

INVITED TALK: Embedding Probabilistic Logic for Machine Reading

Sebastian Riedel

University College London, UK

sebastian.riedel@gmail.com

Abstract

We want to build machines that read, and make inferences based on what was read. A long line of the work in the field has focussed on approaches where language is converted (possibly using machine learning) into a symbolic and relational representation. A reasoning algorithm (such as a theorem prover) then derives new knowledge from this representation. This allows for rich knowledge to be captured, but generally suffers from two problems: acquiring sufficient symbolic background knowledge and coping with noise and uncertainty in data. Probabilistic logics (such as Markov Logic) offer a solution, but are known to often scale poorly.

In recent years a third alternative emerged: latent variable models in which entities and relations are embedded in vector spaces (and represented "distributionally"). Such approaches scale well and are robust to noise, but they raise their own set of questions: What type of inferences do they support? What is a proof in embeddings? How can explicit background knowledge be injected into embeddings? In this talk I first present our work on latent variable models for machine reading, using ideas from matrix factorisation as well as both closed and open information extraction. Then I will present recent work we conducted to address the questions of injecting and extracting symbolic knowledge into/from models based on embeddings. In particular, I will show how one can rapidly build accurate relation extractors through combining logic and embeddings.

Bio

Dr. Riedel is a Senior Lecturer in the Department of Computer Science at University College London, leading the Machine Reading lab. He received his MSc and PhD (in 2009) in Computer Science from the University of Edinburgh. He was a researcher at the University of Tokyo, and a postdoc with Andrew McCallum at the University of Massachusetts Amherst. He is an Allen Distinguished Investigator, a Marie Curie CIG fellow, was a finalist for the Microsoft Research Faculty Award in 2013 and recently received a Google Focused Research award. Sebastian is generally interested in the intersection of NLP and machine learning, and particularly interested in teaching machines to read, and to reason with what was read.