DIPLOMA THESIS REVIEW

Author: Bc. Michaela Urbanovská

Name: Learning Classical Planning Transition Functions by Deep Neural Networks

The review was elaborated by: MSc. Alexander Shleyfman, PhD candidate

Michaela Urbanovská's diploma thesis focuses on the field of Artificial Intelligence (AI), namely, the application of machine learning techniques, such as various neural networks (NN), in the setting of the grid-like domains of classical planning (Maze, multi-agent Maze, Sokoban). The presented work tackles two distinct problems: learning of a deterministic transition function (given a state\action pair return a successor state), and learning of a heuristic function (evaluate the distance from a given state to a goal). The convolutions neural network (CNN) that approximates the transition function for the maze domains was obtained by chaining multiple CNN layers, usage of 3x3 kernel, and exploitation of residual connections modification, for the Sokoban domain the same network was used, except the kernel was increased to 5x5 and 7x7. Logistic entropy loss was chosen as a loss function. The heuristic function was built using CNN layers (with 3x3 kernels) interleaved with attention blocks (each with its own mask). The experiments compared the approach proposed by the author, to well known BFS algorithms such as $GBFS$ and A^*, equipped with state-of-the-art heuristics, e.g., h^{FF}, LM-cut.

The thesis is divided into six chapters. The first chapter is dedicated to introduction of the work, and establishes the goal of the thesis. The second chapter provides theoretical background information on the techniques used in the thesis. This chapter has various minor errors and inconsistencies. The third chapter, which is the main contribution of the thesis provides information on the structure of the NN used to solve the problems established in chapter one. Chapter four presents the the empirical evaluation of the proposed techniques. Lastly, chapters five and six are dedicated to result discussion and conclusion, correspondingly. The code for the experiments was developed in Julia, and its description is given in the appendix. The work is well motivated and overall interesting. On the other hand, The experimental evaluation of the NN-heuristic looks rather discouraging. Note that neither h^{FF} and LM-cut are domain independent, and does not require learning time. Thus, the fact that NN-heuristic does not favourably compete with them seems somewhat disappointing. Note also that the coverage of all domains show that all algorithms solved all problems with standard successor generation. This implies that the problems were not chosen up to scale. Lastly, it seems that the proposed successor generation not yet ready to be deployed.

The thesis is sufficiently well-written and its structure was chosen appropriately: most of the chapters are well-organized and have a good flow; however, there are still typographical errors, incomplete definitions and redundant subsections. I.e.,

1. In Definition 2.1.1, $c(o) : o \rightarrow R$ should be written as $c : o \rightarrow \mathbb{R}^+$, since the author does not intend to work with negative costs. Moreover, it looks like the whole thesis concerns only unit-cost domains.
2. In Definition 2.1.3, The definition of plan lacks the initial state and the applicability of actions.
3. In Definition 2.1.4, “if no subsequence of π is a solution for Π”. Does the author assume zero-cost actions? If so, this should be explicitly specified, and if not, this part should be removed altogether.
4. p. 12, “From a mathematical perspective, a universal approximation theorem states, that neural networks can learn any continuous function.” This is misleading. The theorem says that we can learn an approximation to every continuous function.
5. In Figure 2.3, in the kernel K, the 1 in the upper right corner should be 0.
6. Subsections 3.4.3 and 3.4.4 does not provide information on the structure of the NN.
7. The citation [2], “Classical planning in deep latent space: Bridging the subsymbolic-symbolic boundary”, lacks authors – Masataro Asai and Alex Fukunaga.

The bullet points above show sloppy work on the finalization of the text, and problems with both planning the work and providing an adequate overview of the relevant literature. The student aptly quotes the relevant literature. However, renaming the GBFS algorithm that uses $f + h$ as its priority queue function, seems rather drastic to me. The heuristic that the student named “none” is known in the literature as the “blind” heuristic. The student chose a difficult, yet very important topic and the research conducted was solid. The experimental part was executed sufficiently well. The work assignment has been fulfilled in most of the points.

The work **meets the requirements for a Thesis** of an Very Good quality. I evaluate the diploma thesis by the mark **Very Good (B)**.

I would like to present the following questions:

1. The maze problem presented in the thesis can be viewed as a variations of the shortest path and traffic routing problems. In their work Rauch and Winarske, ”Neural networks for routing communication traffic” (1988), proposed to RNN, namely, Hopfield network for the soultion of such. Have you considered to evaluate your approach against the Hopfield networks?
2. Given that the IPC domains include domains such as Sokoban, Grid, and Visit-all, have you considered integrating the experiments on this domains in you thesis?

Toronto, University of Toronto, 16.01.2020

Alexander Shleyfman, MSc.