SUMMARY OF MASTER THESIS
to the "Computer Informatics" department, Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski"

Topic: Solving a continuous-state object reinforcement learning task with the NEST neurobiological simulator

Graduate: Borislav Stoyanov Markov, master's program "Artificial Intelligence", faculty number 0MI3400048
Research supervisor: Prof. Petya Koprinkova-Hristova, Institute of Information and Communication Technologies (IICT), BAS

Neurobiology is increasingly gaining momentum in the world of artificial intelligence. There is a growing body of research into the functioning of nerve cells that has led to the creation of biologically valid spike timing models of neurons, as well as much knowledge about the structural organization and decision-making function of the mammalian brain. It has been proven that many decisions are made by the method of Reinforcement Learning. The aim of this thesis is to develop a biologically inspired (spike timing, SNN) neural network model using the NEST Simulator library, which is able to solve an optimization task for object decision making with continuous states from the Gym package. In particular, a CartPole task was chosen, where the goal is to balance a vertical pole for as long as possible. The pole is attached to a car, with movement in two directions - left and right.

In the course of the development, an introduction was made to the main scientific areas of artificial intelligence, necessary to solve the task. A mathematical apparatus for solving reinforcement learning problems is given, as well as a biological rationale for some of the mechanisms for using learning. From the side of reinforcement learning, explained areas are the temporal error (TD) methods, approximation of value functions, how to represent the task as episodical one, as well as the Actor-Critic algorithm. On the neuroscience side, topics covered are Spike-timing-dependent plasticity (STDP), the role of dopamine in error signal processing, volume transmitters of dopamine, the role of the basal ganglia in problem solving and fine motor control in reinforcement learning. Also presented are Winner Takes All (WTA) neuron connections as well as mechanical interpretation of dynamical systems.

Two variants for solving the task with the NEST neurosimulator are considered. The first variant does not give the required results, but still outperforms an agent with a chaotic policy (baseline). The second variant gives a solution that outperforms many other solutions with classical artificial neural networks. Spike diagrams of neurons from the network of different neural units are examined. Also considered are training subvariants when tuning some hyperparameters.

The thesis includes a brief overview of the field of Spike Timing Neural Networks, description of the theoretical formulation, Python code using the NEST Simulator library, and analysis of the results. In the development process of the master's thesis, various parameters of the biosimilar neurons are tested and the solution is illustrated with appropriate visualizations and graphics accompanying the training process.