Internship topic, master level, 2020

Title: Design of an Interactive Demo on Multi-task Deep Reinforcement Learning using Procedurally Generated Environments

Supervision: PY Oudeyer and R Portelas
Team: Inria Flowers, Bordeaux
Duration: 6 months

How to apply: send an email to pierre-yves.oudeyer@inria.fr and remy.portelas@inria.fr with a CV and letter of motivation

Research Context. Training autonomous agents able to generalize an action policy to a multiplicity of tasks is a key desiderata in current Deep Reinforcement Learning (Deep RL) research. This challenge is significantly harder than learning to be proficient at a single problem, in which case a naive single-strategy policy is often sufficient. Towards this objective, several parallel lines of works showcased the benefits of jointly using continuously parameterized procedural generation, to generate a diversity of tasks, and Automatic Curriculum Learning (ACL), to scaffold Deep RL learners within the task space. Our team recently proposed one such ACL algorithm, named ALP-GMM, which is based on tracking the learning progress of its learner on the task space, such that it focuses its task sampling on sub-spaces of appropriate difficulty (Videos and code are available at https://github.com/flowersteam/teachDeepRL). Following this work (in collaboration with a previous Intern) we built a set of procedural Box2D parkour environments that we used to conduct a benchmark of current ACL approaches.

General Public Context. Artificial Intelligence softwares, predominantly in the form of computer vision and natural language processing systems, are now being extensively used for various applications such as virtual assistants or chatbots, shopping, security and surveillance, ... While these fields of AI are quite visible to the general public, the domain of embodied autonomous agents trained using Deep Reinforcement Learning (DRL) algorithms is far more obscure since out-of-the-lab applications are currently anecdotal. As such, providing demonstrations of the field of DRL and its main challenges (e.g. multi-tasking) would constitute an interesting popular science project.

Project. In this engineering internship, the objective will be to design an interactive popular science experiment showcasing the challenge of generalization to unknown tasks for Deep Reinforcement Learning agents. The core idea is to let visitors select parameters encoding the procedural generation of a parkour task (i.e. landscape, obstacles, water lever, …), and to show them live how a Deep RL agent previously trained on similar tasks performs on this never-seen-before situation. The experiment will first be made available online, using in-browser interactions implemented in Javascript. See these online demos for examples: demo1 (https://magenta.tensorflow.org/sketch-rnn-demo, demo2 (https://demos.pragnakalp.com/gpt2-text-generation/)). Then, the experiment will be adapted in a real world setup (e.g. TV +

tactile interface). The Intern will have multiple opportunities to show and test the system in public spaces like science museums. In more details, the objectives will be:

1) **Training a high-performance DRL agent with Automatic Curriculum Learning**
   Which will require to re-use open-access (Github) State of the art DRL algorithms implemented in either PyTorch or Tensorflow (using Python), such as IMPALA (https://arxiv.org/abs/1802.01561), and use one or several of them to train a learning agent on a GPU-powered computing cluster on our Box2D procedural environments.
2) **Designing and implementing an interactive website**
   Which will require to create an interactive website providing sliders & buttons so that visitors can a) select parameters that are used in live to create a parkour landscape and b) "drop" a trained agent on the resulting parkour and visualize in live how this learner deals with the never-seen-before environment. The website, implemented in Javascript, will have to use technologies such as tensorflow.js or ml5js (https://ml5js.org/) to enable the use of in-browser deep reinforcement learning agents. Optionally, the intern could also investigate how to perform in-browser interactive training (e.g. to fine-tune a trained agent on a never-seen-before environment).
3) **Designing and constructing a real-world version of the experiment**
   Once the demonstration site is online, a real-world setup will be designed to render the experiment locally on a screen + ipad. This will require to imagine the setup, order the electronic components, assemble everything and make it appealing. The setup will have to be simple to use (plug-and-play).
4) **Presentation of the setup**
   Once the experiment is completed, the intern will have the opportunity to present its experiment at various public spaces, e.g. during the "Villages des sciences" event of Cap Sciences (https://www.cap-sciences.net/), which attracted more than 3000 people in 2019. In parallel to these short-term presentations, we will also propose to Cap Sciences and other science museums to expose the setup in their collection. Additionally, this internship could also include writing a blog post describing and presenting the experiment.

References:

Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments
Rémy Portelas, Cédric Colas, Katja Hofmann, Pierre-Yves Oudeyer
CoRL 2019 - Conference on Robot Learning, Oct 2019, Osaka, Japan
https://hal.archives-ouvertes.fr/hal-02370165/file/1910.07224.pdf

IMPALA: Scalable Distributed Deep-RL with Importance Weighted Actor-Learner Architectures.

Lasse Espeholt, Hubert Soyer, Remi Munos, Karen Simonyan, Volodymir Mnih, Tom Ward, Yotam Doron, Vlad Firoiu, Tim Harley, Iain Dunning, Shane Legg, Koray Kavukcuoglu

https://arxiv.org/abs/1802.01561