Recommendation for the defense of Ph.D. thesis

Ph.D. candidate: **Karel Veselý**
Advisor: **Lukáš Burget**

I am happy that Karel Veselý has submitted his thesis under my supervision, and that his long-years efforts have converged to this day.

Karel's thesis addresses the important area of **semi-supervised training for automatic speech recognition (ASR)**. Based on solid theoretical and implementation knowledge of DNN training, he has extensively studied confidence measures for modern DNN-based ASR systems and compared granularities of data selection, as well as selection vs. probabilistic weighting approaches. His thesis proposes a recipe that is general enough and simple to be implemented in a broad scope of training data and scenarios. He has thoroughly evaluated the suggested approaches on several standard data-sets. I am happy that he managed to consolidate his findings into a document that I consider very helpful to the community.

The thesis, however, cover only part of Karels's activities. He has been around the BUT Speech@FIT group for over a decade, starting with a bachelor project on simple hybrid recognizer. He obtained further know-how during his Erasmus stay in Avignon and started to focus on DNNs and their training. The result of his diploma project – **TNet** (a parallel tool for DNN training) has been since used in the group. With Dan Povey and others, Karel was one of the founders of Kaldi toolkit. His Kaldi module `nnet1`, allowing for sequence discriminative training of DNN acoustic models, has brought him international recognition and made him a "Kaldi DNN guru". During the DARPA RATS project, he developed a winning **voice-activity detector**, which allowed the team to obtain excellent results on all follow-up technologies. During the IARPA Babel project, he developed a new paradigm for **multi-lingual training of DNN-based features and acoustic models**, that allowed the BUT group and our Babel team to take advantage of multi-lingual training over single target language only.

Although Karel presented these results always in his typical shy way (so different from our American colleagues), they had an impact – according to Scopus, he has **802 citations** (the paper on discriminative training of DNNs alone was cited 254 times) and has **h-index 13** which is more than most of senior faculty members at FIT BUT. Karel has been also active pedagogically - besides teaching (the usual set of signals and systems and speech processing), he led the successful Bc. thesis of Katka Žmolíková, who is now with the group as PhD student.

From personal point of view, I appreciate Karel's sense of responsibility and availability – he has always been available for help in case anyone needed to dig deep into DNN training and practical applications. He has also been one of the key contributors to Kaldi, and has helped and consulted seniors and students during numerous international workshops and stays. I also appreciate his help and organization of social activities for the group, especially its foreign staff, students and visitors.

To conclude, **I fully recommend Karel Veselý's Ph.D. thesis for the defense**, I am very pleased that he will stay with the group, and I am looking forward to future cooperation with him.

In Brno, September 4th 2017

Lukáš Burget, Ph.D.
BUT Speech@FIT, Department of Computer Graphics and Multimedia
Faculty of Information Technology, Brno University of Technology
Božetěchova 2, 612 66 Brno, Czech Republic
Tel: +420 5 41141269 Fax: +420 5 41141290
mailto:burget@fit.vutbr.cz, http://www.fit.vutbr.cz/~burget