

Mention Detection: Heuristics for the OntoNotes Annotations

Jonathan K. Kummerfeld, Mohit Bansal, David Burkett, and Dan Klein

{jkk,mbansal,burkett,klein}@cs.berkeley.edu

Overview

The system our submission was based on, Haghighi and Klein (2010), was designed to consider all non-trivial syntactic NPs. Here we present examples of the heuristic filters we introduced to determine which NPs fit the OntoNotes annotation scheme.

Approach

We construct the input mention set by extracting the maximal NP for each possible head, and applying filters to reject some mentions. After running the Haghighi and Klein (2010) system we use additional filters to reject mentions from the output.

Removed before resolution

Appositive

The case, *Hodgson vs. Minnesota*, will be argued Nov 29.

Good removal

Appositive

Tribe, *Gore's attorney*, said the state court did nothing illegal.

Good removal

WP Tag

It was Justice Breyer *who* asked the most straightforward question of all.

Good removal

EX Tag

Although the picture generally looks pretty good, *there* is still much work to be done.

Good removal

Copular Verb

Fangchenggang City is a recently established coastal open port city at the regional level approved by the State Council in May 1993.

Good removal

Copular Verb

After all, come January, these people are going to have to be in Washington whether Gore or Bush is *the President*.

Bad removal

Removed after resolution

RB Tag

Okay *here's* my problem.

Good removal

Non-Year Cardinal

The Dow Jones industrials skidded 190.58, to *2569.26*.

Good removal

Non-Year Cardinal

Two were married on June twenty-fifth in Newport Rhode Island.

Bad removal

Generic you

And I feel like I got it off my chest *you know*.

Good removal

NNS Tag

So, *infants* are strictly forbidden from sitting in the front seat.

Good removal

NNS Tag

The networks and *studios* have bickered ever since.

Bad removal

Generic it

Perhaps *it* is because utopian nostalgia evokes both military experience and the social gospel.

Good removal

Generic Head

Business as usual, and *war* is good business.

Good removal

Generic Head

It strives to desensitize *people* towards others' suffering by surrounding them with constant killing.

Bad removal

