

GroRef: Rule-Based Coreference Resolution for Dutch

Rob van der Goot, Hessel Haagsma, Dieke Oele

Rijksuniversiteit Groningen

Introduction

Stanford's Multi-Pass Sieve Coreference Resolution System (Lee et al. [2011], Lee et al. [2013])

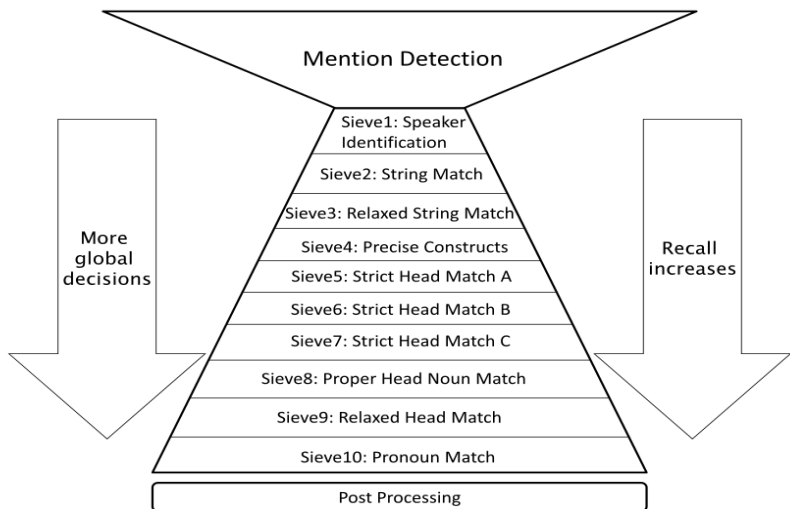
- Sieve-based architecture.
- Deterministic coreference models, stacked on top of each other.
- Each model builds on the previous model's clustering output.

Mention Detection

Alpino [Van Noord, 2006]

- Noun Phrases
- Names
- Subjects
- Pronouns

Sieves



Taken from Lee et al. [2013]

Results (Blanc)

| Corpus | Mention detection | | | Coreference | | |
|-------------|-------------------|----|----|-------------|----|----|
| | R | P | F1 | R | P | F1 |
| Apple (dev) | 65 | 57 | 61 | 37 | 28 | 31 |
| Boeing | 60 | 58 | 60 | 32 | 31 | 31 |
| GM | 64 | 58 | 61 | 35 | 29 | 32 |
| Stock | 62 | 48 | 53 | 35 | 20 | 26 |

Mention Detection Errors

- Errors in approach:
[Hooggeplaatst manager bij [Apple]] ...
- Annotation inconsistencies
... de fabrikant van de iPhone ([die] op woensdag voor het eerst ...
[de [iTunes Apps Store]] vs. [de iPhone]
- Mistakes of Alpino

Example Output

Hooggeplaatst manager bij [Apple] verlaat [bedrijf] na antenneproblemen [iPhone 4]

[Mark Papermaster] , [de manager bij [Apple Inc.]] [die] toezicht houdt op [hardware engineering van [de iPhone]] , is weg bij [het bedrijf] . Dit volgde op de kritiek [die Apple] vorige maand kreeg vanwege de positie van de antenne op het meest recente model van de [iPhone] , [de iPhone 4]] . [Apple] bevestigde dat [Papermaster] (49) [het bedrijf] had verlaten maar wilde niet zeggen of [hij] uit eigen beweging wegging of was ontslagen . [Papermaster] wilde geen commentaar geven op de situatie .

Conclusion

- Stanford's coreference resolution system can easily be adapted to other languages
- This system is robust across different domains
- Unsupervised: no training data needed

Heeyoung Lee, Yves Peirsman, Angel Chang, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. Stanford's multi-pass sieve coreference resolution system at the conll-2011 shared task. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, pages 28–34. Association for Computational Linguistics, 2011.

Heeyoung Lee, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Computational Linguistics*, 39 (4):885–916, 2013.

Gertjan Van Noord. At last parsing is now operational. 2006.