# Coreference Resolution

Rule-based approaches

# Rule-based

- Before machine learning methods were used, rule-based methods were the go-to methods in coreference resolution
- Until 2011, these were even state-of-the-art (Stanford sieve)
- E.g. CoNLL:

| | | MUC | | | $B^3$ | | | |
|---|---|---|---|---|---|---|---|---|
| Track | Mentions | P | R | F1 | P | R | F1 | durchschn. F1 |
| Geschlossen | Nicht Gold | 57.5 | 61.8 | 59.6 | 68.2 | 68.4 | 68.3 | 57.8 |
| Offen | Nicht Gold | 59.3 | 62.8 | 61.0 | 69.0 | 68.9 | 68.9 | 58.3 |
| Geschlossen | Gold | 62.1 | 65.9 | 63.9 | 70.6 | 69.5 | 70.0 | 60.7 |
| Offen | Gold | 63.9 | 66.9 | 65.4 | 71.5 | 70.1 | 70.8 | 61.4 |

- There as the best system!

# Rule-based

- Main idea:
  - Resolve the mentions in the order they are found in the text
  - For each Mention:
    - Find a candidate set ("aggregate")
    - Filter the candidates by hard constraints ("filter")
    - And sort the candidates by soft constraints ("rank")
  - Main idea of the improvements of the system from Stanford:
    Resolve mentions by their security on "Easy-First" to minimize consequential errors!

# Rule-based - Constraints

- Hard constraints
  - Morphological constraints (gender, numerus)
  - Syntactic constraints (binding theory: "he combs his ..." vs "he combs himself")
  - Semantic constraints ("buildings cannot walk")

- Soft constraints ("preferences")
  - Syntactic ("resolving to subjects is better than resolving to objects")
  - Semantic ("Older people draw pensions more often than younger people")
  - Morphological (German example: "A neuter can be resolved to masculine antecedents as well as feminine antecedents")
    - Example with the German diminutive form
      - „Das Bürschchen" (the little boy) → antecedent masculine
      - „Das Mädchen" (the little girl) → antecedent feminine

# Lappin and Leass 1994

- Idea: Maintain a discourse model, in which there are representations for potential referents.
- Lappin and Leass 1994 propose a discourse model in which potential referents have degrees of salience.
- They try to resolve (**pronoun**) references by finding highly salient referents compatible with pronoun agreement features.
- In effect, they incorporate:
  - recency
  - syntax-based preferences
  - agreement, but no (other) semantics

# Lappin and Leass 1994

- First, we assign a number of salience factors & salience values to each referring expression.

- The salience values (weights) are derived by experimentation on a certain corpus.

# Lappin and Leass 1994

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | 80 |
| Existential emphasis | 70 |
| Accusative emphasis | 50 |
| Indirect object emphasis | 40 |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Lappin and Leass 1994

- Non-adverbial emphasis is to penalize “demarcated adverbial PPs” (e.g., “In his hand, …”) by giving points to all other types.
- Head noun emphasis is to penalize embedded referents.
- Other factors & values:
  - Grammatical role parallelism: 35
  - Cataphora: -175

# Lappin and Leass 1994

- The algorithm employs a simple weighting scheme that integrates the effects of several preferences:
  - For each new entity, a representation for it is added to the discourse model and salience value computed for it.
  - Salience value is computed as the sum of the weights assigned by a set of salience factors.
    - The weight a salience factor assigns to a referent is the highest one the factor assigns to the referent's referring expression.
  - Salience values are cut in half each time a new sentence is processed.

# Lappin and Leass 1994

- The steps taken to resolve a **pronoun** are as follows:
  - Collect potential referents (four sentences back); ("aggregate")
  - Remove potential referents that don't semantically agree;
  - Remove potential referents that don't syntactically agree; ("filter")
  - Compute salience values for the rest of potential referents;
  - Select the referent with the highest salience value. ("rank")

# Lappin and Leass 1994

- Salience factors apply per NP, i.e., referring expression
- However, we want the salience for a potential referent
  - So, all NPs determined to have the same referent are examined
- The referent is given the sum of the highest salience factor associated with any such referring expression
- Salience factors are considered to have scope over a sentence
  - So references to the same entity over multiple sentences add up
  - While multiple references within the same sentence don’t

# Example

- John saw a beautiful Acura Integra at the dealership.
- He showed it to Bob.
- He bought it.

# Example

- John saw a beautiful Acura Integra at the dealership.

| Referent | Phrases | Value |
|---|---|---|
| John | {John} | ? |
| Integra | {a beautiful Acura Integra} | ? |
| dealership | {the dealership} | ? |

# John

John saw a beautiful Acura Integra at the dealership.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | 80 |
| Existential emphasis | |
| Accusative emphasis | |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Integra

John saw a beautiful Acura Integra at the dealership.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | |
| Existential emphasis | |
| Accusative emphasis | 50 |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# dealership

John saw a beautiful Acura Integra at the dealership.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | |
| Existential emphasis | |
| Accusative emphasis | |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- John saw a beautiful Acura Integra at the dealership.

| Referent | Phrases | Value |
|---|---|---|
| John | {John} | 310 |
| Integra | {a beautiful Acura Integra} | 280 |
| dealership | {the dealership} | 230 |

# Example

- He showed it to Bob.

| Referent | Phrases | Value |
|---|---|---|
| John | {John} | 310/2 |
| Integra | {a beautiful Acura Integra} | 280/2 |
| dealership | {the dealership} | 230/2 |

| Referent | Phrases | Value |
|---|---|---|
| John | {John} | 155 |
| Integra | {a beautiful Acura Integra} | 140 |
| dealership | {the dealership} | 115 |

# He

He showed it to Bob.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | 80 |
| Existential emphasis | |
| Accusative emphasis | |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- He showed it to Bob.
- Gender constraint:
  "Integra" and "dealership" can't be "he"

| Referent | Phrases | Value |
| --- | --- | --- |
| John | {John, $he_1$} | 465 |
| Integra | {a beautiful Acura Integra} | 140 |
| dealership | {the dealership} | 115 |

# It

He showed it to Bob.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | |
| Existential emphasis | |
| Accusative emphasis | 50 |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- He showed it to Bob.

| Referent | Phrases | Value |
|---|---|---|
| John | {John, $he_1$} | 465 |
| Integra | {a beautiful Acura Integra} | 140 |
| dealership | {the dealership} | 115 |

- Gender constraint: “John” can’t be “it”
- Since “Integra” is more salient than “dealership” (140>115):
  → “it” refers to “Integra”

# Example

- He showed it to Bob.

| **Referent** | **Phrases** | **Value** |
|---|---|---|
| John | {John, $he_1$} | 465 |
| Integra | {a beautiful Acura Integra, $it_1$} | 420 |
| dealership | {the dealership} | 115 |

# Bob

He showed it to Bob.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | |
| Existential emphasis | |
| Accusative emphasis | |
| Indirect object emphasis | 40 |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- He showed it to Bob.

| **Referent** | **Phrases** | **Value** |
| --- | --- | --- |
| John | {John, $he_1$} | 465 |
| Integra | {a beautiful Acura Integra, $it_1$} | 420 |
| Bob | {Bob} | 270 |
| dealership | {the dealership} | 115 |

# Example

- He bought it.

| Referent | Phrases | Value |
|---|---|---|
| John | {John, $he_1$} | 465/2 |
| Integra | {a beautiful Acura Integra, $it_1$} | 420/2 |
| Bob | {Bob} | 270/2 |
| dealership | {the dealership} | 115/2 |

| Referent | Phrases | Value |
|---|---|---|
| John | {John, $he_1$} | 232.5 |
| Integra | {a beautiful Acura Integra, $it_1$} | 210 |
| Bob | {Bob} | 135 |
| dealership | {the dealership} | 57.5 |

# He

He bought it.

| Salience Factor | Salience Value |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | 80 |
| Existential emphasis | |
| Accusative emphasis | |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- He bought it.

| Referent | Phrases | Value |
| --- | --- | --- |
| John | {John, $he_1$} | 232.5 |
| Integra | {a beautiful Acura Integra, $it_1$} | 210 |
| Bob | {Bob} | 135 |
| dealership | {the dealership} | 57.5 |

Since “John” is more salient than “Bob” (232.5>135):

→ “He” refers to “John”

# Example

- He bought it.

| **Referent** | **Phrases** | **Value** |
|---|---|---|
| John | {John, $he_1$,He} | 542.5 |
| Integra | {a beautiful Acura Integra, $it_1$} | 210 |
| Bob | {Bob} | 135 |
| dealership | {the dealership} | 57.5 |

# It

He bought it.

| Salience Factor | Salience Value |
| --- | --- |
| Sentence recency | 100 |
| Subject emphasis | |
| Existential emphasis | |
| Accusative emphasis | 50 |
| Indirect object emphasis | |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

# Example

- He bought it.

| **Referent** | **Phrases** | **Value** |
| --- | --- | --- |
| John | {John, $he_1$,$he_2$} | 542.5 |
| Integra | {a beautiful Acura Integra, $it_1$} | 210 |
| Bob | {Bob} | 135 |
| dealership | {the dealership} | 57.5 |

- Gender constraint: “John” and “Bob” can’t be “it”
- Since “Integra” is more salient than “dealership” (210>57.5):
  → “it” refers to “Integra”

# Example

- He bought it.

| **Referent** | **Phrases** | **Value** |
|---|---|---|
| John | {John, $he_1$,$he_2$} | 542.5 |
| Integra | {a beautiful Acura Integra, $it_1$,$it_2$} | 490 |
| Bob | {Bob} | 135 |
| dealership | {the dealership} | 57.5 |

We should have added 35 for grammatical role parallelism, but we ignore this.

# Evaluation of Lappin and Leass 1994

- Weights were arrived at by experimentation on a corpus of computer training manuals
- Combined with other filters, algorithm achieve 86% accuracy (74% / 89% inter- / intra-sentential):
  - applied to unseen data of same genre

# Current state of the art

- Coreference resolution remains an unsolved problem:
- On CoNLL: (as of December 2017 Lee: "End-to-End Neural Coreference Resolution").

| | MUC | | | $B^3$ | | | $CEAF_{\phi_4}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Prec. | Rec. | F1 | Prec. | Rec. | F1 | Prec. | Rec. | F1 | Avg. F1 |
| Our model (ensemble) | **81.2** | **73.6** | **77.2** | **72.3** | **61.7** | **66.6** | **65.2** | **60.2** | **62.6** | **68.8** |
| Our model (single) | 78.4 | 73.4 | 75.8 | 68.6 | 61.8 | 65.0 | 62.7 | 59.0 | 60.8 | 67.2 |
| Clark and Manning (2016a) | 79.2 | 70.4 | 74.6 | 69.9 | 58.0 | 63.4 | 63.5 | 55.5 | 59.2 | 65.7 |
| Clark and Manning (2016b) | 79.9 | 69.3 | 74.2 | 71.0 | 56.5 | 63.0 | 63.8 | 54.3 | 58.7 | 65.3 |
| Wiseman et al. (2016) | 77.5 | 69.8 | 73.4 | 66.8 | 57.0 | 61.5 | 62.1 | 53.9 | 57.7 | 64.2 |
| Wiseman et al. (2015) | 76.2 | 69.3 | 72.6 | 66.2 | 55.8 | 60.5 | 59.4 | 54.9 | 57.1 | 63.4 |
| Clark and Manning (2015) | 76.1 | 69.4 | 72.6 | 65.6 | 56.0 | 60.4 | 59.4 | 53.0 | 56.0 | 63.0 |
| Martschat and Strube (2015) | 76.7 | 68.1 | 72.2 | 66.1 | 54.2 | 59.6 | 59.5 | 52.3 | 55.7 | 62.5 |
| Durrett and Klein (2014) | 72.6 | 69.9 | 71.2 | 61.2 | 56.4 | 58.7 | 56.2 | 54.2 | 55.2 | 61.7 |
| Björkelund and Kuhn (2014) | 74.3 | 67.5 | 70.7 | 62.7 | 55.0 | 58.6 | 59.4 | 52.3 | 55.6 | 61.6 |
| Durrett and Klein (2013) | 72.9 | 65.9 | 69.2 | 63.6 | 52.5 | 57.5 | 54.3 | 54.4 | 54.3 | 60.3 |