Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce   Micha de Rijk

Název práce   Codenames: a practical application for modelling word association

Rok odevzdání   2020

Studijní program   Informatika   **Studijní obor**   Matematická lingvistika

Autor posudku   David Mareček   **Role**   vedoucí

Pracoviště   Ústav formální a aplikované lingvistiky

Text posudku:

Modelling word associations for the purposes of developing an artificial player of the game Codenames is quite a challenging task. When we were starting, we did not know whether it is even possible to create any reasonably playing artificial intelligence that would be able to give hints targeting more than one word.

Micha implemented a very nice online version of the game suitable for testing various word-association models by human players. It took much more time than we expected since many optimizations were needed to be done so that the models do not consume much memory and the responses were fast enough.

Micha finally tested three models: one is based on word-embeddings and two are based on sentence-level and dependency-level collocations. Each model provides different types of hints. A natural continuation of work would be combining this models so that one hint can target different words using different word-association models at the same time. This turned out to be a hard problem since the models have different scaling and, also due to lack of time, no reasonable solution was found. I also regret a bit that Micha has not tested other methods based on manually annotated data as WordNet or human-generated data. He argues that such methods have not sufficient coverage on the words in the game. However, they could play at least a good role in a combined model.

The thesis itself has 72 pages including attachments. It is well organized and written in good English. However, a substantial part of the thesis was written in the last days before the deadline, so there was not much time for fine-tuning the text. There are some inconsistencies across the chapters and some paragraphs are rather redundant. On the other hand, all the decisions made during the implementation are well explained in detail and the whole story of the thesis is clear.

Overall, Micha had proven that he can work on a given task independently. Even though I expected that more methods and combinations could be explored, there turned out to be many

difficulties on the way, which Micha was struggling with and finally solved. The number of experiments performed is fully sufficient for a diploma thesis. Besides the basic and improved models, Micha implemented a nice and playable game, which can be further used for evaluation of other models for word-associations in the future. I recommend this thesis to be defended.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 31. 1. 2020

Podpis: