

COMPUTER SCIENCE  
SEMINAR

*Pattern-Based Mining of Entity/Relation Structures from  
Massive Text*

Qi Li, PhD

University of Illinois at Urbana-Champaign

**Abstract:** Majority of information nowadays is carried by massive and unstructured text, in the form of news, articles, reports, or social media messages. This poses a major research challenge on mining entity/relation structures from unstructured text. Manual curation or labeling cannot be scalable to match the rapid growth of text. Most existing information extraction approaches rely on heavy human annotations, which can be too expensive to tune and not adaptable to new domains. In this talk, I will present a pattern-based methodology that conducts information extraction from the massive corpora using existing resources with little human effort. The first component, WW-PIE, discovers meaningful textual patterns that contain the entities of interest. The second component, TruePIE, discovers high quality textual patterns for target relation types. I will demonstrate how semi-supervised methods can empower information extraction for broad applications and provide explainable results.

Thursday, February 28, 2019, 10:00 am  
Mathematics and Science Center: W303

COMPUTER SCIENCE  
EMORY UNIVERSITY