Playing Object-Centric Atari Games with Reinforcement Learning

Hector Kohler 2nd Year PhD
hector.kohler@inria.fr

1 Aim of the Project

Recent advances in deep learning allowed the community to learn very good policies for the game of Go, Atari video games, or robot control in simulation.

In these applications, policies are implemented with deep neural networks and a reinforcement learning algorithm is used to learn their parameters. In that case the deep neural network returns in-game actions based on pixel inputs e.g: "if pixel 34 green level ≤ 87, then move right". To help interpretability by a human, the deep reinforcement learning policy should take semantically meaningful game states as input, e.g: "if the character is left of key, then move right". Object-centric representations turn pixel inputs into object-like inputs.

This project aims at benchmarking different reinforcement learning algorithms to play Obect-Centric Atari Games [1] (Fig. 1).
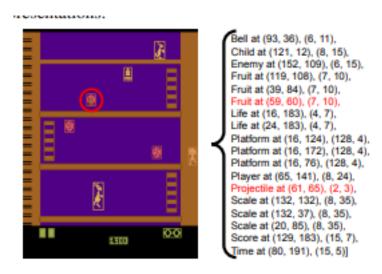

Figure 1: Object-Centric Atari Game

2 Guidelines and Supervision

The student should be familiar with reinforcement learning `https://en.wikipedia.org/wiki/Reinforcement_learning` ,`pytorch`. The student will get familiar with the `stable-baselines3` library that implements most deep reinforcement learning algorithms in `pytorch`. Object-Centric Atari games [1] are implements in `gym` here `https://github.com/k4ntz/OC_Atari/tree/master`. Great attention will be given to the rigor of the experimental process, coding quality, and results presentation with plotting tools such as `matplotlib`. All sources and codes will be given to the students.

The main supervisor of this project will be Hector Kohler, a 2nd year PhD student in the internationally recognised Scool team. As such, he will have plenty of time to supervise and guide the student through in-person meetings and `discord` discussions.

If the student works well, there is a possibility to pursue the research during an internship at Inria.

References

[1] Quentin Delfosse et al. *OCAtari: Object-Centric Atari 2600 Reinforcement Learning Environments*. 2023. arXiv: `2306.08649 [cs.LG]`.