Automatic Design of Behaviors for Khepera IV Robots Emulating Automotive Vehicles

Amélie Martin

Professor : Alcherio Martinoli
Assistant(s) : Cyrill Baumann

In the field of robotics, automatically generated behaviors can be used to control the actions of a robot in a way that is flexible and adaptable to changing circumstances. This can be useful in situations where it is not practical or feasible to manually specify every action that the robot should take.
In this project, we aim to automatically generate a library of basic behaviors for Khepera IV Robots. We then want to take it a step further and let the robot be able to compose from this library of basic behaviors to adapt itself into a new environment. The combined efficacy of the resulting behavior-based control architecture will be tested in a realistically simulated traffic scenario using Khepera IV robots to represent motor vehicles.

After some literature research, we found that our aim is similar to the Dynamics-Aware Unsupervised Discovery of Skills (DADS) algorithm, recently proposed in [1]. This unsupervised Reinforcement Learning Algorithm would allow us to be able to learn diverse and predictable skills in a given environment. Then, the agent would be able to compose these skills to adapt itself in a new environment. Here, to allow the policy to be exploratory and the skills to be predictable, the intrinsic reward function developed in DADS encourages discovery of 'predictable and 'diverse' skills. Therefore, it is high if:

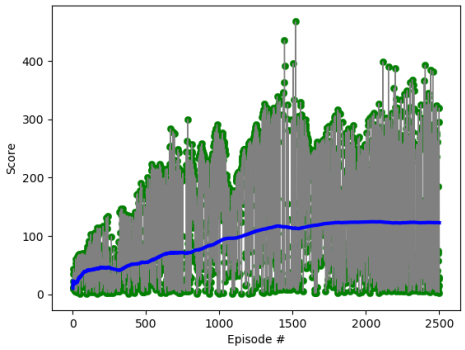
1. the changes in the environment are different for different skills (encouraging *diversity*)
2. the changes in the environment for a given skill are predictable (encouraging *predictability*).

Unfortunately, after studying the code provided with the algorithm we were left with a two main reserves:

- the algorithm might not be the most resistant to noise, as in our case, we would like to apply it to several agents and not only to one as it was done in [1].
- the computation cost of this algorithm and the time it would take to run on our simulations to learn a basic behavior as none of this information are specified in the original paper.

Therefore, we decided to implement a simple reinforcement learning algorithm (the Reinforce algorithm), with an arbitrary reward to begin with and then use the theory of the DADS algorithm to implement a new reward function that would allow us to automatically generate diverse and predictable skills.

After setting up our environment for our simulation, we defined and built our Class Policy and the Reinforce Algorithm that had to be adapted to fit with our given environment. After training for one basic behavior, we obtained the following results:



Resulting score after training our agent for 2'500 episodes

Due to time constraints, we didn't have the time to finish like we would have wanted to this work and stopped before adding the DADS computation reward to our Reinforce algorithm. Therefore, our agent now knows how to learn policies with respect to a reward but, unfortunately, the other main objectives of the project have not been achieved.

[1] Dynamics-Aware Unsupervised Discovery of Skills, A. Sharma et al., ICLR, 2020