מכון ויצמן למדע

THE WEIZMANN INSTITUTE OF SCIENCE

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 1 ,Ziskind Building
on Wednesday, Jan 21, 2015
at 11:15

Omri Abend
University of Edinburgh

Machine-Learning the Universal Semantics of Natural Languages

Abstract:

The field of Natural Language Processing (NLP) has recently been pivotal in producing important language technologies such as machine translation and question answering. Such technologies are based on elaborate structural representations of text, detected by statistical methods. However, common approaches to structural representation are language-specific or even domain-specific, limiting the applicability of NLP tools and models. How to represent both the idiosyncrasies of specific domains and languages as well as their commonalities is still an open question. In my talk I will address these questions and propose an approach for learning a level of representation shared by all languages using latent variable models for structured prediction. Under this approach, learning starts from universally-applicable coarse-grained logical structures, which is used to bootstrap the learning of more fine-grained semantic distinctions, as well as the learning of the specifics of individual languages. I will discuss the value of universal semantic structures both to the computational modeling of child language acquisition and to leading NLP applications, focusing on machine translation. Joint work with Ari Rappoport, Shay Cohen and Mark Steedman.