Language Technologies Institute

Thesis Defense

NSH 1507| Monday, February 12| 12:00 pm

Event Extraction for Document-Level Structured Summarization

Andrew Hsi

COMMITTEE:

Yiming Yang, (co-chair)

Jaime Carbonell, (co-chair)

Alex Hauptmann

Michael Mauldin
(external member)

Abstract

Event extraction has been well studied for more than two decades, through both the lens of document-level and sentence-level event extraction. However, event extraction methods to date do not yet offer a satisfactory solution to providing concise, structured, document-level summaries of events in news articles. Prior work on document-level event extraction methods have focused on highly specific domains, often with great reliance on handcrafted rules. Such approaches do not generalize well to new domains. In contrast, sentence-level event extraction methods have applied to a much wider variety of domains, but generate output at such fine-grained details that they cannot offer good document-level summaries of events.

In this thesis, we propose a new framework for extracting document-level event summaries called macro-events, unifying together aspects of both information extraction and text summarization. The goal of this work is to extract concise, structured representations of documents that can clearly outline the main event of interest and all the necessary argument fillers to describe the event.

We propose three novel methods to address the macro-event extraction task. First, we introduce a structured prediction model based on the Learning to Search framework for jointly learning argument fillers both across and within event argument slots. Second, we propose a multi-layer neural network that is trained directly on macro-event annotated data. Finally, we propose a deep learning method that treats the problem as machine comprehension, which does not require training with any on-domain macro-event labeled data. Our experimental results on a variety of domains show that such algorithms can achieve stronger performance on this task compared to existing baseline approaches. Furthermore, under scenarios of limited training data, we find that machine comprehension models can offer very strong performance compared to directly supervised algorithms, while requiring very little human effort to adapt to new domains.

http://www.cs.cmu.edu/~ahsi/andrew-hsi-thesis.pdf