

# The Life and Death of Discourse Entities: Identifying Singleton Mentions

Authors

Marta Recasens, Marie-Catherine de Marneffe, Christopher Potts

Paper Presentation  
by Gregory Katsios

# Presentation Overview

- Datasets
- Motivation
- Problem
- Previous efforts
- Authors approach
  - Predicting lifespans with linguistic features
- Results
- Contribution
- Questions

# Datasets

- CoNLL-2012 Shared Task
  - Including 1.6M English words from OntoNotes v5.0

# Motivation

- The progress of
  - Coreference resolution models
  - Textual entailment
  - Discourse coherence

# What is the problem?

- What factors influence the lifespan of a discourse referent
- Predict if a discourse referent is coreferent or singleton

# Previous Engineering Efforts

- Predict the role that different mentions play in coreference chains
- Detecting four different target mentions
  - non-referential, non-anaphoric, discourse-new and non-antecedent

# Authors Approach

- Cross-cut the four categories
- Try to identify
  - non-referential and referential noun phrases
    - where the referent is mentioned only once

# Authors Approach

- Feature representations
  - Close to the original insights
    - Extract features from full syntactic parses
    - Including feature interaction terms

# Predicting Lifespans with Linguistic Features

- Binary distinction between coreferents and singletons
- Features of the model
  - Morphosyntax, grammatical role and semantic environment of the mentions

# Morphosyntax of the Mentions

- Internal morphology and syntactic structure of the mention
- Extracted features
  - Type, animacy, person, number, quantification, number of modifiers and named entities

# Grammatical Role of the Mentions

- Coreferent mentions appear as core verbal arguments and favor sentence-initial positions
- Grammatical relation of the mention
- Sentence position of the mention
- Whether or not the mention is in a conjunct

# Semantic Environment of the Mentions

- No access to the semantic scope
  - Dependency representations capturing the syntactic scope for
    - Negation, modal auxiliaries and attitude predicates

# Results

- The lifespan model confirms the claims concerning how semantic operator interact with specific kinds of mentions
- The model can distinguish singleton and coreferent mentions apart
  - 78% accuracy (baseline of 55.1%)

# Contribution

- Perform singleton / coreferent detection
- The combination of
  - Linguistic theories into the “Lifespan model”

# Greg's Opinion

This paper was hard!

## Suggestion

Read “*Modeling the Lifespan of Discourse Entities with Application to Coreference Resolution*”  
by the same authors.

# Questions

1. Does the classification of an entity as singleton depend on the length of research data? (ie. If text is really long, can two mentions of NE still be considered singleton)
2. Is the tradeoff between precision and recall worth it if, by decreasing recall, we miss significant relevant details about a topic because of false negatives?
3. How does the authors' approach deal with non-antecedent mentions
4. What could be the real life applications for this kind of logistic regression model?
5. How does having singletons as the majority of the data help in differentiating singletons from coreferents or predicting their lifespans?
6. Considering the fact that singleton entity can never be coreferent and they considered every singleton entity as only non-referential NPs but not verbal mentions.
7. Can ever singleton entity be considered as referential or verbal mentions ? Why are they always non-referential?