

Coreference Resolution

Chapter 22

Is this text coherent?

“Consider, for example, the difference between passages (18.71) and (18.72). Almost certainly not. The reason is that these utterances, when juxtaposed, will not exhibit coherence. Do you have a discourse? Assume that you have collected an arbitrary set of well-formed and independently interpretable utterances, for instance, by randomly selecting one sentence from each of the previous chapters of this book.”

Or, this?

“Assume that you have collected an arbitrary set of well-formed and independently interpretable utterances, for instance, by randomly selecting one sentence from each of the previous chapters of this book. Do you have a discourse? Almost certainly not. The reason is that these utterances, when juxtaposed, will not exhibit coherence. Consider, for example, the difference between passages (18.71) and (18.72).”

What makes a text coherent?

- Discourse structure
 - In a coherent text the parts of the discourse exhibit a sensible ordering and hierarchical relationship
- Rhetorical structure
 - The elements in a coherent text are related via meaningful relations (“coherence relations”)
- Entity structure (“Focus”)
 - A coherent text is **about** some entity or entities, and the entity/entities is/are referred to in a structured way throughout the text.
 - *Appropriate use of referring expressions*

Pronouns and Reference Resolution

A Reference Joke

Gracie: Oh yeah ... and then Mr. and Mrs. Jones were having matrimonial trouble, and my brother was hired to watch Mrs. Jones.

George: Well, I imagine **she** was a very attractive woman.

Gracie: **She** was, and my brother watched **her** day and night for six months.

George: Well, what happened?

Gracie: **She** finally got a divorce.

George: Mrs. Jones?

Gracie: No, my brother's wife.

Coreference Resolution Examples

(from Jacob Eisenstein)

- Apple Inc Chief Executive Tim Cook has jetted into China for talks with government officials as he seeks to clear up a pile of problems in the firm's biggest growth market... Cook is on his first trip to the country since taking over...

Heuristics: recency

- The doctor found an old map in the captain's chest.
- Jim found an even older map hidden on the shelf.
- It described an island.

Heuristics: prominence

- Asha loaned Mei a book on Spanish.
- She is always trying to help people.

Heuristics: parallelism

- Asha loaned Mei a book on Spanish.
- Olya loaned her a book on Portuguese.

Beyond heuristics

- The city council denied the protesters a permit because they feared violence.
- The city council denied the protesters a permit because they advocated violence.

Pronouns that do not refer to entities

- They told me I was too ugly for show business, but I didn't believe **it**.
- Asha saw Babak get angry, and I saw **it** too.
- Asha said she worked in security. **That's** one way to put **it**.
- Every farmer who owns a donkey beats **it**.
- **It's** too bad we have to work so hard.

Names

- Match on heads?
- [Apple Inc Chief Executive [Tim Cook]] ... [Cook] is on his first trip ...

Nominals

Apple Inc Chief Executive Tim Cook has jetted into China for talks with government officials as he seeks to clear up a pile of problems in the firm's biggest growth market... Cook is on his first trip to the country since taking over...

- *Talks, the firm, the firm's biggest growth market, the country*
- These seem to require world knowledge to address.

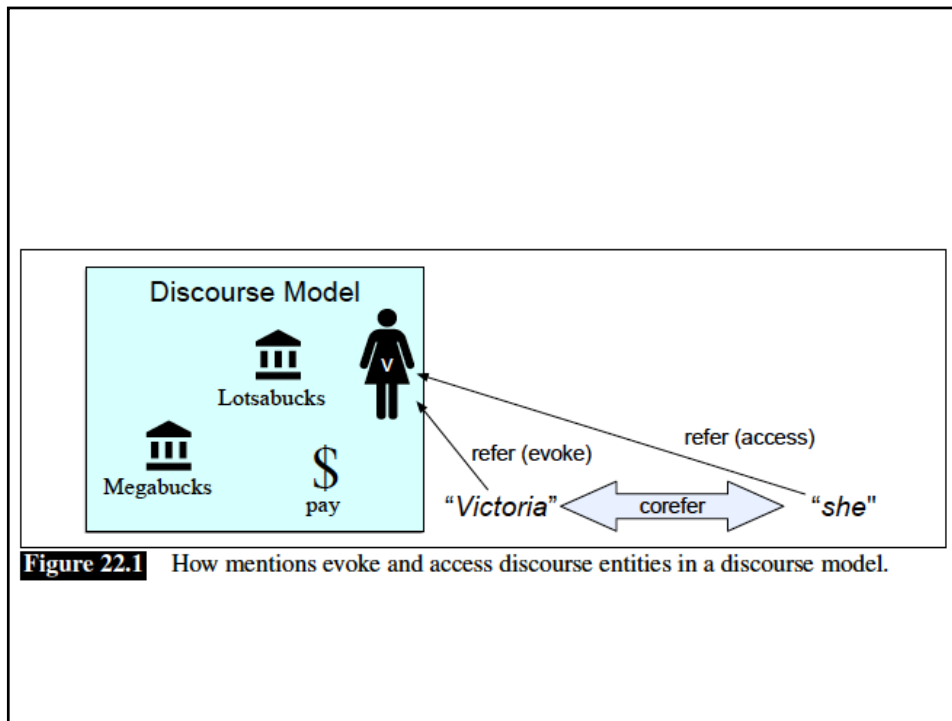
Reference Resolution: Vocabulary

- Process of associating **Bloomberg/he/his** with particular person and **big budget problem/it** with a concept
Guiliani left **Bloomberg** as mayor of a city with a **big budget problem**. *It's unclear how he'll be able to handle it during his term.*
- Referring exprs./mentions: **Guilani, Bloomberg, a big budget problem, he, it, his**
- **Presentational it:** non-referential
- Referents: the **person named Guiliani**, the concept of a **big budget problem**

- Co-referring referring expressions: Bloomberg, he, his
- Antecedent: Bloomberg
- Anaphors: he, his

Discourse Models

- Needed to model reference because referring expressions (e.g. Giuliani, Bloomberg, he, it, budget problem) encode information about beliefs about the referent
- When a referent is first mentioned in a discourse, a representation is evoked in the model
 - Information predicated of it is stored also in the model
 - On subsequent mention, it is accessed from the model



Types of Referring Expressions

- Entities, concepts, places, propositions, events, ...
 - According to John, Bob bought Sue an Integra, and Sue bought Fred a Legend.
 - But **that** turned out to be a lie. (a speech act)
 - But **that** was false. (proposition)
 - **That** struck me as a funny way to describe the situation. (manner of description)
 - **That** caused Sue to become rather poor. (event)
 - **That** caused them both to become rather poor. (combination of multiple events)

Reference Phenomena:

5 Types of Referring Expressions

- Indefinite NPs

A homeless man hit up Bloomberg for a dollar.

Some homeless guy hit up Bloomberg for a dollar.

This homeless man hit up Bloomberg for a dollar.

- Definite NPs

The poor fellow only got a lecture.

- Demonstratives

This homeless man got a lecture but that one got carted off to jail.

- Names

Prof. Litman teaches on Tuesday.

Pronouns

A large tiger escaped from the Central Park zoo chasing a tiny sparrow. It was recaptured by a brave policeman.

- Referents of pronouns usually require some degree of *salience* in the discourse (as opposed to definite and indefinite NPs, e.g.)
- How do items become *salient* in discourse?

Saliency vs. Recency

E: So you have the engine assembly finished. Now attach the rope. By the way, did you buy the gas can today?

A: Yes.

E: Did it cost much?

A: No.

E: OK, good. Have you got it attached yet?

Reference Phenomena: Information Status

- Givenness hierarchy / accessibility scales ...
- But complications

Inferables

- I almost bought an Acura Integra today, but a door had a dent and the engine seemed noisy.
- Mix the flour, butter, and water. Knead the dough until smooth and shiny.

Discontinuous Sets

- Entities evoked together but mentioned in different sentence or phrases

John has a St. Bernard and Mary has a Yorkie. They arouse some comment when they walk them in the park.

Generics

I saw two Corgis and their seven puppies today. They are the funniest dogs

Constraints on Pronominal Reference

- Number agreement
John's parents like opera. John hates it/John hates them.
- Person agreement
George and Edward brought bread. They shared it.

- Gender agreement

John has a Porsche. He/it/she is attractive.

- Syntactic constraints

John bought himself a new Volvo. (himself = John)

John bought him a new Volvo (him = not John).

Preferences in Pronoun Interpretation

- Recency

John bought a new boat. Bill bought a bigger one.
Mary likes to sail it.

- But...grammatical role raises its ugly head...

John went to the Acura dealership with Bill. He bought an Integra.

- And so does...repeated mention
 - John needed a car to go to his new job. He decided that he wanted something sporty. Bill went to the dealership with him. He bought a Miata.
 - Who bought the Miata?
 - What about grammatical role preference?
- Parallel constructions
 - Saturday, Mary went with Sue to the farmer's market.
 - Sally went with her to the bookstore.
 - Sunday, Mary went with Sue to the mall.
 - Sally told her she should get over her shopping obsession.
- Selectional restriction
 - John left his plane in the hangar.
 - He had flown it from Memphis this morning

- Verb semantics/thematic roles
 - John telephoned Bill. He'd lost the directions to his house.
 - John criticized Bill. He'd lost the directions to his house.

Summary: What Affects Reference Resolution?

- Lexical factors
 - Reference type: Inferrability, discontinuous set, generics, one anaphora, pronouns,...
- Discourse factors:
 - Recency
 - Focus/topic structure, digression
 - Repeated mention
- Syntactic factors:
 - Agreement: gender, number, person, case
 - Parallel construction
 - Grammatical role
- Semantic/lexical factors
 - Selectional restrictions
 - Verb semantics, thematic role

Reference Resolution Algorithms

- Given these types of features, can we construct an algorithm that will apply them such that we can identify the correct referents of anaphors and other referring expressions?

Reference Resolution Task

- Finding in a text all the referring expressions that have one and the same denotation
 - Pronominal anaphora resolution
 - Anaphora resolution between named entities
 - Full noun phrase anaphora resolution

Issues

- Which constraints/features can/should we make use of?
- How should we order them? I.e. which override which?
- What should be stored in our discourse model? I.e., what types of information do we need to keep track of?
- How to evaluate?

Coreference Resolution

- Input: text
- Output: all entities (via mention detection) and the coreference links between them (create clusters)

(22.43) [Victoria Chen]_a¹, CFO of [Megabucks Banking]_a², saw [[her]_b¹ pay]_a³ jump to \$2.3 million, as [the 38-year-old]_c¹ also became [[the company]_b²'s president. It is widely known that [she]_d¹ came to [Megabucks]_c² from rival [Lotsabucks]_a⁴.

1. { *Victoria Chen, her, the 38-year-old, She* }
2. { *Megabucks Banking, the company, Megabucks* }
3. { *her pay* }
4. { *Lotsabucks* }

Mention Detection

- Finding the spans of text
- Usually liberal (focus on recall), e.g.,
 - run parser and NER and return all spans that are NPs, possessive pronouns, or named entity
 - Or, N-grams up to some value of N
- Followed by filtering, e.g.,
 - rules
 - learned classifiers using hand-labeled datasets
 - end-to-end via neural methods

OntoNotes

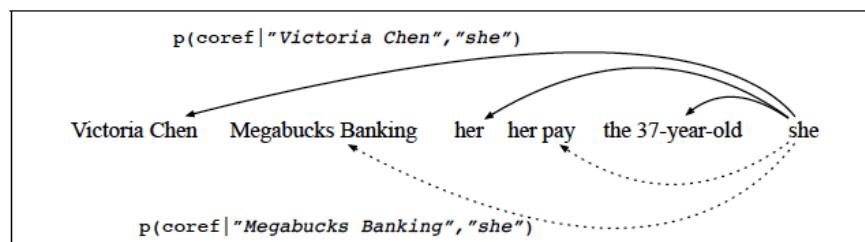
- A popular hand-labeled coreference dataset
 - 1 million words each of English and Chinese, plus less Arabic
 - Does not label singletons (an entity that has only a single mention, e.g. cluster with only one member; majority of data)

Architectures for Coreference Algorithms

- Mention-based
 - Consider each mention independently, then rank
- Entity-based

Mention-Pair Architecture

- Input: A candidate anaphor and a candidate antecedent
- Output: Probabilistic binary decision about coreference



Approaches

- Machine learning supervised classifiers
- Need a heuristic for sampling training examples due to class imbalance
 - E.g., choose pair with closest antecedent as positive example, and all intervening pairs as negative examples

Evaluation

- Task-dependent metrics have been developed, e.g.

$$\text{Precision} = \sum_{i=1}^N w_i \frac{\# \text{ of correct mentions in hypothesis chain containing entity}_i}{\# \text{ of mentions in hypothesis chain containing entity}_i}$$
$$\text{Recall} = \sum_{i=1}^N w_i \frac{\# \text{ of correct mentions in hypothesis chain containing entity}_i}{\# \text{ of mentions in reference chain containing entity}_i}$$

Other Comments

- Entity Linking – link mentions to Wikipedia, gazeteer, etc.
- Difficult evaluation sets
 - Winograd schema (e.g., feared vs advocating violence example) has examples more likely to require world knowledge and reasoning methods from AI
- Detecting and mitigating bias (e.g., gender)