

# Center for Cognitive Science

University at Buffalo, State University of New York

**Wednesday, November 9, 2005**

280 Park Hall  
North Campus  
2:00 pm –4:00 pm

Cassandra Creswell, Ph.D.

Janya, Inc., Buffalo, NY

*(Sponsored by the Department of Linguistics, UB)*

## **"Automatically detecting nominal mentions of events**

The basic task in information extraction (IE) is to automatically extract mentions of entities, relations, and events from text. Usually, events are treated as textually anchored by verb phrases or sentences, as in (1), but of course they can also be referred to with noun phrases, as in (2).

- 1) [Amr Moussa was appointed Secretary General of the Arab League last year.]
- 2) [Amr Moussa's appointment as Secretary General of the Arab League] took place last year.

Automatic detection of nominal mentions of events faces several challenges. The one we address here is the systematic polysemy unique to nominal (vs. verbal) mentions, such that they can be used to denote either an event or the result, outcome, or product of an event.

- 3) Things are getting back to normal in the Baywood Golf Club after [a chemical spill](=event). Clean-up crews said [the chemical spill](=result) was 99 percent water and shouldn't cause harm to area residents.

This ambiguity is problematic for IE systems because the ultimate goal in event extraction is to create abstract event objects containing information from multiple coreferring mentions, such as their time, location, and participants. Information should not be propagated from mentions of non-events to events, and event mentions should not be collapsed with non-event mentions.

Tackling this ambiguity requires more fine-grained syntactic and semantic evidence than is used in standard word sense disambiguation techniques (Schuetze, 1998) because the event-result ambiguity does not obey the One Sense Per Discourse principle (Gale, et al., 1992).

The goal of this project is to develop a classifier that can label NPs as event-denoting or non-event-denoting based on their local lexico-syntactic context. Because learning from hand-annotated data alone is not sufficient to overcome the sparse data problem, we use a weakly-supervised bootstrapping technique using known unambiguous terms to learn to distinguish the contexts which make an event vs. non-event interpretation more probable. In this talk I will present some initial experimental results from our ongoing project on nominal event detection.

**Refreshments will be available; Open to the public!**