Internship project, master, 2021

Title: Time-Extended Goals for Language-Conditioned Deep Reinforcement Learning Agents.

Supervision: Pierre-Yves Oudeyer and Cédric Colas

Team: Flowers team, Inria Bordeaux

Duration: 6 months, around January - June 2021.

Keywords: Deep reinforcement learning; language grounding; artificial curiosity; multi-goal RL; sequence learning; recurrent neural networks; Deep Sets; Transformers; Graph Neural Networks

How to apply: contact pierre-yves.oudeyer@inria.fr and cedric.colas@inria.fr with a CV and letter of motivation.

Context: Deep Reinforcement Learning is a family of methods that leverage new advances in Deep Learning to tackle problems where a learning agent interacts with its environment so as to maximize a cumulative measure of reward. Such algorithms were successfully applied to a variety of domains and achieved impressive results (e.g. in chess, Go, Atari games, but also autonomous cars (Arulkumaran et al., 2017)). However, such agents are often trained on a very specific task, using learning signals provided by a hand-defined reward function externally provided by researchers.
Recently, complementary approaches have emerged. Multi-goal RL, for example, aims at training agents to solve a set of tasks instead of a single one (Schaul et al., 2015). Intrinsic motivations train agents to be autonomous, to generate their own internal learning signals (e.g. Colas et al., 2019). Finally, language-conditioned RL proposes to represent tasks by instructions in natural language, which strongly facilitate the instruction and explainability (Luketina et al., 2019). In this project, we aim at extending the set of goals that learning agents can target to time-extended goals: goals that require an extended period of time to be achieved.

Project:
We are interested in artificial autonomous learning agents that generate, represent and pursue their own goals. The term “goal” can be understood as a “problem” the agent has to solve (e.g. opening the door, finding a red object, etc). Each goal comes with a *goal-achievement function* (or reward function), that provides feedback to the agent, estimating whether the goal is reached (binary feedback), or how close it is from being reached (continuous feedback). Autonomous agents cannot rely on externally-provided goal-achievement functions and need to learn them.

Multi-goal RL methods enable agents to learn about multiple goals at once. In these approaches, the goal space is usually hand-defined and goals are expressed as parameters (e.g. in a 2D Euclidean goal space, goals are x-y coordinates). Recently, Instruction-Following RL proposed to express goals by sentences in natural language. This approach allows experimenters to specify more abstract goals “find a red object”, or “get a sharp object” and represent goals in a more comprehensive way (e.g. Lynch et al., 2020 https://language-play.github.io/).

The IMAGINE approach was developed in the lab to enable artificial agents to imagine new goals in natural language by composing known ones. It was shown to improve the ability of the agent to explore its environment and generalize to new goals (Colas et al., 2020).

Current RL approaches are limited to “snapshot goals”: goals whose corresponding goal-achievement function only requires to observe the current state of the world and of the agent (e.g. “go to the kitchen”, or “open the door”). However, humans often use “time-extended goals”. For these goals, the goal-achievement function needs to consider a trajectory of states, as a single state is not enough to judge the goal achievement (e.g. “jump”, “turn around twice”).

This project aims at extending the set of goals that artificial agents can target by designing new learning architectures to consider time-sensitive goals.

The objectives of this project will be:

- to do a comprehensive review of the relevant literature
- to design goal-achievement functions for time-sensitive goals using new learning architecture (e.g. involving graph neural networks, deep sets, transformers, recurrent neural networks etc).
- to design one or several environment to show the benefits of the proposed approach
- to present the new approach to the community (e.g. through a submission at top-tier conference, open source codebase, blog post etc).

In this project, the student will use state-of-the-art deep learning methods including recurrent networks, transformers, graph neural networks and deepsets. He/she will get familiarized with existing benchmarks (Baby AI, Alfred, Text-World) and will learn to modify them of his/her needs. The project will also involve experimental campaigns on large-scale computing clusters (Jean Zay). This project can lead to a submission to one of the top-tier AI conferences (ICLR, NeurIPS, ICML etc) and, as such, can be a first step in the research career of the student.

Requirements: We are looking for motivated MSc students (Master II). Ideally, he/she has prior experience with Python and deep learning frameworks (Pytorch, Tensorflow). Prior experience in research is, of course, **not** mandatory.

References

- Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep

reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, *34*(6), 26-38.
- Schaul, T., Horgan, D., Gregor, K., & Silver, D. (2015, June). Universal value function approximators. In *International conference on machine learning* (pp. 1312-1320).
- Colas, C., Fournier, P., Chetouani, M., Sigaud, O., & Oudeyer, P. Y. (2019, May). CURIOUS: intrinsically motivated modular multi-goal reinforcement learning. In *International conference on machine learning* (pp. 1331-1340).

- Colas, C., Karch, T., Lair, N., Dussoux, J. M., Moulin-Frier, C., Dominey, P. F., & Oudeyer, P. Y. (2020). Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration. *arXiv preprint arXiv:2002.09253*.
- Lynch, C., & Sermanet, P. (2020). Grounding Language in Play. *arXiv preprint arXiv:2005.07648*.
- Luketina, J., Nardelli, N., Farquhar, G., Foerster, J., Andreas, J., Grefenstette, E., ... & Rocktäschel, T. (2019). A survey of reinforcement learning informed by natural language. *arXiv preprint arXiv:1906.03926*.