

Supplementary material for Evaluation of Named Entity Coreference

Anonymous NAACL submission

Pronoun Resolution Accuracy We also track the NEC F1 when only mentions of given syntactic type are preserved in the chain—name, pronoun and nominal. Of special interest is to track performance when resolving pronouns. Many of these issues arise due to the need for common-sense knowledge and reasoning.

A school official₁ talked about the Vice President’s chances during an interview with the Boston Globe. He₂ says it’s unlikely Gore will be selected, because he₃ doesn’t have enough experience in the academic world.

The pronoun *he₃* is incorrectly resolved to the school official instead of Al Gore. To correctly resolve this, we need to know that not having enough experience would be the reason for Gore not getting selected and not the reason for the school official making a statement.

Maybe Lily became so obsessed with where people slept and how because her own arrangements kept shifting. When Rosie₁ died, her₂ uncles moved in and let her₃ make the sleeping and other household arrangements.

Rosie is Lily’s mother (explained earlier in the article). *Her_{2,3}* are incorrectly resolved to Rosie, despite the common sense that a person can not make household arrangements after their death.

Over-splitting/combination of entities We tracked the over-splitting (systems produce multiple clusters for the same name) and the over-combination of entities as well (placing mentions to different named entities in the same cluster. This error usually occurs when different people have the same last name but also occasionally when the names are completely different but the roles of the people are similar). However, the overall such errors were quite small and similar for all systems and have thus not been included in the later tables with results.

In the following example, mentions 2 and 3 to Frank Curzio form one cluster while the other three form another one -

Frank Curzio₁. Many people now claim to have predicted the 1987 crash. Queens newsletter writer Francis X. Curzio₂ actually did it: He₃ stated in writing in ... Mr. Curzio₄ says what happens now will depend a good deal on the Federal Reserve Board. If it promptly cuts the discount rate it charges on loans to banks, he₅ says, “That could quiet things down.”

Although it is clear from the below text snippet that Kofi Annan and Danilo Turk are two different people, they have similar titles and recognized to be coreferent.

UN Secretary General Kofi Annan₁ said Wednesday, it is important to help spread democracy In a new report, the Secretary General₂ says democratization Commenting on the report, UN Assistant Secretary General for Political Affairs Danilo Turk₃ told reporters

Clusters in Ontonotes Statistics, in OntoNotes (nw,bn,mz), on coreference for PERson, ORGa-nization, geopolitical entity (GPE) and DATE named entities. Non-singleton entities are mentioned at least twice in a text and so require coreference. Entities is the percentage of coreference clusters to entities of the given type. Entity mentions is the percentage of all individual references to any entity of a given type. Average cluster size is the average number of coreferent mentions.

	PER	ORG	GPE	DATE
Non-singleton Entities	67.98%	50.76%	52.02%	21.3%
Entity Mentions	15.44%	11.35%	11.36%	4.24%
Named mentions	24.47%	14.33%	13.62%	3.41%
Non-named mentions	45.12%	55.79%	68.43%	82.69%
Avg cluster size	54.88%	44.21%	31.57%	17.31%
	5.74	4.57	4.34	2.89

Table 1