Review of Thesis Proposal

Reviewer:	RNDr. Ondřej Bojar, Ph.D.; bojar@ufal.mff.cuni.cz ÚFAL MFF UK, Malostranské náměstí 25, Praha 1, 181 00
Date:	31. 10. 2017
Thesis Title:	Improving Neural Machine Translation with External Information
Candidate:	Mgr. Jindřich Helcl
Supervisor:	prof. RNDr. Hajič Jan, Dr. ÚFAL MFF UK

The thesis proposal by Jindřich Helcl summarizes very well the current state of the research in neural machine translation and clearly documents that the author has already significantly contributed to the state of the art and progress in this area.

I have some detailed questions:

- When describing the work by Caglayan et al. (2016b), which shows some gains in multi-modal translation, Jindřich mentions that the object recognition network is different: ResNet-50 vs. VGG-16 used in previous works. Calixto et al. (2017) uses VGG-19. It would be good if Jindřich could comment on the differences between the three networks (in terms of architecture but more importantly of their performance in object recognition) and their impact on the downstream translation task. (Was the improvement thanks to the different visual network, or due to other changes?) When considering the inclusion of linguistic annotation in NMT training, choosing a particular grammatical framework (esp. constituency vs. dependency representation) and particular tools is likely to considerably influence the effectiveness of the method, so it would be good to have some sense of the magnitude of differences in performance we can expect.
- When describing the Eriguchi et al. (2017) paper, the use of SyntaxNet parser instead of a manually annotated corpus is mentioned. Is only the single-best parse fed to the system? Would Jindřich expect performance gains if the multi-task training was performed on a corpus that is both manually annotated and parallel? Or would it make sense to automatically translate a treebank and use this synthetic parallel corpus to train the system? (Forward translation is risky, but backtranslation is very effective; training a system on back-translated treebank so that it both translates and parses the target seems rather promising to me.)

I value high the work devoted to the toolkit Neural Monkey, although it is arguably very hard to keep up the pace with much larger research teams such as the one at Google. Neural Monkey will therefore probably not deliver the top performance for unconstrained tracks at competitions like WMT. Its value should therefore be seen (and advertised for) in the educational use. I would recommend applying for e.g. faculty grants to support masters students helping to polish, document and popularize the toolkit.

The proposal is currently very succinct regarding the proposed future research. The general directions are clear and I fully support them but I would still like to see some more specific proposals and particularly promising first options relying on Czech linguistically-annotated data. I would like to ask the author to present some concrete experimental setups at the defense: which of the Praguian treebanks, other annotated corpora (e.g. discourse) or tools could best serve in which network architectures and towards which challenges in NMT. Some prioritization or proposed stages of this future research would be also very desirable to ensure timely progress towards finishing the thesis.

The text of the proposal is well structured and equipped with clear illustrations and all important formulas. It is written in overall very good English, with only a small number of grammatical errors. I have a number of smaller corrections to the text which I will pass directly to the author.

In sum, the thesis proposal by Jindřich Helcl is of sufficient quality and documents Jindřich's expertise in the area. Despite the lack of a detailed work plan, I fully recommend to accept the proposal.

In Prague, October 31, 2017.

Ondřej Bojar