

Topics for Today

- Pragmatics of discourse context
 - reference resolution
 - noun phrase coreference resolution
 - machine learning approach to NP coreference resolution

The problem of reference resolution

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.

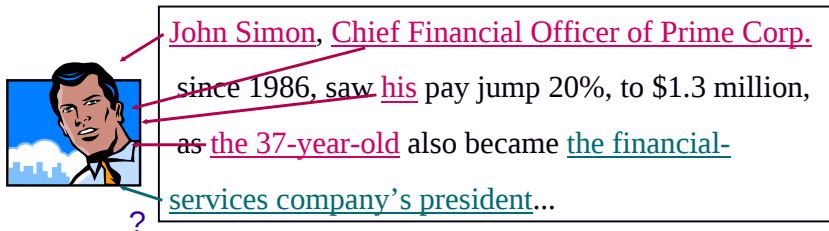
George: Mrs. Jones?

Gracie: No, my brother's wife.

George Burns and Gracie Allen in *The Salesgirl*

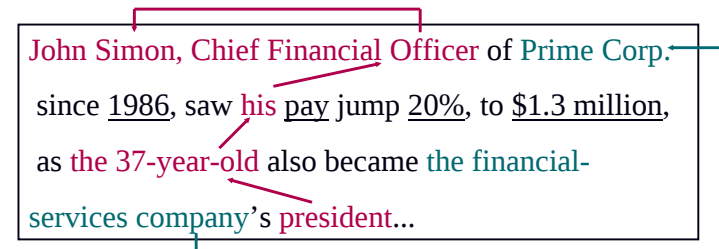
Reference resolution

- **Reference:** the process by which speakers use expressions like “John Simon” and “his” to denote a real-world entity
 - **Referring expressions:** NL expression used to perform reference
 - **Referent:** the entity that is referred to
 - **Shorthand form:** *his* refers to John Simon



Coreference

- **Coreference:** two referring expressions that are used to refer to the same entity are said to corefer
- *John Simon* is the **antecedent** of *his*.
- Reference to an entity that has been previously introduced into the discourse is called **anaphora**; and the referring expression used is said to be **anaphoric**.



Types of referring expressions

- Indefinite noun phrases
 - Introduce entities that are new to the hearer into the discourse context
 - » I saw *a Subaru WRX* today.
 - » I saw *this awesome Subaru WRX* today.
- Definite noun phrases
 - Refer to an entity that is identifiable to the hearer
 - » It has already been mentioned in the discourse
 - » It is contained in the hearer's set of beliefs about the world
 - » The uniqueness of the object is implied by the description itself
 - ◆ I saw a Subaru WRX today. *The WRX* was blue and needed a wash.
 - ◆ *The Indy 500* is the most popular car race in the US.
 - ◆ *The fastest car in the Indy 500* was a Subaru WRX.

Types of referring expressions

- Pronouns
 - Another form of definite reference
 - Referent must have a high degree of activation or **salience** in the discourse model
 - » John went to Bob's party, and parked next to a beautiful Subaru WRX. He went inside and talked to Bob for more than an hour. Bob told him that he recently got engaged.
 - (a)?? He also said that he bought *it* yesterday.
 - (a') He also said that he bought *the WRX* yesterday.
 - Cataphora: referring expression is mentioned before its referent
 - » Before *he* bought *it*, John checked over the WRX carefully.

Types of referring expressions

- Demonstrative pronouns
 - Behave somewhat differently from simple definite pronouns
 - » Can appear alone or as determiners
 - » Choice of *this* or *that* depends on some notion of spatial or temporal proximity
 - ◆ I bought a WRX yesterday. It's similar to the one I bought a year ago. *That one* was really nice, but I like *this one* even better.
- One-anaphora
 - Blends properties of definite and indefinite reference
 - » I saw no fewer than 6 Subaru WRX's today. Now I want *one*.
 - May introduce a new entity into the discourse, but it is dependent on an existing referent for the description of this new entity.

Noun Phrase Coreference Resolution

- Identify all phrases that refer to each real-world entity mentioned in the text

John Simon, Chief Financial Officer of Prime Corp. since 1986, saw *his* pay jump 20%, to \$1.3 million, as *the 37-year-old* also became *the financial-services company's president...*

Why It's Hard

Many sources of information play a role

- head noun matches
 - » IBM executives = the executives
 - » Microsoft executives
- syntactic constraints
 - » John helped himself to...
 - » John helped him to...
- discourse focus, recency, syntactic parallelism, semantic class, agreement, world knowledge, ...

Why It's Hard

No single source is a completely reliable indicator

- semantic preferences
 - » Mr. Callahan = president =? the carrier
- number and gender
 - » assassination (of Jesuit priests) = these murders
 - » the woman = she = Mary =? the chairman

Why It's Hard

Coreference strategies differ depending on the type of referring NP

- definiteness of NPs
 - » ... Then Mark saw the man walking down the street.
 - » ... Then Mark saw a man walking down the street.
- pronoun resolution alone is notoriously difficult
 - » resolution strategies differ for each type of pronoun
 - » some pronouns refer to nothing in the text

I went outside and it was snowing.

Types of referents: complications

- Inferrables
 - A referring expression does not refer to an entity in the text, but to one that is inferentially related to it.
 - » I almost bought a WRX today, but a door had a dent and the engine seemed noisy.
 - » Mix the flour, butter, and water. Stir the batter until all lumps are gone.
- Discontinuous sets
 - Referents may have been evoked in discontinuous phrases
 - » John has a Volvo, and Mary has a Mazda. They drive them all the time.
- Generics – refer to a class of entities
 - I saw no fewer than 6 WRX's today. They are the coolest cars.

Topics for today

- Pragmatics of discourse
 - reference resolution
 - noun phrase coreference resolution
- ➔ machine learning approach to NP coreference resolution
 - just the basics

Traditional Knowledge-Based Approaches

Lappin and Leass [1994]

- hand-crafted heuristics and filters
 - syntactic filters [Lappin and McCord 1990a]
 - morphological filter
 - pleonastic pronoun filter (“It was raining.”)
 - procedure for identifying possible antecedents [Lappin and McCord 1990b]
 - salience assignment w.r.t. grammatical role, proximity, parallelism, etc.
- decision procedure

Problems

- Portability
- Robustness
- Few large-scale evaluations
- Evaluations make a number of simplifying assumptions
 - perfect parse
 - omit many difficult cases, e.g. pleonastic pronouns
- **Impose coreference resolution strategies rather than learn them empirically**

A Machine Learning Approach

- Classification
 - given a description of two noun phrases, NP_i and NP_j , classify the pair as *coreferent* or *not coreferent*

[John Simon], [Chief Financial Officer] of [Prime Corp.]

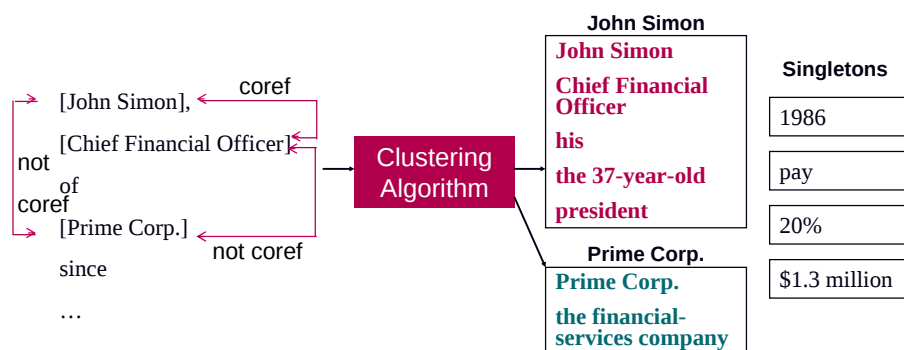


since 1986, saw his pay jump 20%, to \$1.3 million,
as the 37-year-old also became the

Aone & Bennett [1995]; Connolly et al. [1995]; McCarthy & Lehnert [1995];
Soon, Ng & Lim [2001]; Ng & Cardie [2002]

A Machine Learning Approach

- Clustering
 - coordinates pairwise coreference decisions



Issues

- Training data
- Instance representation
- Learning algorithm
- Clustering approach

Training Data Creation

- Creating training instances
 - texts annotated with coreference information
- candidate antecedent anaphor
- one instance $inst(NP_i, NP_j)$ for each *ordered* pair of NPs
- » NP_i precedes NP_j
 - » feature vector: describes the two NPs and context
 - » class value:
 - coref* pairs on the same coreference chain
 - not coref* otherwise

Instance Representation

- 25 features per instance
 - lexical (3)
 - » string matching for pronouns, proper names, common nouns
 - grammatical (18)
 - » pronoun_1, pronoun_2, demonstrative_2, indefinite_2, ...
 - » number, gender, animacy
 - » appositive, predicate nominative
 - » binding constraints, simple contra-indexing constraints, ...
 - » span, maximalnp, ...
 - semantic (2)
 - » same WordNet class
 - » alias
 - positional (1)
 - » distance between the NPs in terms of # of sentences
 - knowledge-based (1)
 - » naïve pronoun resolution algorithm

Learning Algorithm

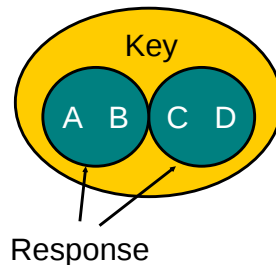
- RIPPER (Cohen, 1995)
C4.5 (Quinlan, 1994)
 - rule learners
 - » input: set of training instances
 - » output: coreference classifier
- Learned classifier
 - » input: test instance (represents pair of NPs)
 - » output: classification
confidence of classification

Clustering Algorithm

- Start with each NP in its own partition
- For each NP in the document
 - Consider each NP to its left
 - If ML algorithm says “coreferent”, merge the partitions for the two NPs.

Evaluation

- MUC-6 and MUC-7 coreference data set
- documents annotated w.r.t. coreference
- 30 + 30 training texts (dry run)
- 30 + 20 test texts (formal evaluation)
- scoring program
 - recall
 - precision
 - F-measure: $2PR/(P+R)$



Baselines...

	MUC-6		
	R	P	F
Match Any Word			41.3
Match Head Word			45.7
Single Cluster	93.8	33.4	49.2
Top System	59	72	64.9

Results

	MUC-6			MUC-7		
	R	P	F	R	P	F
Ng & Cardie	63.3	76.9	69.5	54.2	76.3	63.4
Best MUC System	59	72	65	56.1	68.8	61.8

Classifier for MUC-6 Data Set

```
ALIAS = C: +
ALIAS = I:
| SOON_STR_NONPRO = C:
| | ANIMACY = NA: -
| | ANIMACY = I: -
| | ANIMACY = C: +
| SOON_STR_NONPRO = I:
| | PRO_STR = C: +
| | PRO_STR = I:
| | | PRO_RESOLVE = C:
| | | | EMBEDDED_1 = Y: -
| | | | EMBEDDED_1 = N:
| | | | PRONOUN_1 = Y:
| | | | | ANIMACY = NA: -
| | | | | ANIMACY = I: -
| | | | | ANIMACY = C: +
| | | | PRONOUN_1 = N:
| | | | | MAXIMALNP = C: +
| | | | | MAXIMALNP = I:
| | | | | WNCLASS = NA: -
| | | | | WNCLASS = I: +
| | | | | WNCLASS = C: +
| | | PRO_RESOLVE = I:
| | | | APPOSITIVE = I: -
| | | | APPOSITIVE = C:
| | | | GENDER = NA: +
| | | | GENDER = I: +
| | | | GENDER = C: -
```

Summary

- Performs better than the best non-learning approaches on two standard data sets
- Still lots of room for improvement
 - common noun resolution remains a major limiting factor