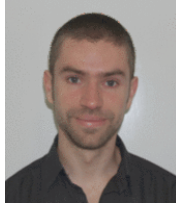




CLaS Talk



Dr Phil Blunsom

University of Oxford Computing Laboratory

My research interests lie at the intersection of machine learning and computational linguistics. I apply machine learning techniques, such as graphical models, to a range of problems relating to the understanding, learning and manipulation of language. Recently I have focused on structural induction problems such as grammar induction and learning statistical machine translation models.

Date: 11th of April (Mon) 2011

Time : 11:00 am – 12:30 pm

Venue: Bldg E6A Room 202

Topic: Inducing Synchronous Grammars for Machine Translation.

Abstract:

In this talk I'll introduce the current state of the art in statistical machine translation (SMT) and outline my work modeling machine translation as a probabilistic machine learning problem.

Although SMT systems have made large gains in translation quality in recent years, most are currently induced using a hand engineered pipeline of disparate models linked by heuristics. Although such techniques are effective for translating between related languages (e.g. English and French), they fail to capture the latent structure necessary to translate between languages which diverge significantly in syntactic structure, such as Chinese and English.

I'll present non-parametric Bayesian models for inducing synchronous context free grammars. These models are capable of learning the latent structure of translation equivalence from a corpus of parallel string pairs. I'll discuss the difficult inference problems posed by such models and describe Monte Carlo sampling techniques that can help solve them. Finally I'll present experiments demonstrating competitive results on full scale translation evaluations..