SSY 340 - Project Planning

Group 23:
Anton Karlsson (antka)
Calle Hellberg (hcalle)
Tommy Sy
Anton Jansson
Examiner: Lennart Svensson

October 7, 2020

1 The Problem

The problem we are faced with is to create a modified snake game environment and then train an agent to play said game using reinforcement learning. It will be a challenge to design a good network for this problem and interface it with our game environment. Once that is achieved training can begin to find optimal policy using Q-learning.

2 Data

Since we are doing RL we will not need any external dataset.

3 Available code

The game snake is a well established game with numerous sources of code that can be used to build the environment. In terms of the uniqueness of the code, certain alternations and adjustment will be added to avoid using already available code and fit the algorithm to the reinforcement learning problem. The adjustment will be to add new elements that are not used in the original game, E.g power ups, punishments or restrictions. The adjustments will not be deterministic and can be modified over the course of the project since it is hard to define the extent of the project for beginners. To facilitate the creation of the code and graphics we will be using the pygame package.

The reinforcement learning algorithm will more or less be written from scratch with the use of the pytorch, gym and numpy packages.

4 Relevant papers

We listed a couple of papers that deal with Q-learning that seems relevant to our problem.

- https://www.toptal.com/deep-learning/pytorch-reinforcement-learning-tutorial
- https://towardsdatascience.com/snake-played-by-a-deep-reinforcement-learning-agent-53f2c4331d36
- https://www.diva-portal.org/smash/get/diva2:1342302/FULLTEXT01.pdf

- https://arxiv.org/abs/1312.5602
- https://arxiv.org/abs/1507.06527
- https://arxiv.org/pdf/1909.04751.pdf
- https://arxiv.org/pdf/1905.04127.pdf
- https://arxiv.org/abs/1808.05032?
- https://ieeexplore.ieee.org/abstract/document/8460004
- https://ieeexplore.ieee.org/abstract/document/9070827
- http://cs229.stanford.edu/proj2016spr/report/060.pdf
- https://ieeexplore.ieee.org/document/9070827
- https://www.researchgate.net/publication/327638529_Autonomous_Agents_in_Snake_G

5 Evaluate result

The evaluation of performance will be based on the number of training iterations and the time required for each iteration. This should conclude if the network is efficient enough if one wants to train it within a reasonable time frame, or if the algorithm within the network is too computationally expensive for our particular case. Limitations could be explored in a similar fashion. The evaluation of the result will be based on observations. Every observation will tell whether the RL meets our expectations, that is to say if the RL is considered effective in solving problems with similar context or not. Problems with similar context could be avoidance of pedestrians when designing algorithms for self driving cars where MC-learning is not suitable. Comparison of observations will be made by comparisons of "scores" or roughly speaking the magnitude of e.g survival time.

6 Time plan

The project will be carried out according to the time plan specified in the Gantt schedule 1 below.

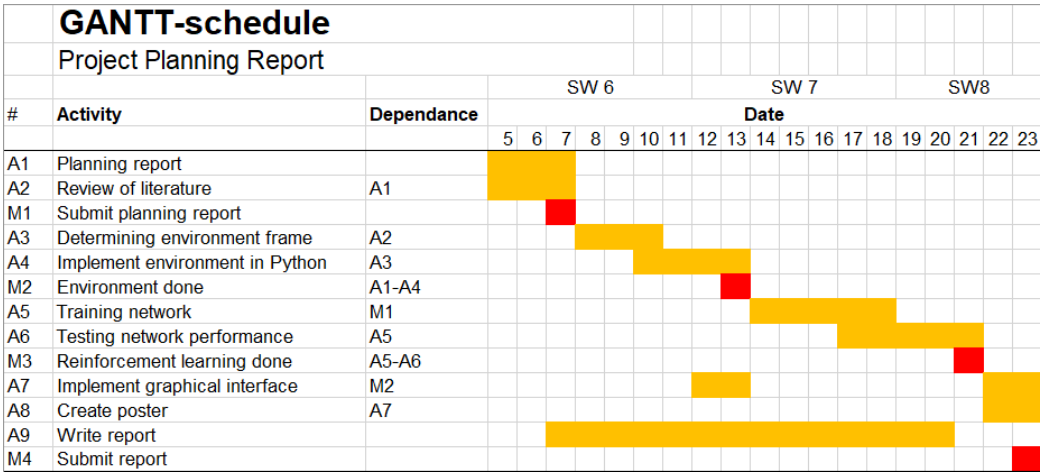

Figure 1: Gantt schedule for project planning