

Assessment of Masters' Thesis as an External Examiner

Title: **Deep Reinforcement Learning for Autonomous Off-road Driving in Simulation**

Author: **Jacques Valentin**

Supervisor: **Karel Zimmermann, doc. Ing., Ph.D.**

External examiner: **Dr. Gaël Écorchard**

Fulfillment of Assigned Tasks

All assigned tasks though being of a high difficulty level have been fulfilled. The choice of the simulator is well argued. Both methods of machine learning have been correctly implemented.

Resolution Methods

The neural network architecture which is used along the work is based on a publicly available neural network which was used for self-driving cars with a monocular camera. The network is then simplified to better fit the task at hand. This is a reasonable approach. However, the simplification method was not explained. In particular, there is no mention whether layers were dropped out or their depth reduced, or both. Also, it is not specified whether only the architecture of the neural network was adapted or if the network weights could also be used as a pre-trained network. In the thesis, two and a half pages concern the technique of manually driving the car, probably directly through the simulator, and trying to gather data to train the neural network. The student probably spend a lot of time on this and wanted to explain this technique in the manuscript. According to me, however, this part is not necessary because the method using the API is much clearer and requires only a few lines of codes, it gives better results, and , being an API, is a more direct way to control the simulator and get data from it. Moreover, driving a car with a joystick rather than with the keyboard seems more practical.

Obtained Results

The results of the imitation learning are quite detailed. A few details are missing though. In Fig. 12, the effect of the learning rate is given but the *keep_prob* value used for the trainings is not given. As stated by the student himself, the results of the training with 10^{-4} and 10^{-5} are very similar. It seems to me, however, that a network obtained by slow training but steady improvements is better than a overfitting network. This is actually confirmed with Fig. 16 where the training had to be stopped very early before overfitting.

I miss an explanation of the behavior of the learning process for *keep_prob* = 0.7 in Fig. 14. The behaviors with *keep_prob* = 0.6 and 0.8 are very similar and very different from the one with *keep_prob* = 0.7.

At p.31 is unclear in which the previous weights were suboptimal. According to the text, the conditions of the training Fig. 14 with *keep_prob* = 0.9 are the same, number of training steps included. It is then unclear why a further training process was required and this is confirmed by the fact that the results in Fig. 16 look very similar.

On the contrary to the results of the imitation learning, the results about reinforcement learning are rather sparse. I admit though that the implementation of three reinforcement learning algorithms are already a good result requiring a lot of work.

In general, it would have been interesting to show at least one example of the behavior of the both trained algorithms by plotting the position of the vehicle on the road.

In p. 45, the student states that the training of the reinforcement learning algorithms takes less time than the imitation learning with ADAM optimizer. I would like the student to elaborate on this during the defense of his thesis.

Practical Requirements

The thesis is written with very good English level and only a minimal amount of grammatical errors could be found. There seems to be a unfinished sentence p. 17 though.

There are unfortunately a lot of stylistic errors or inconsistencies along the document which could have been corrected in a short time. Some paragraphs have a double line spacing for no apparent reason. Most tables are larger than the text, especially Table 1 that contains a lot of empty space. The font used in tables should be the same as in the text or, at least, it should be consistent across all tables. Section numbering is missing. Chapters must start on a right odd page. Right pages must have odd numbers, not the opposite. Page 44 has a large empty space as before a new section but page 45 does not start with any section title. The insertion of mathematical notations within the text provokes inconsistent line spacing.

The figures are clear and appropriate. I particularly appreciate that the plot of the learning processes that are difficult to read in the text are also given in annexes. It is a pity though that these plots do not use the same scale, which would facilitate the comparison among them.

In contradiction to the overall mostly error-prone text of the thesis, the bibliography really lacks some love:

- authors of all citations are identified by their first name and the last name is only given as initial,
- most of the citations do not contain the type of publication, [1], [2], [6], [7], [8], [9], [12], [13], [15], [16],
- the title of [7] is incomplete,
- wrong journal name in [3] and [4],
- [15] miss the date of publication,
- there are a lot of typographic errors:
 - [1], [6] and [9] start with a period ("."),
 - [3] and [6] have a comma at the beginning of a line,
 - a period follow a comma in [8],
 - wrong punctuation in [9].

General Comments and Conclusion

The presented thesis is of very good quality. The results are well presented but there are a few points that would need to be clarified, as the different measures taken to simplify the neural network on which this work is based. More results for the reinforcement learning would have been welcome and both machine learning methods miss the presentation of an example trajectory and the required statistical information associated with the presented average results.

As a conclusion, I advise the commission to evaluate the presented Masters' thesis with the grade

B - Very Good.

Prague, June 15, 2018

Dr. Gaël Écorchard
ČVUT, CIIRC