Predicting and Discovering Linguistic Structure with Neural Networks.
Promotor is dr. C. Monz. Copromotor is dr. A. Bisazza
M.K. Tran
In the field of natural language processing (NLP), recent research has shown that deep neural network models are quite brittle and can only model linguistic principles to a limited degree. In order to build NLP systems that can generalize and work well in practice, it is important to integrate linguistic knowledge into such systems as well as investigate the ability of current models in capturing linguistic phenomena. This thesis attempts to address those two aspects from four different angles.

First, this thesis demonstrates that neural network models enable the integration of morphological knowledge seamlessly into phrase-based machine translation systems without any feature engineering.

Second, this thesis investigates what linguistic phenomena are implicitly captured by recurrent neural networks by augmenting them with an external memory. This thesis also studies the impact of recurrent vs non-recurrent architectures in modeling hierarchical structure.

Third, while neural networks are well known to be powerful supervised learners, this thesis investigates whether they offer the same benefits for unsupervised structure learning. This thesis proposes an unsupervised Neural Hidden Markov Model for the purpose of part-of-speech induction.

Finally, this thesis asks whether neural networks can induce meaningful structure from non-annotated text. This thesis proposes structured attention models that induce a dependency-like tree representation of the input sentence for the purpose of translation. Moreover, this thesis shows that the models learn some basic elements of the source language grammar.