Curiosity driven exploration

Coordinators	Nr. Students
Asist. drd. Ing. Tudor Berariu Andrei Nica (andreic.nica@gmail.com)	1

1. Introduction/Context

From a biological perspective behavior is the result of multiple competing needs that together provide an evolutionary advantage. These needs arise because they provide particular rewards to the organism. One can distinguish between primary rewards which are those necessary for the survival of one's self and offspring, extrinsic rewards, which are conditioned rewards that motivate behavior but are not inherently pleasurable, and intrinsic rewards which are unconditioned rewards that are attractive and motivate behavior because they are inherently pleasurable.

One of the greatest difficulties in reinforcement learning is the fundamental dilemma of exploration versus exploitation: When should the agent try out (perceived) non-optimal actions in order to explore the environment (and potentially improve the model), and when should it exploit the optimal action in order to make useful progress?

One direction of research which wishes to solve part of the problem is intrinsic rewards generation. So far, there are various ways that have beens studied in the literature to equip agents with the drive to explore simulated environments [2]. For example, agents can manifest curiosity depending on how *boring* an environment is, avoiding fully predictable states ([1]). Other agents can be viewed as *information-seeking*. There are studies which analyze the concept of *flow* where an agent tries to maintain a state where learning is challenging, but not overwhelming [3]. The *free-energy principle* states that an agent seeks to minimize uncertainty by updating its internal model of the environment and selecting uncertainty-reducing actions. *Empowerment* is founded on information-theoretic principles and quantifies how much control an agent has over its environment.

Recent vast investment in deep reinforcement learning have led to promising results in simulated environments and games. However, in the current state, it is very challenging to apply RL to complex real world problems. Thus, it is easy to encounter various limitations for agents to efficiently explore even the most basic simulated environments.

Through this research we wish to build upon current state of the art research in *curiosity driven exploration* and investigate questions such as: *what are still some of the basic exploration problems in solving current RL benchmarks and how can we improve? Is it a problem of the information encoded in the state representation? Can we improve it with a population based approach? Or is it the learning process?*

2. Objective

- Review state of the art solutions for intrinsic reward generation in reinforcement learning frameworks.
- Analyse current limitations for exploration in RL and exemplify them on games of two different levels of complexity (e.g. four room problem and Atari game (Mnih et al., 2015)).
- Investigate how curiosity driven exploration can be improved (e.g. with a multi-agent approach).
- Propose and evaluate solution to solve the previously identified limitations.

3. Required and Learned Skills

- Requirements
  - Good knowledge of Python
  - Prior Machine Learning and Reinforcement learning knowledge is preferred
  - Fast learner, proactive mindset, theoretical research
  - Comfortable working in a team

- Learned skills
  - Experience working with RL algorithms and training of deep neural networks
  - Experience working with frameworks such as: Pytorch, Pandas, Opencv
  - Participation in reinforcement learning study groups that review state-of-the-art articles
  - Support and guidance for writing academic research papers

4. References

[1] Pathak, Deepak, et al. "Curiosity-driven exploration by self-supervised prediction." International Conference on Machine Learning (ICML). Vol. 2017. 2017.

[2] Burda, Yuri, et al. "Large-Scale Study of Curiosity-Driven Learning." arXiv preprint arXiv:1808.04355 (2018).

[3] de Abril, Ildefons Magrans, and Ryota Kanai. "Curiosity-driven reinforcement learning with homeostatic regulation." arXiv preprint arXiv:1801.07440 (2018).

[4] Conti, Edoardo, et al. "Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents." arXiv preprint arXiv:1712.06560 (2017).

[5] Ostrovski, Georg, et al. "Count-based exploration with neural density models." arXiv preprint arXiv:1703.01310 (2017).