Learning to Learn: Meta Reinforcement Learning in Fast and Slow Timescales

Summary: While current reinforcement learning (RL) systems can achieve super-human performance in various domains, they possess two major interconnected drawbacks for real-world applicability: First, they tend to specialize within the task domain that they are trained in, and perform poorly even in only slightly different tasks, usually requiring a whole training process from scratch within the new task domain. Second, mainstream RL systems require large amounts of experience to learn, to the extent that it becomes practically impossible to train in real-world scenarios. The combination of these two factors means that, an RL system deployed in the real-world will need to perform a long and data-intensive training for each task variation it encounters.

Meta-Learning has been proposed and obtained a lot of attention in the recent years as a promising solution to these two drawbacks. The goal of meta-learning can be summarized as "learning to learn:" instead of learning individual specific tasks like mainstream RL techniques, Meta-Learning systems try to learn how to quickly adapt its behavior in any task environment (within a given domain, i.e. with structural similarities) and therefore "learn" new tasks rapidly, with a minimal amount of experience needed. A particular state-of-the-art technique proposes Meta-Learning in two timescales: In the "slow" timescale across multiple episodes, a classical RL algorithm trains a recurrent neural network (RNN), which in turn acts as a learner itself in the "fast" timescale – adapting the agent behavior quickly within an episode. (See Duan et al. (2016) & Wang et al. (2016) for details.)

This project will start by implementing this two-timescale Meta-Learning techniques and experimenting in domains similar to those in the literature. Following the first results, this method will be applied to more complex domains and problems, potentially including simple multi-agent environments as well. Drawbacks and limitations in these domains will be addressed accordingly. (The exact direction of the project will depend on the results obtained on the way.)

Who is this project aimed for?

- Students with interest in machine learning or reinforcement learning techniques, their applications and limitations,
- Students with interest in promising, non-mainstream and "composite" reinforcement learning techniques.
- **Proposed as:** Semester or Master project

Requirements:

- Good understanding of basic machine learning concepts
- Familiarity with reinforcement learning framework (literacy about basic concepts is sufficient)
- Familiarity with recurrent neural networks (literacy about basic concepts is sufficient)
- Familiarity with Python (the more experience the better, but not obligatory)

Note: In case of sufficiently early contact, the student can familiarize himself/herself with the missing required areas before the beginning of the project.

References with brief explanations:

Andrey Kurenkov, "How to fix reinforcement learning", The Gradient, 2018.
An easy-to-read, popular-language article about how current limitations of RL can be addressed, discussing also the importance of meta-learning in this context.

Duan, Y., Schulman, J., Chen, X., Bartlett, P. L., Sutskever, I., & Abbeel, P. (2016). RL2: Fast reinforcement learning via slow reinforcement learning. arXiv preprint arXiv:1611.02779.
The paper that originally proposes the discussed two-timescale meta-reinforcement learning method.

Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., ... & Botvinick, M. (2016). Learning to reinforcement learn. arXiv preprint arXiv:1611.05763.
Further development and evaluation of the two-timescale meta-reinforcement learning method by Google DeepMind.

Li, X., Sun, Z., Xue, J. H., & Ma, Z. (2020). A Concise Review of Recent Few-shot Meta-learning Methods. arXiv preprint arXiv:2005.10953.
A more-general recent review of meta-learning techniques for those interested.

Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. arXiv preprint arXiv:1703.03400.
A meta-reinforcement learning method different from the two-timescale method, to give an example of a different technique.

Contact person
Zeki Doruk Erden | *zeki.erden@epfl.ch*