

## Evaluation of doctoral thesis

Student: Mgr. Nathan David Green

Thesis title: Improvements to Syntax-based Machine Translation  
using Ensemble Dependency Parsers

Supervisor: doc. Ing. Zdeněk Žabokrtský, Ph.D.  
Institute of Formal and Applied Linguistics, MFF UK  
Malostranské náměstí 25  
118 00 Praha 1

The main aim of Nathan's research was to improve the performance of dependency parsing, in order to improve the performance of a machine translation system that uses such parses in one of its intermediate steps.

The center of gravity of Nathan's approach lies in using a set (ensemble) of already existing parsers, and combining their outputs. Building a mere combination of already existing systems is sometimes perceived as a low-hanging-fruit task. However, those who tried it seriously know how far from true this expectation can be, especially in well established disciplines that already reach high state of the art accuracy. Even if it is not novel to expect that different systems produce different error distributions, it is by far not guaranteed that one can successfully identify criteria under which this or that system (or a subset of systems) is more trustful than the others.

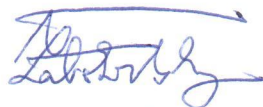
Nathan introduced, implemented and systematically evaluated several combination techniques. He showed which techniques lead to better parsing quality and/or better machine translation quality. This is the biggest contribution of the thesis. Perhaps the most surprising observation from this part is how much one can gain by using features that capture agreement between individual systems - very simple to implement, yet very helpful.

Apart from the main research direction, Nathan performed several other experiments, such as measuring the impact of using deeper noun phrases on the quality of syntax-based machine translation. Another important outcome from his work is his hand-checked dataset of English dependency trees.

During his PhD study, Nathan became an author or co-author of about 10 scientific papers published in reviewed proceedings of international conferences or workshops. Last but not least, Nathan has verified that if a PhD student focuses fully on his own research topic, a solid dissertation can be produced in considerably shorter time than what is our department standard.

To conclude, in my opinion Nathan's work presents a valuable scientific contribution and I find it fully adequate for a PhD defense.

Rychnov nad Kněžnou, 20. 8. 2013



doc. Ing. Zdeněk Žabokrtský, Ph.D.