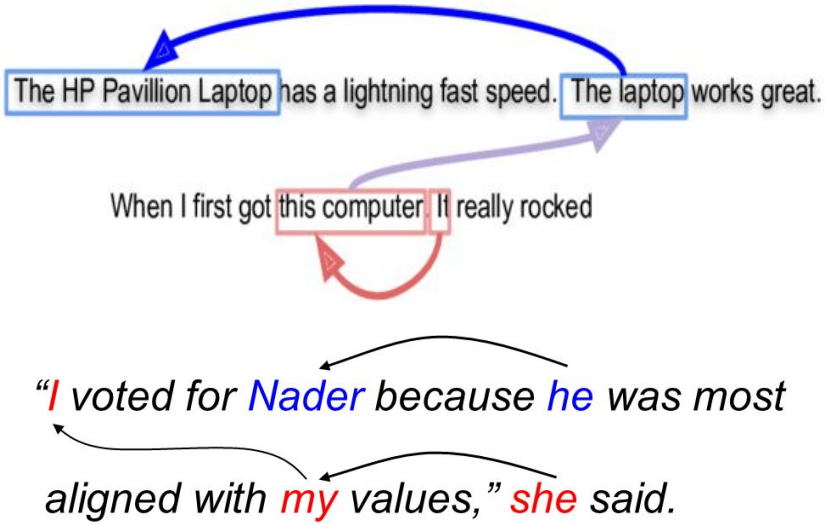
Natural Language Processing (CS5803)

Lecture 10
(Coreference Resolution)

Reference Resolution

- **Reference**: the process by which speakers use expressions to denote an entity.
- **Referring expression**: expression used to perform reference .
- **Referent**: the entity that is referred to.
- **Coreference**: referring expressions that are used to refer to the same entity.
- **Anaphora**: reference to a previously introduced entity.



Examples of References

mention	**John** told **Sally** that **she** should come watch **him** play the **violin**.
antecedent	John told **Sally** that **she** should come watch him play the violin.
coreferent	John told **Sally** that **she** should come watch him play the violin.
cluster	**John** told **Sally** that **she** should come watch **him** play the **violin**.
anaphoric	John told **Sally** that **she** should come watch him play the violin.
non-anaphoric	ϵ John told Sally that she should come watch him play the **violin**.

Types of References

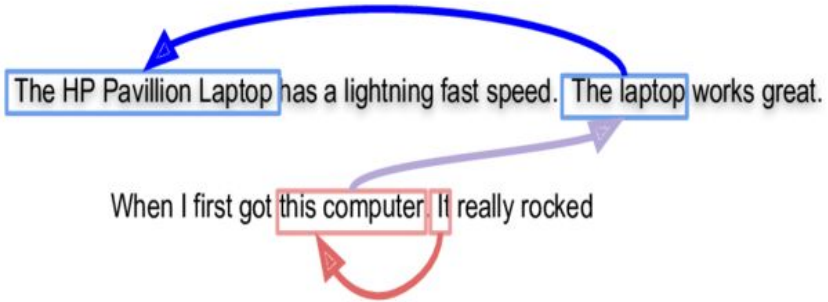
Five common types of referring expression	
Type	**Example**
Indefinite noun phrase	I saw **a** Ford Escort today.
Definite noun phrase	I saw a Ford Escort today. **The** Escort was white.
Pronoun	I saw a Ford Escort today. **It** was white.
Demonstratives	I like **this** better than **that**.
One-anaphora	I saw 6 Ford Escort today. Now I want **one**.
Three types of referring expression that complicate the reference resolution	
Type	**Example**
Inferrables	I almost bought a Ford Escort, but a **door** had a dent.
Discontinuous Sets	John and Mary love their Escorts. **They** often drive them.
Generics	I saw 6 Ford Escorts today. **They** are the coolest cars.

Coreference Chain

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling by a plane.

Coreference chains

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer}



Coreference Chain

Toni Johnson pulls a tape measure across the front of what was once a stately Victorian home. A deep trench now runs along its north wall, exposed when the house lurched two feet off its foundation during last week's earthquake.

Once inside, she spends nearly four hours measuring and diagramming each room in the 80-year-old house, gathering enough information to estimate what it would cost to rebuild it.

While she works inside, a tenant returns with several friends to collect furniture and clothing.

One of the friends sweeps broken dishes and shattered glass from a countertop and starts to pack what can be salvaged from the kitchen.

Anaphora

Exophora: (extra linguistic feature: deitic markers, this that)

"what is this?"

Anaphora:

"I used to have the key. But I lost it."

Cataphora:

"It is your turn, John"

- The Process of finding the antecedent for an Anaphor is Anaphora resolution
  - Anaphor-The reference that point to the previous item.
  - Antecedent-The entity to which the anaphor refers

Constraints

- **Number Agreement**:
  - To distinguish between singular and plural references.
    - *John has a new car. They are red.

- **Gender Agreement**:
  - To distinguish male, female, and non-personal genders.
    - John has a new car. It is attractive. [It = the new car]

- **Person and Case Agreement**:
  - To distinguish between three forms of person;
    - *You and I have Escorts. They love them.
  - To distinguish between subject position, object position, and genitive position.

Constraints

- Syntactic Constraints:
  - Syntactic relationships between a referring expression and a possible antecedent noun phrase
    - John bought himself a new car. [himself=John]
    - John bought him a new car. [him≠John]

- Selectional Restrictions:
  - A verb places restrictions on its arguments.
    - John parked his Acura in the garage. He had driven it around for hours. [it=Acura, it≠garage];
    - I picked up the book and sat in a chair. It broke.

Preferences in Pronoun Interpretation

- **Recency**:
  - Entities introduced recently are more salient than those introduced before.
    - John has a Legend. Bill has an Escort. Mary likes to drive it.

- **Grammatical Role**:
  - Entities mentioned in subject position are more salient than those in object position.
    - Bill went to the Acura dealership with John. He bought an Escort. [he=Bill]

- **Repeated Mention**:
  - Entities that have been focused on in the prior discourse are more salient.
    - John needed a car to get to his new job. He decided that he wanted something sporty. Bill went to the Acura dealership with him. He bought an Integra. [he=John]

Preferences in Pronoun Interpretation

- **Verb Semantics**:
  - Certain verbs appear to place a semantically-oriented emphasis on one of their argument positions.

    - John telephoned Bill. He had lost the book in the mall. [He = John]
    - John criticized Bill. He had lost the book in the mall. [He = Bill]

    - David praised Hans because he … [he = Hans]
    - David apologized to Hans because he… [he = David]

- **World knowledge in general**:

  - The city council denied the demonstrators a permit because they {feared|advocated} violence.
  - **The city council** denied the demonstrators a permit because **they** {**feared**|advocated} violence.
  - The city council denied **the demonstrators** a permit because **they** {feared|**advocated**} violence.

Anaphora Resolution

- Anaphora resolution
  - Determine the antecedent of an anaphor
- Coreference resolution
  - Identify all coreference chains
- Anaphora in case of coreference chains
  - If any of the preceding entries in the coreference chains is detected as the antecedent, the answer is considered correct

Coreference chain resolution

<Coref_chain Id="1">Eagle Clothes Inc.</Coref_chainId="1"> , which is operating under Chapter 11 of the federal Bankruptcy Code, said <Coref_chain Id="1">it</Coref_chain Id="1"> reached an agreement with <Coref_chain Id="1">its</Coref_chain Id="1"> creditors.

Under the accord , <Coref_chain Id="2">Albert Roth, chairman and chief executive officer</Coref_chain Id="2"> , and Arthur Chase, Sam Beigel , and Louis Polsky will resign as officers and directors of the <Coref_chain Id="1">menswear retailer</Coref_chain Id="1">.

<Coref_chain Id="2">Mr. Roth, who has been on leave from his posts ,</Coref_chain Id="2"> will be succeeded by <Coref_chain Id="3">Geoffrie D. Lurie of GDL Management Inc. , which is <Coref_chain Id="1">Eagle 's</Coref_chain Id="1"> crisis manager </Coref_chain Id="3">. <Coref_chain Id="3">Mr. Lurie </Coref_chain Id="3"> is currently co-chief executive .

Datasets

- CoNLL 2012
  - News articles, conversational logs, web data, Magazines
- LitBank
  - Novels
- ARRAU Corpus
  - Dialog, Fictions

Winograd Schema Challenge

- Contains
  - A pair of sentences that differ in a single word or phrase.
  - A coreference question

- **Example 1**
  - **Sentences:** The trophy didn't fit into the suitcase because it was too large.
  - **Question:** What was too large? Answer: The trophy
- **Example 2:**
  - **Sentences:** The trophy didn't fit into the suitcase because it was too small.
  - **Question:** What was too small? Answer: The suitcase
- **Example 3/4:**
  - **Sentences:** Bill passed the gameboy to John because his turn was [over/next].
  - **Question:** Whose turn was [over/next]? Answers: Bill/John

- Getting correct answers requires commonsense reasoning ability

Evaluation Measures

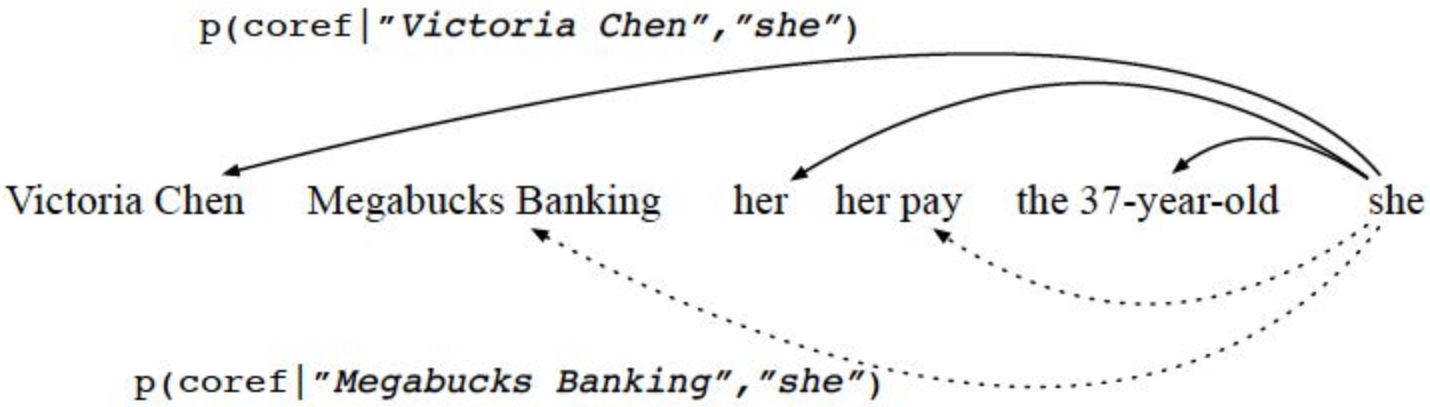
- MUC
  - Link-based
- B^3
  - Mention-based

- Often average of metric values are used

Features

- For word pair (i,j) predict if it is the referred object
- Sent-Distance, Mark-Distance i-Pronoun, j-Pronoun, Str-match, DEF-NP, DEM-NP, Number-match, Gender-match, etc.
- Pass to a binary classifier
- Filtering out of non-anaphoric pronoun tends to improve results

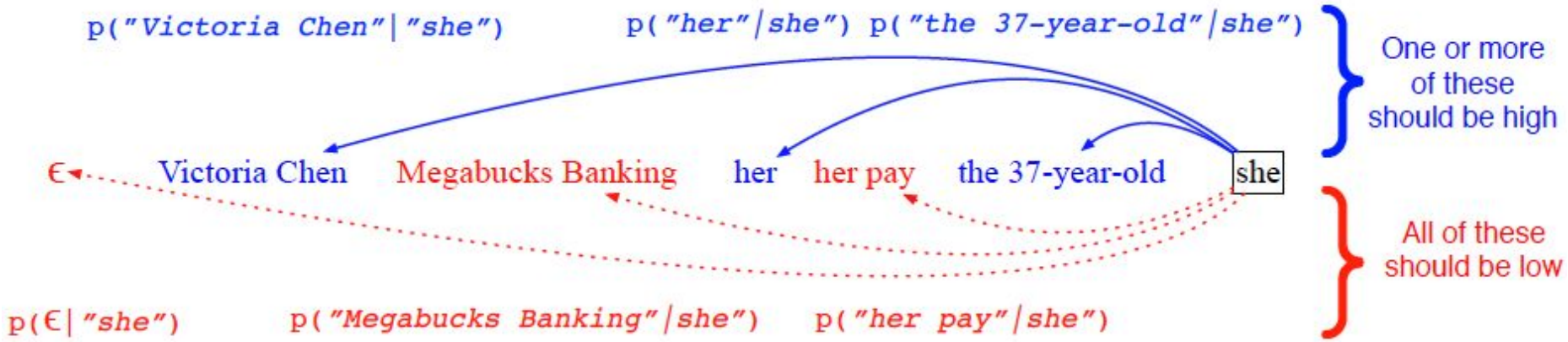
Mention-pair architecture



For a pair of mentions, predicts whether it is a coreference.

Modeled as a classifier, needs to be trained.

Mention-pair architecture

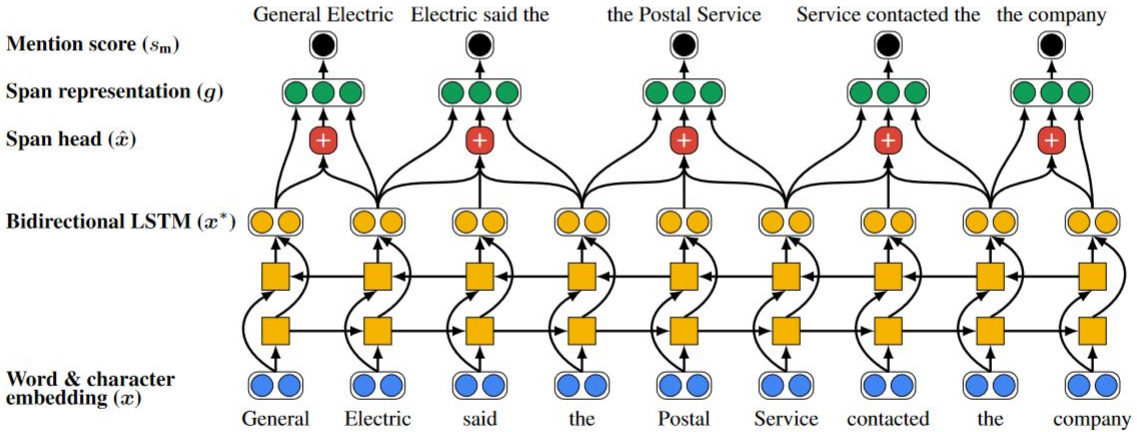


For pairs of mentions, ranks them based on how likely they are coreferences.

End-to-end Neural Coreference Resolution

- Text: $x_1, x_2, \ldots, x_T$
- Consider all $\frac{T(T-1)}{2}$ possible spans
  - Often length restriction is imposed
- Order the spans in lexicographic manner
- i^{th} span y_i ranges from $START(i)$ to $END(i)$
- For y_i, possible antecedents are from $\{\epsilon, 1, 2, .., i-1\}$
- ϵ is a dummy span which is output when
  - y_i is not an entity mention, or
  - y_i is a span but is not co-referent with any previous span

End-to-end Neural Coreference Resolution



- For spans y_i and y_j with $i \neq j$
  - $s(i,j) = s_m(i) + s_m(j) + s_a(i,j)$
- Score is 0 for the dummy antecedent ϵ

$$\alpha_t = \boldsymbol{w}_\alpha \cdot \text{FFNN}_\alpha(\boldsymbol{x}_t^*)$$

$$a_{i,t} = \frac{\exp(\alpha_t)}{\sum_{k=\text{START}(i)}^{\text{END}(i)} \exp(\alpha_k)}$$

$$\hat{\boldsymbol{x}}_i = \sum_{t=\text{START}(i)}^{\text{END}(i)} a_{i,t} \cdot \boldsymbol{x}_t$$

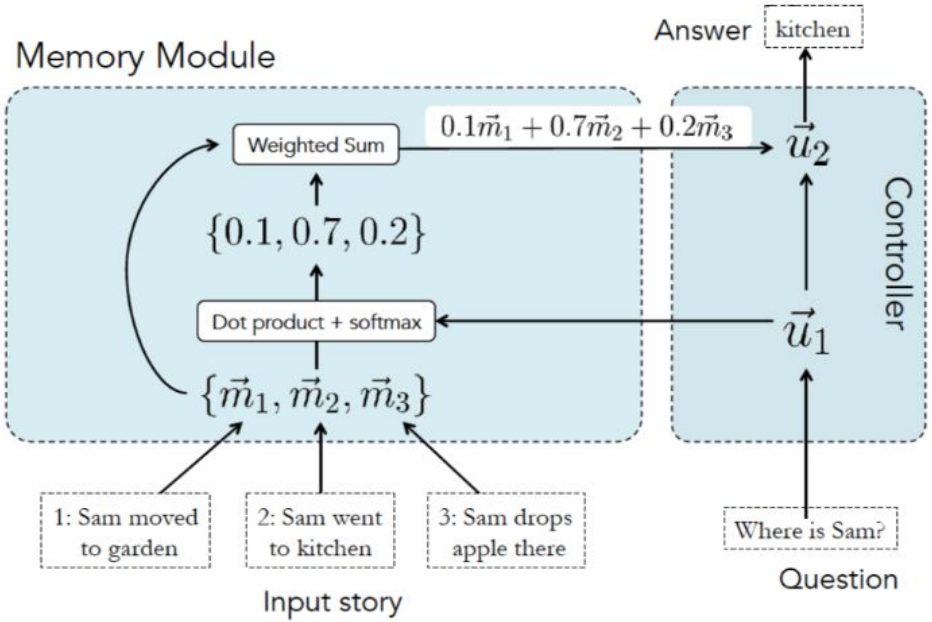
- $s_m(i) = FFFN(g_i)$
- $s_a(i,j) = w_a \cdot FFNN_a([g_i, g_j, g_i \circ gj, \Phi(i,j)]$
- $g_i = [x^*_{START(i)}, x^*_{END(i)}, \hat{x_i}, \Phi(i)]$

Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. 2017. End-to-end Neural Coreference Resolution. EMNLP 2017

Coreference resolutions using QA

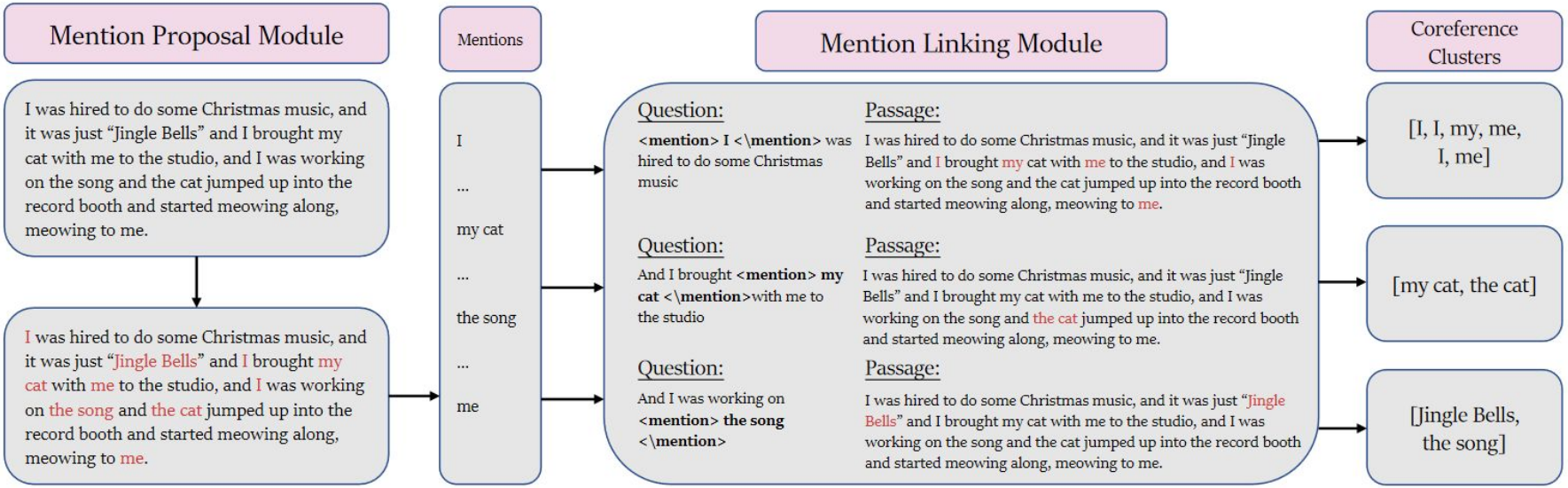
- Lack of proper large scale datasets
  - Linguistic overhead in preparing them
- Can be posed as QA
- "Reasoning (such as synonymy, world knowledge, syntactic variation, and multiple sentence reasoning) required to answer the questions are also indispensable for coreference resolution"
- Existing QA frameworks evaluated on large scale datasets can be leveraged

Conference Resolution using memory network



- Use memory network to get o_k
- $u_{k+1} = u_k H + o_k$
- $A = softmax\big(W(u_L + o_k)\big)$
- $A' = softmax\big(W(u_L + o_k)\big)$

Coreference Resolution as Query-based Span Prediction



- $s_m(i) = FFFN\big(\Phi(b_i); \Phi(e_i)\big)$
- B-I-O tagging for sentence tokens using $p_i^{tag} = FFFN_{tag}(\Phi(P;Q))$

- $s_a(i,j) = \frac{1}{|e_j|}\left[\log p^B_{b(i)} + \sum_{k=b_i+1}^{e_i} \log p^I_k\right]$
- $s(i,j) = s_m(i) + s_a(i,j) + s_a(j,i)$
- Multiple runtime optimization decisions

References

- A Dataset of Literary Coreference
- A generalized framework for anaphora resolution in Indian languages
- A Ranking Approach to Pronoun Resolution
- A Machine Learning Approach to Coreference Resolution of Noun Phrases
- Towards Harnessing Memory Networks for Coreference Resolution
- End-to-end Neural Coreference Resolution
- Bridging Anaphora Resolution as Question Answering
- CorefQA: Coreference Resolution as Query-based Span Prediction
- A Neural Entity Coreference Resolution review