Latent variables in RNNG

Source

See `What do recurrent neural networks learn about syntax?`: * The RNNG can be trained on *unlabeled* trees. * The composition function produces representations for each constituent phrase. * These representations can be clustered and *recover* the main grammatical categories. * This despite the fact that the model has never been trained on labels.

Idea

Let latent variables induce the labels. * Put latent variables in the composition function. * Put priors on these variables that induce a labeling of the reduction. * Use something sparsity inducing priors. * Use semi-discrete latent variables for this.

Sources about (non neural network) latent variable models for this are: * Klein and Manning 2002: A Generative Constituent-Context Model for Improved Grammar Induction * Petrov et al. 2006: Learning Accurate, Compact, and Interpretable Tree Annotation