

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Maria Obedkova
Název práce Data-Driven Pronunciation Generation for ASR
Rok odevzdání 2019
Studijní program Computer Science **Studijní obor** Computational Linguistics

Autor posudku Ondřej Plátek **Role** Vedoucí
Pracoviště Institute of Formal and Applied Linguistics MFF UK

Text posudku:

The goal of the thesis is to explore methods for automatic lexicon generation for Automatic Speech Recognition (ASR). This thesis focuses on methods with as little supervision as possible.

Chapters 1, 2 provide well written theoretical background and current research overview. Chapter 3 describes from a high-level point of view the structure and goal of experiments described in this thesis.

Chapter 4 introduces the corpus that we chose for our experiments.

In chapters 5, 6 and 7, a very low-level set of experiments is well described but the number of experiments may be confusing with regards to the overall task of generating data-driven pronunciation with as little supervision as possible.

The core of the thesis lies in a novel approach to generating pronunciation candidates based on acoustic similarities and clustering the candidates. Several extensions of this idea using a pre-trained acoustic model for ASR are explored.

Chapter 8, nicely summarizes the conducted experiments and put previously described experiments into context.

In general, I appreciate easy to read thesis with well-placed figures, tables, and diagrams. A little more work on the organization of the thesis, more specific examples, and spelling would make the thesis even better.

The author boldly suggests a novel approach and describes precisely conducted experiments.

There are two major setbacks which reinforce each other:

1. The author cannot provide source code which is highly non-standard and unfortunate given.
2. The result of the experiments did not bring performance improvement and it is hard to decide if the cause of obtaining worse results against a well-polished implementation from the state-of-the-art toolkit Kaldi is the fault of the author's implementation or the chosen method because the author implementation is not available for review.

Possibly more extensive experimentation may better explain the cause of the results but it would require substantially more time.

Overall, the author understood numerous subfields of ASR and was able to conduct meaningful experiments and implement algorithms and components in ASR pipeline as well as build the final ASR evaluation pipeline using the Kaldi toolkit.

In conclusion, the author implemented a large number of high-quality experiments which unfortunately have not finished with wished results. To fully understand the quality of the proposed method, more experimentation is required.

Consequently, I recommend the thesis for the defense and suggest grade two - very good.

Práci doporučuji k obhajobě.

Práci na zvláštní ocenění.

Nenavrhuji.

**2.9.
2019**

Podpis Ondřej Plátek