

Recommendation for the defense of Ph.D. thesis

Ph.D. candidate: Lucas Ondel
advisor: Lukáš Burget

I am happy that Lucas Ondel has submitted his thesis under my supervision, and that his long-years efforts have converged to the submission of doctoral dissertation at FIT BUT.

Lucas's thesis is in the important area of automatic unit discovery that allows for work with under-resourced languages without extensive linguist or phonetician knowledge. With its sound theoretical framework (Bayesian techniques, subspace projections, sequence modeling) it however touches a much broader area of machine learning and I am persuaded that it will make a lasting international impact. Lucas has worked on these approaches at BUT under my supervision (that later turned into taking over most of the initiative and putting me in a (desired!) position of scholar), but he also received valuable feedback and supervision from Najim Dehak and Hynek Hermansky (both at the Johns Hopkins University).

The scientific content of the thesis is sufficiently covered in the document itself and in the reviews, let me therefore concentrate on other, more personal remarks.

Lucas came to FIT as an Erasmus/Socrates student from an industrial (read "not very mathematical") master program at UAPV in France. By his curiosity and hard work, he gradually became one of the machine learning gurus of our group, always ready to discuss deep theoretical concepts, able to cover lots of real-estate on our whiteboards with equations, and able to turn these into a working (and public!) code and interesting experimental results. Moreover, he has always been keen to discuss projects and cooperations on an international level.

During his PhD, Lucas moved around quite a bit:

- He spent considerable time at the Johns Hopkins University where he first participated in the 2016 JSALT workshop (work-group "Building Speech Recognition System from Untranscribed Data" led by Lukas Burget) and then collaborated with Najim Dehak and Hynek Hermansky
- He was a member of team "The Speaking Rosetta Stone - Discovering Grounded Linguistic Units for Languages without Orthography" led by Emmanuel Dupoux at the 2017 JSALT workshop held at Carnegie Mellon University.

He was also instrumental in our participation at several international technology evaluations:

- MediaEval (Query by Example Search on Speech track). The task involved searching for audio content within audio content using an audio content query. This task was particularly interesting for speech researchers in the area of spoken term detection or low-resource speech processing. At the beginning of his Ph.D., in 2013, Lucas helped us to train a collection of modules that were used in the BUT keyword spotting system. Thanks to his help, we were able to be in the top ranking of the evaluation.
- DARPA LORELEI - The techniques he has developed were used to build a topic identification system for spoken documents. His method for acoustic units discovery combined with n-gram-based document classification were among the best performing systems in zero-resource setting in 2017 and 2018 evaluations, avoiding the need to build a full-fledge ASR pipeline.

Lucas was also active pedagogically - besides usual involvement in Signals and Systems, he replaced me in lectures and exercises of machine learning courses and he co-advised several successful students and visitors (Bolaji Yusuf, Martin Kocour and Martin Šústek).

From a personal point of view, I appreciate Lucas's sense of responsibility and availability – he has always been available for help in case anyone needed to dig deep into machine learning, Bayesian statistics or Python and Julia coding.

To conclude, **I fully recommend Lucas Ondel's Ph.D. thesis for the defense.** I am wishing him all the best in future professional and personal life, and I am looking forward to continuing to cooperate with him.

In Brno, September 3rd 2020

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>