

Event Coreference for Information Extraction

Humphreys, Gaizauskas, and Azzam (1997)

Presented by Will Corvey

Motivation

- Coreference resolution has traditionally treated objects: nouns, pronouns
- Information Extraction (IE) seeks to fill templates with key information.
 - Templates for a “scenario” include:
 - The event type
 - Roles pertaining to event type
 - NLP systems should be concerned with this information, too

Identifying Events

- Simple events:
 - “Mr. Jones **succeeds** M. James Bird, 50, as president of Wholistic Therapy.”
- Complex events:
 - “Daniel Wood was named president and chief executive officer of EFC Records Group, a unit of London’s Spear EFC PLC. He succeeds Charles Paulson, who was recently made chairman and chief executive officer of EFC Records Group North America.”
 - “The sell-off followed the resignation late Monday of Freddie Heller, the president of Renard Broadcasting Co. Yesterday, Renard named Susan B. Kemptham, chairman of Renard Inc.’s television production arm, to succeed him.”

The Problem

- Events in a variety of syntactic contexts
 - In the main clause
 - In a relative clause
 - As nominalizations or infinitives
- The syntax obscures the point that event coreference should be as “easy” as object coreference.

The LaSIE System (DARPA 1995)

- NE recognition, coreference resolution, template filling, auto-summarization
- Pipeline architecture:
 - Lexical preprocessing
 - Tokenization, POS tagging, morphological analysis, proper name matching
 - Parsing and semantic interpretation
 - Syntactic parsing, NE recognition
 - Discourse interpretation
 - (From world model) Presupposition information, coreference resolution ("instances in world model"), inference from input

General Coref. Resolution

1. Add the sentence to the world model, giving a **discourse model**
2. New instances should match older instances in the discourse model

How?

- Semantic type consistency (semantic path length) ==> similarity score
 - Attribute consistency (attributes must match) ==> similarity score
3. Sum the similarity score, and merge the highest scoring pair in the discourse model

The Semantic Ontology

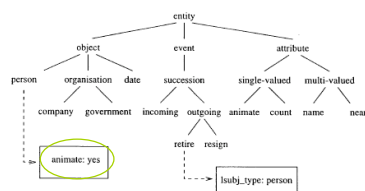


Figure 1: Upper ontology for the management succession task

The Semantic Ontology

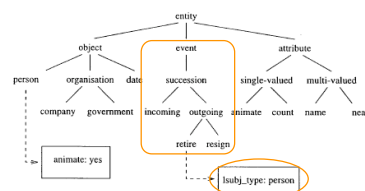


Figure 1: Upper ontology for the management succession task

Constraints for Events

- Task independent:
 - Time stamps
 - Tense resolution (temporal phases)
- Task-specific:
 - Organizational or management positions (on lower nodes)
- Verb-specific constraints

Example

"The sell-off followed the resignation late Monday of Freddie Heller, the president of Renard Broadcasting Co.

```
% Sentence 3a
sell-off(e2), number(e2,sing), det(e2,the),
follow(e1,time(e1,past)),
lsobj(e1,e2), lobj(e1,e3),
resignation(e3), number(e3,sing),
det(e3,the),
date(e5), name(e5,'Monday'),
person(e7), name(e7,'Freddie Heller'),
title(e8,president),
company(e10),
name(e10,'Renard Broadcasting Co.')
```

Yesterday, Renard named Susan B. Kempham, chairman of Renard Inc.'s television production arm, to succeed him."

```
% Sentence 3b
yesterday(e11), number(e11,sing),
name(e15,'Renard'),
name(e12), time(e12,past),
lsobj(e12,e13),
person(e15), name(e15,'Susan B. Kempham'),
apposed(e15,e16),
title(e16,chairman),
arm(e18), number(e18,sing),
qual(e18,e19),
production(e19), number(e19,sing),
qual(e19,e20), of(e19,e21),
television(e20),
company(e21), name(e21,'Renard Inc. '),
succeed(e14), time(e14,present),
lsobj(e14,e15), lobj(e14,e22),
pronoun(e22,him)
```

Notes on Example

- Presuppositions encoded as attributes of outgoing event
- From ontology: Incoming_event, changeover_event, outgoing_event
- Object coreference also necessary
 - Notion of the logical subject for each event (may be hypothesized if not explicit)
- Focus on resolution (merging) of events in the ontology

Discussion

Does this really work? Sort of...

	Succession Events		Overall		
	Recall	Precision	Recall	Precision	Combined
No Event Coref	66%	72%	42%	60%	48.88%
With Evgnt Coref	65%	77%	42%	60%	49.40%

Table 1: Upper Ontology for the management succession task

The next step would be manual annotation of coreference events in a large corpus; "not feasible" here.

- * **Advantage:** powerful extension of existing IE architecture
- * **Challenge:** Difficult task, results in marginal improvement; best results shown in small domain.