Rule-based Coreference Resolution with BART

Julian Baumann, Xenia Kühling, Sebastian Ruder
Institute for Computational Linguistics, Univ. Heidelberg

Goal

We improve coreference resolution through integration of the rule-based, entity-centric sieve architecture developed by the Stanford NLP group into the existing BART machine learning system for coreference resolution.

Stanford Deterministic Coreference Resolution System

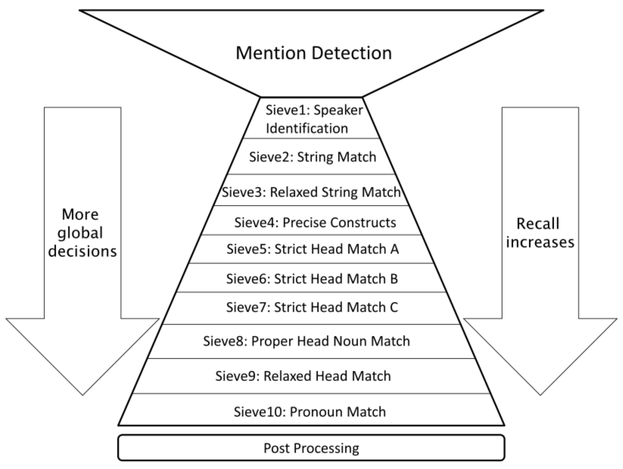

Fig. 1: Functionality of Stanford Coreference System

- Input (mentions) passes ten independent precision-oriented coreference models ("sieves")
- Entity-centric approach uses previous sieves' output and information to make decisions
- top ranked system at the CoNLL-2011 shared task

Examples

Speaker Identification Sieve

"**[Ich]**$^1_1$ schließe jetzt ab", sagt **[der Standesbeamte Rolf Paschen]**$^1_2$ resolut, "sonst wird das hier nie was."

```
TRUE! Antecedent of '[der Standesbeamte Rolf Paschen]':'[Ich]'
```

Precise Constructs Sieve

Dafür spricht **[[ihre]**$^1_1$ **klassische Ausbildung]**$^2_2$, **[die]**$^2_3$ nicht mit [Wegwerfkultur]$^3_4$ und platten Melodien zusammen paßt.

```
TRUE! Antecedent of '[die]':'[ihre klassische Ausbildung]'
```

Pronoun Match Sieve

[Der koreanische Autokonzern Daewoo]$^1_1$ wollte auf [keinen Fall]$^2_2$ mit [[**seinem**]$^1_3$ Autoumschlag]$^3_4$ in [Bremerhaven]$^4_5$ bleiben und mit [[**seiner**]$^1_6$ Konzern-Zentrale]$^5_7$ auch nicht nach [Bremerhaven]$^4_8$ gehen.

```
TRUE! Antecedent of '[seinem]':'[Der koreanische Autokonzern Daewoo]'
TRUE! Antecedent of '[seiner]':'[Der koreanische Autokonzern Daewoo]'
```

Entities that require more or commonsense knowledge

[Der Saatgutkonzern Pioneer Hi-Bred]$^1_1$ hat in [Süddeutschland]$^2_2$ [nicht zugelassenen Gentech-Mais]$^3_3$ verkauft.
[Der Weltmarktführer für [Saatgut]$^4_4$**]**$^1_5$ verstößt damit gegen [das Gentechnikgesetz]$^5_6$, [...].

```
FALSE! No Antecedent for '[Der Weltmarktführer für Saatgut]'
ANTECEDENT: [Der Saatgutkonzern Pioneer]
```

BART

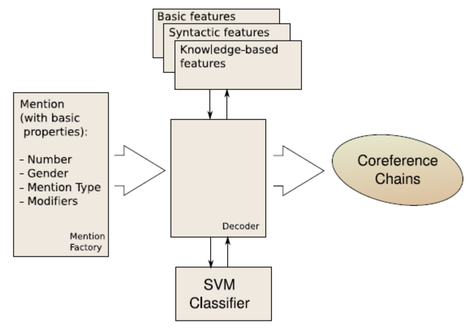

Fig. 2: Functionality of BART

- Mentions containing semantic information about a markable (gender, number, etc.) are generated
- Machine Learning employs syntactic and semantic features to generate pair instances (anaphor, antecedent) which are assembled in coreference chains

Evaluation

We used the following data for evaluation.
GERMAN: first 100 documens of TüBa-D/Z (2008)
ENGLISH: CoNLL-2011 Shared Task training set

	MUC-Score		
	Recall	**Precision**	**F_1**
Our system	0.644	0.691	0.667
BART	0.721	0.532	0.612

Fig. 3: *GERMAN*: Comparison with BART ML Configuration (`XMLExperiment`)

	MUC-Score		
	Recall	**Precision**	**F_1**
SpeakerIdentification	0.004	0.637	0.008
+StringMatch	0.157	0.857	0.265
+RelaxedStringMatch	0.180	0.825	0.295
+PreciseConstructs	0.241	0.822	0.372
+HeadMatchA	0.295	0.809	0.432
+HeadMatchB	0.355	0.775	0.487
+HeadMatchC	0.357	0.771	0.488
+ProperHeadNounMatch	0.358	0.771	0.489
+RelaxedHeadMatch	0.383	0.771	0.512
+PronounMatch	0.644	0.691	0.667

Fig. 4: *GERMAN*: Performance of individual sieves

	MUC-Score
	F_1
Our system	0.420
Stanford	0.603

Fig. 5: *ENGLISH*: Comparison with Stanford System

Conclusion

The rule-based sieve approach exceeds BART's Machine Learning performance. Since our system has been primarily designed using specific German linguistic constants, there is still a lot of room for improvement of the English language version.
Due to the nature of the rule-based approach, the system is easy to extend. We leave this along with its adaptation to English, Italian, and other languages as future work.

References

Broscheit, S. et al. (2010), BART: A multilingual anaphora resolution system, in 'Pro- ceedings of the 5th International Workshop on Semantic Evaluation', SemEval '10, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 104–107.
Lee, H. et al. (2013), 'Deterministic coreference resolution based on entity-centric, precision-ranked rules', Comput. Linguist. 39(4), 885–916.
Versley, Y. et al. (2008), BART: A modular toolkit for coreference resolution, in 'Pro- ceedings of the ACL-08: HLT Demo Session', Association for Computational Lin- guistics, Columbus, Ohio, pp. 9–12.