Local Coherence and Coreference Resolution

Regina Barzilay

MIT

October, 2005

SCIgen: An Automatic CS Paper Generator

- An output of a system that automatically generates scientific papers (Stribling et al., 2005):

> Active networks and virtual machines have a long history of collaborating in this manner. The basic tenet of this solution is the refinement of Scheme. The disadvantage of this type of approach, however, is that public-private key pair and red-black trees are rarely incompatible.

- The paper was accepted to a conference (not ACL!)

What's wrong?

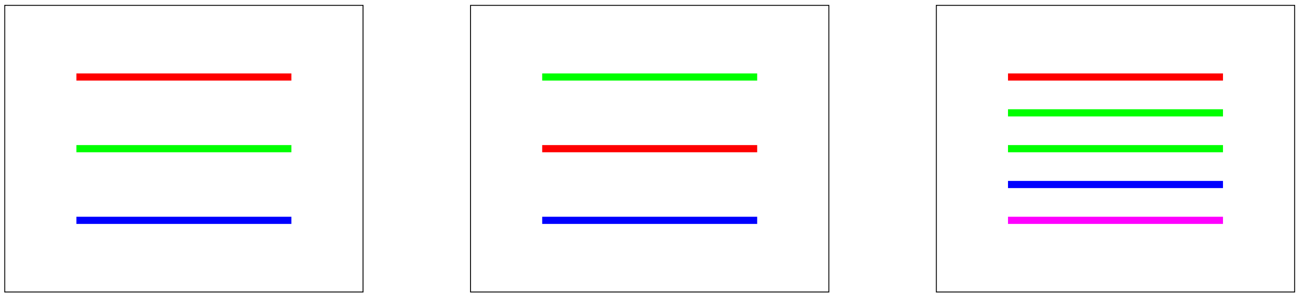
Active networks and virtual machines have a long history of collaborating in this manner. The basic tenet of this solution is the refinement of Scheme. The disadvantage of this type of approach, however, is that public-private key pair and red-black trees are rarely incompatible.

- **Coherence** is a property of well-written texts that makes them easier to read and understand than a sequence of randomly strung sentences

- **Local coherence** captures text organization at the level of sentence-to-sentence transitions

Modeling Local Coherence

Input: N alternative text realizations



Task: Find the most coherent alternative

Entity-based Approaches to Discourse

- Constraints on the entity distribution in a coherent text
  - Focus is the most salient entity in a discourse segment
  - Transition between adjacent sentences is characterized in terms of focus switch
- Constraints on linguistic realization of focus
  - Focus is more likely to be realized as subject or object
  - Focus is more likely to be referred to with anaphoric expression

Centering Theory

(Grozs&Joshi&Weinstein'95)

- Goal: to account for differences in perceived discourse

- Focus: local coherence
  global vs immediate focusing in discourse (Grosz'77)

- Method: analysis of reference structure

Phenomena to be Explained

John went to his favorite music store to buy a piano.

He had frequented the store for many years.

He was excited that he could finally buy a piano.

He arrived just as the store was closing for the day.

John went to his favorite music store to buy a piano.

It was a store John had frequented for many years.

He was excited that he could finally buy a piano.

It was closing just as John arrived.

Analysis

- The same content, different realization
- Variation in coherence arises from choice of syntactic expressions and syntactic forms

Another Example

John really goofs sometimes.
Yesterday was a beautiful day and he was excited about trying out his new sailboat.
He wanted Tony to join him on a sailing trip.
He called him at 6am.
He was sick and furious at being woken up so early.

Centering Theory: Basics

- Unit of analysis: centers
- "Affiliation" of a center: utterance (U) and discourse segment (DS)
- Function of a center: to link between a given utterance and other utterances in discourse

Center Typology

- Types:
  - Forward-looking Centers C_f (U, DS)
  - Backward-looking Centers C_b (U, DS)
- Connection: C_b (U_n) connects with one of C_f (U_{n-1})

Constraints on Distribution of Centers

- C_f is determined only by U;
- C_f are partially ordered in terms of salience
- The most highly ranked element of C_f (U_{n-1}) is realized as C_b (U_n)
- Syntax plays role in ambiguity resolution: subj $>$ ind obj $>$ obj $>$ $others$
- Types of transitions: center continuation, center retaining, center shifting

Center Continuation

Continuation of the center from one utterance not only to the next, but also to subsequent utterances

- $C_b(U_{n+1}) = C_b(U_n)$
- $C_b(U_{n+1})$ is the most highly ranked element of $C_f(U_{n+1})$ (thus, likely to be $C_b(U_{n+2})$)

Center Retaining

Retention of the center from one utterance to the next

- $C_b(U_{n+1}) = C_b(U_n)$
- $C_b(U_{n+1})$ is not the most highly ranked element of $C_f(U_{n+1})$ (thus, unlikely to be $C_b(U_{n+2})$)

Center Shifting

Shifting the center, if it is neither retained nor continued

- $\mathrm{C}_b(\mathrm{U}_{n+1}) <> \mathrm{C}_b(\mathrm{U}_n)$

Coherent Discourse

Coherence is established via center continuation

John went to his favorite music store to buy a piano.	John went to his favorite music store to buy a piano.
He had frequented the store for many years.	It was a store John had frequented for many years.
He was excited that he could finally buy a piano.	He was excited that he could finally buy a piano.
He arrived just as the store was closing for the day.	It was closing just as John arrived.

Centering Theory: Corpus-based Implementation

Key Premise: the distribution of entities in locally coherent discourse exhibits certain regularities

- Abstract a text into an entity-based representation that encodes syntactic and distributional information
- Learn properties of coherent texts, given a training set of coherent and incoherent texts

Text Representation

- Entity Grid — a two-dimensional array that captures the distribution of discourse entities across text sentences
- Discourse Entity — a class of coreferent noun phrases

Input Text

1 Former Chilean dictator Augusto Pinochet, was arrested in London on 14 October 1998.

2 Pinochet, 82, was recovering from surgery.

3 The arrest was in response to an extradition warrant served by a Spanish judge.

4 Pinochet was charged with murdering thousands, including many Spaniards.

5 He is awaiting a hearing, his fate in the balance.

6 American scholars applauded the arrest.

Input Text with Syntactic Annotation

Use Collins' parser(1997):

1. [Former Chilean dictator Augusto Pinochet]$_{\mathbf{S}}$, was arrested in [London]$_{\mathbf{X}}$ on [14 October]$_{\mathbf{X}}$ 1998.
2. [Pinochet]$_{\mathbf{S}}$, 82, was recovering from [surgery]$_{\mathbf{X}}$.
3. [The arrest]$_{\mathbf{S}}$ was in [response]$_{\mathbf{X}}$ to [an extradition warrant]$_{\mathbf{X}}$ served by [a Spanish judge]$_{\mathbf{S}}$.
4. [Pinochet]$_{\mathbf{S}}$ was charged with murdering [thousands]$_{\mathbf{O}}$, including many [Spaniards]$_{\mathbf{O}}$.
5. [He]$_{\mathbf{S}}$ is awaiting [a hearing]$_{\mathbf{O}}$, [his fate]$_{\mathbf{X}}$ in [the balance]$_{\mathbf{X}}$.
6. [American scholars]$_{\mathbf{S}}$ applauded the [arrest]$_{\mathbf{O}}$.

Notation: **S**=subjects, **O**=object, **X**=other

Input Text with Coreference Information

Use noun-phrase coreference tool (Ng and Cardie, 2002):

1. [Former Chilean dictator Augusto Pinochet]$_{\mathbf{S}}$, was arrested in [London]$_{\mathbf{X}}$ on [14 October]$_{\mathbf{X}}$ 1998.
2. [Pinochet]$_{\mathbf{S}}$, 82, was recovering from [surgery]$_{\mathbf{X}}$.
3. [The arrest]$_{\mathbf{S}}$ was in [response]$_{\mathbf{X}}$ to [an extradition warrant]$_{\mathbf{X}}$ served by [a Spanish judge]$_{\mathbf{S}}$.
4. [Pinochet]$_{\mathbf{S}}$ was charged with murdering [thousands]$_{\mathbf{O}}$, including many [Spaniards]$_{\mathbf{O}}$.
5. [He]$_{\mathbf{S}}$ is awaiting [a hearing]$_{\mathbf{O}}$, [his fate]$_{\mathbf{X}}$ in [the balance]$_{\mathbf{X}}$.
6. [American scholars]$_{\mathbf{S}}$ applauded the [arrest]$_{\mathbf{O}}$.

Output Entity Grid

	Pinochet	London	October	Surgery	Arrest	Extradition	Warrant	Judge	Thousands	Spaniards	Hearing	Fate	Balance	Scholars	
1	s	x	x	–	–	–	–	–	–	–	–	–	–	–	1
2	s	–	–	x	–	–	–	–	–	–	–	–	–	–	2
3	–	–	–	–	s	x	x	s	–	–	–	–	–	–	3
4	s	–	–	–	–	–	–	–	o	o	–	–	–	–	4
5	s	–	–	–	–	–	–	–	–	–	o	x	x	–	5
6	–	–	–	–	o	–	–	–	–	–	–	–	–	s	6

Comparing Grids

S	S	S	X	X	–	–	–	–	–	–	–	–	–	–
–	–	S	–	–	X	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	S	X	X	O	–	–	–	–	–
–	–	S	–	–	–	–	–	–	–	O	O	–	–	–
–	–	S	–	–	–	–	–	–	–	–	–	O	X	X
–	–	–	–	–	–	O	–	–	–	–	–	–	–	–

S	S	X	X	X	–	–	–	–	–	–	–	–	–	X
–	–	X	–	–	X	–	–	–	–	–	–	–	–	X
–	–	X	–	–	–	–	X	X	O	–	–	–	–	X
–	–	X	–	–	–	–	–	–	–	O	O	–	–	X
–	–	X	–	–	–	–	–	–	–	–	–	O	X	X
–	–	X	–	–	–	O	–	–	–	–	–	–	–	X

Coherence Assessment

- Text is encoded as a distribution over entity transition types
- Entity transition type — $\{\text{s}, \text{o}, \text{x}, \text{–}\}^n$

	S S	S O	S X	S –	O S	O O	O X	O –	X S	X O	X X	X –	– S	– O	– X	– –
d_1	0	0	0	.03	0	0	0	.02	.07	0	0	.12	.02	.02	.05	.25
d_2	.02	0	0	.03	0	0	0	.06	0	0	0	.05	.03	.07	.07	.29

How to select relevant transition types?:

- Use all the unigrams, bigrams, ... over $\{\text{s}, \text{o}, \text{x}, \text{–}\}$
- Do feature selection

Text Encoding as Feature Vector

	S S	S O	S X	S –	O S	O O	O X	O –	X S	X O	X X	X –	– S	– O	– X	– –
d_1	0	0	0	.03	0	0	0	.02	.07	0	0	.12	.02	.02	.05	.25
d_2	.02	0	0	.03	0	0	0	.06	0	0	0	.05	.03	.07	.07	.29

Each grid rendering x_{ij} of a document d_i is represented by a feature vector:

$$\Phi(x_{ij}) = (p_1(x_{ij}), p_2(x_{ij}), \ldots, p_m(x_{ij}))$$

where m is the number of all predefined entity transitions, and $p_t(x_{ij})$ the probability of transition t in the grid x_{ij}

Learning a Ranking Function

- **Training Set**
  Ordered pairs (x_{ij}, x_{ik}), where x_{ij} and x_{ik} are renderings of the same document d_i, and x_{ij} exhibits a higher degree of coherence than x_{ik}

- **Training Procedure**
  - Goal: Find a parameter vector $\vec{w}$ that yields a "ranking score" function $\vec{w} \cdot \Phi(x_{ij})$ satisfying:

    $$\vec{w} \cdot (\Phi(x_{ij}) - \Phi(x_{ik})) > \ 0$$
    $$\forall (x_{ij}, x_{ik}) \text{ in training set}$$

  - Method: Constraint optimization problem solved using the search technique described in Joachims (2002)

Evaluation: Text Ordering

- Goal: recover the most coherent sentence ordering
- Basic set-up:
  - Input: a pair of a source document and a permutation of its sentences
  - Task: find a source document via coherence ranking
- Data: Training 4000 pairs, Testing 4000 pairs (Natural disasters and Transportation Safety Reports)

Evaluation: Summarization

- Goal: select the most coherent summary among several alternatives
- Basic set-up:
  - Input: a pair of system summaries
  - Task: predict the ranking provided by human
- Data: 96 summary pairs for training, 32 pairs for testing (from DUC 2003)

Results

Tasks:

- O_1=ordering(Disasters)
- O_2=ordering(Reports)
- S=summary ranking

Model	O_1	O_2	S
Grid	87.3	90.4	81.3

Varying Linguistic Complexity

- What is the effect of syntactic knowledge?
  - Reduce alphabet to { **X**,**–** }

Model	O_1	O_2	S
+Syntax	87.3	90.4	68.8
-Syntax	86.9	88.3	62.5

- What is the contribution of coreference resolution?
  - Assume that entities are coreferent only if they have the same surface form

Model	O_1	O_2	S
+Coreference	87.3	90.4	68.8
-Coreference	83.4	89.7	81.3

Reference Resolution: Example

The Salesgirl (Burns and Allen)

Gracie: And then Mr. and Mrs. Jones were having matrimonal trouble, and my brother was hired to watch Mrs. Jones.

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

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

George: Well, what happened?

Gracie: She finally got a divorce.

George: Mrs. Jones?

Gracie: No, my brother's wife.

Reference Resolution: Example

The Salesgirl (Burns and Allen)

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

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

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

George: Well, what happened?

Gracie: She finally got a divorce.

George: Mrs. Jones?

Gracie: No, my brother's wife.

Reference Resolution

- Task: determine which noun phrases refer to each real-world entity mentioned in a document
- Goal: partition noun phrases in a text into coreference equivalence classes, with one cluster for each set of coreferent NPs

In the previous example:

{Mrs. Jones, she, she, Mrs. Jones},

{my brother, my brother},

{my brother's wife, she}

Definition: Coreference/Anaphora

- Coreference: Two expressions α_1 and α_2 are coreferent if and only if Referent(α_1)=Referent(α_2)
  - The expressions can be in the same text or different texts, in the same language or different language
- Anaphora: An expression α_1 is in an *anaphoric relation* with expression α_2 if and only if the interpretation of α_1 depends on α_2.
  - The relation holds *within* a text.

Relationship between Anaphora and Coreference

- Some expressions are both coreferential and anaphoric
  - *A bus had to divert to the local hospital when one of the passengers had a heart attack. It got to the hospital in time and the man's life was saved.*
- Some expressions are coreferential but not anaphoric

– *Alberto Gonzales, the White House counsel, intervened directly with Justice Department lawyers to obtain a legal ruling on the extent of president's authority to permit extreme interrogation practices. . . Mr.Gonzales's role in seeking a legal opinion on the definition of torture . . .*

Reference Resolution

Captain Farragut was a good seaman, worthy of the frigate he commanded. His vessel and he were one. He was the soul of it.

- Coreference resolution: {the frigate, his vessel, it}
- Anaphora resolution: {his vessel, it}

Coreference is a harder task!

Today's Topics

- Motivation
- Types of referential expressions
- Syntactic and semantic constraints on coreference
- Algorithms for coreference resolution

Motivation

- Information extraction
- Question-Answering
- Machine-Translation

  pronoun in the Malay language is translated by its antecedent (Mitkov, 1999)
- Summarization

When something goes wrong

The widespread impact of the term fundamentalist is obvious from the following quotation from one of the most influential Encyclopedias under the title 'Fundamentalist': "The term fundamentalist has. . . been used to describe members of militant Islamic groups." Why would the media use this specific word, so often with relation to Muslims? Most of them are radical Baptist, Lutheran and Presbyterian groups.

When something goes wrong

Why would the media use this specific word, so often with relation to Muslims?

Before the term fundamentalist was branded for Muslims, it was, and still is, being used by certain Christian denominations. Most of them are radical Baptist, Lutheran and Presbyterian groups.

Types of referential expressions: Nouns

- Indefinite Noun Phrases:

  I saw an Acura Integra today.

  Some Acura Integras were being unloaded.

  I saw this awesome Acura Integra today.

- Definite Noun Phrases:

  I saw an Acura Integra today. The Integra was white and needed to be washed.

  The fastest car in the Indianapolis 500 was an Integra.

Pronouns

Stronger constraints on using pronouns than on noun phrase references.

- Requires a high degree of activation from a referent
- Has a short activation span

a. John went to Bob's party, and parked next to a Acura Integra.
b. He went inside and talked to Bob for more than an hour.
a. Bob told him that he recently got engaged.
b. ??He also said that he bought *it* yesterday.

Demonstratives and One Anaphora

- Demonstratives (this, that) capture spatial proximity

  I like this one, better than that

- One Anaphora evokes a new entity into the discourse whose description is dependent of this new entity

  I saw no less that 6 Acuras today. Now I want one.

Troublemakers

- Inferables: inferential relation to an evoked entity

  I almost bought an Acura today, but the door had a dent and the engine seemed noisy.

- Discontinuous Sets: refer to entities that do not form a set in a text

  John has an Acura, and Mary has a Mazda. They drive them all the time.

- Generics: refer to general set of entities (in contrast to a specific set mentioned in text)

  I saw no less than six Acuras today. They are the coolest cars.

Today's Topics

- Motivation
- Types of referential expressions
- Syntactic and semantic constraints on coreference
- Algorithms for coreference resolution

Syntactic Constraints on Coreference

- Number Agreement

  John has a new Acura. It is red.

  John has three New Acuras. They are red.

- Person and Case Agreement

  John and Mary have Acuras. We love them.

  You and I have Acuras. We love them.

Syntactic Constraints

- Gender Agreement

  John has an Acura. It is attractive.

- Syntactic Agreement

  * John bought himself a new Acura.

  John bought him a new Acura.

Semantic Constraints

- Selectional restrictions of the verb on its arguments

  (1) John parked his Acura in the garage. He had driven it around for hours.

  (2) John parked his Acura in the garage. It is incredibly messy, with old bike and car parts lying around everywhere.

  (3) John parked his Acura in downtown Beverly Hills. It is incredibly messy, with old bike and car parts lying around everywhere.

Preferences in Pronoun Interpretation

- Recency: Entities introduced in recent utterances are more salient than those introduced further back

  John has an Integra. Bill has a Legend. Mary likes to drive it.

- Repeated mention: Entities that have been focused on in the prior discourse are more likely to continue to be focused on in subsequent discourse

  John needed a car to get his new job. He decided that he wanted something sporty. Bill went to the Acura dealership with him. He bought an Integra.

Preferences in Pronoun Interpretation

- Grammatical Role: Hierarchy of candidate entities based on their grammatical role

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

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

- Parallelism:

  Mary went with Sue to the Acura dealership. Sally went with her to the Mazda dealership.

Preferences in Pronoun Interpretation

Verb Semantics: emphasis on one of verb's arguments

- "implicit causality" of a verb causes change in salience of verb arguments

  John telephoned Bill. He lost the pamphlet on Acuras.

  John criticized Bill. He lost the pamphlet on Acuras.

- thematic roles (Goal, Source) cause change in salience of verb arguments

  John seized the Acura pamphlet from Bill. He loves reading about cars.

  John passed the Acura pamphlet to Bill. He loves reading about cars.

Distribution of NPs in Text

[Fraurud 1990] showed that, for Swedish non-fictional text (brochures, newspapers, texts, debate books)

- Different kinds of NPs occur with different frequencies
- Only a fraction of NPs evoke entities that anchor subsequent coreference
- Most definite NPs were not coreferential and not anaphoric

Distribution of Pronoun Types

Vicedo&Ferrandez'2000 show that for English text:

- Different pronouns occur with different frequencies in the same type of text
- Pronouns occur with different frequencies in different types of text

Text Collection	LAT	TIME	MED	CACM	Cranfield
he, she, they	38.59%	31.2%	15.07%	8.59%	6.54%
his, her, their	25.84%	35.01%	21.46%	15.69%	10.35%
it, its	26.92%	22.42%	57.41%	67.61%	79.76&

Distribution of Pronouns in Text

Pronouns in sentences

Text Collection	LAT	TIME	MED	CACM	Cranfield
0	44.8%	51.37%	77.84%	79.06%	90.95%
1	30.4%	29.46%	15.02%	17.54%	8.1%
2	14.94%	12.26%	4.75%	2.79%	0.85&
2+	9.56%	6.9%	2.39%	0.6%	0.09&

Syntactic properties of Pronouns (Hindu)

Roles	Number	Total	Frequency(%)
Subject-Direct Object	144	149	96
Subject-Indirect Object	50	57	87
Subject-PP Object	128	128	100
Direct Object-PP Object	22	22	100
Possessor-Head	22	22	100

Today's Topics

- Motivation
- Types of referential expressions
- Syntactic and semantic constraints on coreference
- Algorithms for coreference resolution

Hobbs' Algorithm (1976)

- Features: Fully Syntactic
  - search the parse in a left-to-right, breadth-first fashion
  - give a preference to antecedents that are closer to the pronoun
  - give a preference to subjects
- Refinement: When an NP is proposed as antecedent, gender/number agreement is checked

- Accuracy: 300 instances, 88.3% correct resolutions,
  91.7% with selectional restrictions
  132 ambigious cases, 72.2%correct resolutions,
  81.8% with selectional restrictions

Generic Algorithm

- Identification of Discourse Entities

  Identify nouns and pronouns in text

- Characterization of Discourse Entities

  Compute for each discourse entity NP_i a set of values from $\{k_{i1}, \ldots, k_{im}\}$ from m knowledge sources

- Anaphoricity Determination

  Eliminate non-anaporic expressions to cut search space

- Generation of Candidate Antecedents

  Compute for each anaphoric NP_j a list of candidate antecedents C_j

Generic Algorithm(cont.)

- Filtering

  Remove all the members of C_j that violate reference constraints

- Scoring/Ranking

  Order the candidates based on preferences and soft constraints

- Searching/Clustering

  Clustering of instances with the same antecedent

Clustering for Coreference

(Cardie&Wagstaff:1999)

- Each group of coreferent noun phrases defines an equivalence class
- Distance measure incorporates “linguistic intuition” about similarity of noun phrases
- Hard constraints enforce clustering construction

Instance Representation

Based noun phrases (automatically computed) are represented with 11 features:

- Individual Words
- Head Word
- Position
- Pronoun type (nominative, accusative)
- Semantic Class: Time, City, Animal, Human, Object (WordNet)
- Gender (WordNet, specified list)
- Animacy (based on WordNet)

Distance Metric

$$dist(NP_i, NP_j) = \sum_f w_f * incomp_f(NP_i, NP_j)$$

Clustering Algorithm

- Initialization: every noun is a singleton
- From right to left, compare each noun to all subsequent clusters
- Combine "close enough" clusters unless there exist any incompatible NP

Example: The chairman spoke with Ms. White. He ...

Results

MUC-6 (30 documents): Recall 48.8*%, Precision 57.4%, F-measure 52.8%
Baseline: 34.6%, 69.3%, 46.1% Types of Mistakes:

- Parsing mistakes
- Coarse entity representation and mistakes in feature computation
- Greedy nature of the algorithm

Supervised Learning

(Soon et al.,2001)

- Generate pairs of potential coreference expressions
- Represent every pair by a set of features that cpature their similarity
- Apply supervised learning algorithm
- Cluster pairs based on the classification score

Features (Soon et al, 2001)

- distance in sentences between anaphora and antecedent?
- antecedent in a pronoun?
- weak string identity between anaphora and antecedent?
- anaphora is a definite noun phrase?
- anaphora is a demonstrative pronoun?
- number agreement between anaphora and antecedent
- semantic class agreement anaphora and antecedent
- gender agreement between anaphora and antecedent
- anaphora and antecedent are both proper names?
- an alias feature
- an appositive feature

Vector-Based Representation

Example of feature encoding: (Ng&Cardie'2002)

```
0,76,83,C,D,C,D,D,D,D,D,I,I,C,I,I,D,N,N,D,C,D,D,N,N,N,N,N,C,Y,
Y,D,D,D,C,0,D,D,D,D,D,D,D,1,D,D,C,N,Y,D,D,D,20,20,D,D,-.
0,75,83,C,D,C,D,D,D,C,D,I,I,C,I,I,C,N,N,D,C,D,D,N,N,N,N,N,C,Y,
Y,D,D,D,C,0,D,D,D,D,D,D,C,1,D,D,C,Y,Y,D,D,D,20,20,D,D,+.
0,74,83,C,D,C,D,D,D,D,D,I,I,C,I,I,D,N,N,D,C,D,D,N,N,N,N,N,C,Y,
Y,D,D,D,C,0,D,D,D,D,D,D,D,1,D,D,C,N,Y,D,D,D,20,20,D,D,-.
```

Classification Rules

+ 786 59 IF SOON-WORDS-STR = C

+ 73 10 IF WNCLASS = C PROPER-NOUN = D NUMBERS = C SENTNUM <= 1 PRO-RESOLVE = C ANIMACY = C

+ 40 8 IF WNCLASS = C CONSTRAINTS = D PARANUM <= 0 PRO-RESOLVE = C

+ 16 0 IF WNCLASS = C CONSTRAINTS = D SENTNUM <= 1 BOTH-IN-QUOTES = I APPOSITIVE = C

+ 17 0 IF WNCLASS = C PROPER-NOUN = D NUMBERS = C PARANUM <= 1 BPRONOUN-1 = Y AGREEMENT = C CONSTRAINTS = C BOTH-PRONOUNS = C

+ 38 24 IF WNCLASS = C PROPER-NOUN = D NUMBERS = C SENTNUM <= 2 BOTH-PRONOUNS = D AGREEMENT = C SUBJECT-2 = Y

+ 36 8 IF WNCLASS = C PROPER-NOUN = D NUMBERS = C BOTH-PROPER-NOUNS = C

+ 11 0 IF WNCLASS = C CONSTRAINTS = D SENTNUM <= 3 SUBJECT-1 = Y SUBJECT-2 = Y SUBCLASS = D IN-QUOTE-2 = N BOTH-DEFINITES = I

Results

- Feature selection plays an important role in classification accuracy: MUC-6 62.6% (Soon et al., 2001) $\rightarrow$ Ng&Cardie, 2002) 69.1%
- Training size: 30 texts