2020-03122 - Post-Doctoral Research Visit F/M High Performance Deep Reinforcement Learning

Type de contrat : CDI  
Niveau de diplôme exigé : Thèse ou équivalent  
Fonction : Post-Doctorant

A propos du centre ou de la direction fonctionnelle

Grenoble Rhône-Alpes Research Center groups together a few less than 650 people in 37 research teams and 8 research support departments.

Staff is localized on 5 campuses in Grenoble and Lyon, in close collaboration with labs, research and higher education institutions in Grenoble and Lyon, but also with the economic players in these areas. Present in the fields of software, high-performance computing, Internet of things, image and data, but also simulation in oceanography and biology, it participates at the best level of international scientific achievements and collaborations in both Europe and the rest of the world.

Contexte et atouts du poste

- Location: Grenoble or Lille  
- Hosting Teams:  
  - Sequel (INRIA Lille): https://team.inria.fr/sequel/  
  - DataMove (INRIA Grenoble): https://team.inria.fr/datamove  
  - Contact: Bruno.Raffin@inria.fr and Philippe.Pesaux@inria.fr  
- Period: to start somewhere by April 2021  
- Duration: 24 months  
- Requirement: PhD in computer Science

Mission confiée

Reinforcement learning goal is to self-learn a task trying to maximise a reward (a game score for instance). The learning process acts by interacting with a simulation code to explore the space of possible states. As an explicit exploration is not possible too large, the key to success is in building an efficient exploration strategy balancing between exploration (test new states), exploitation (replay actions known to lead to high rewards). Using deep neural networks to encode the decision process as lead to significant progress. This is often referred as Deep Reinforcement Learning (DRL). A classical benchmark where DRL thrives are Atari games. The most visible success of DLR is probably AlphaGo Zero that outperformed the best human players (and itself) after being trained without using data from human games but solely through reinforcement learning. The process requires an advanced infrastructure for the training phase. For instance AlphaGo Zero trained during more than 70 hours using 64 GPU workers and 19 CPU parameter servers for playing 4.9 million games of generated self-play, using 1,900 simulations for each Monte Carlo Tree Search.

The general workflow is the following. To speed up the learning process and enable a wide but thorough exploration of the parameter space, the learning neural network interacts in parallel with several instances of actors, each one consisting of a simulation of the task being learned and a neural network interacting with this simulation through the best winning strategy it knows. Periodically the actor's networks are being updated by the learned neural network. This workflow has evolved through various research works combining parallelisation, asynchronism, replay buffers and learning strategies (SORIL, A3C, IMPALA, ...).

Latest developments have shown that massive parallelism is a key enabler to address more complex problems. The RLlib framework is designed to automatically distribute RL environments at scale. Google/Deepmind recent announcement of the Menger framework goes in the same direction.

The goal of this postdoc is to investigate novel training strategies to learn more rapidly and more complex tasks (multiple heterogeneous tasks at once, non deterministic games, simulations of complex industrial or living systems) relying on massive parallelism to enable. This postdoc is very flexible on the directions it can take. We expect that the candidate bring its own experience and interest to adapt it to the research aspects of Sequel. We expect that the candidate bring its own experience and interest to adapt it to the research aspects of Sequel.

This work will be performed in close collaboration in between the Sequel INRIA team specialised in DRL (https://team.inria.fr/sequel/) and the DataMove team specialised in HPC (https://team.inria.fr/datamove). Datamove and Sequel are involved in an INRIA group focused on the convergence between HPC, AI and Big Data (https://project.inria.fr/hpcbigdata/). The candidate will participate to that group too.

The Sequel team is leading research group on reinforcement learning, either deep or not, ranging from theoretical aspects to applications. For instance Sequel organised the International Summer School on RL in 2019 (https://rlss.inria.fr/). Among other projects, Sequel has collaborated with Mila (https://rlss.inria.fr/).
(Montreal) to design and develop the Guesswhat?! experiment (https://guesswhat.ai/). As early as 2006, Sequel worked on go game and designed the first go program (Crazy Stone) able to challenge a human expert player (https://www.remi-coulom.fr/CrazyStone/).

Datamove has a long experience on high performance computing and data analytics (https://hal.archives-ouvertes.fr/hal-01221186). Datamove is also developing the Melissa (https://melissa-sa.github.io/) solution to manage large ensembles of parallel simulations and aggregate their data on-line in a parallel server. Melissa stands out by its flexibility, efficiency and resilience. Melissa enabled to run tens of thousands of simulations on up to 30 000 cores. Melissa as been used for computing statistics, train deep surrogate models. We expect it to be a sound base for a DRL workflow.

References:
Google Menger: https://ai.googleblog.com/2020/10/massively-large-scale-distributed.html
AlphaGoZero: https://deepmind.com/blog/alphago-zero-learning-scratch/
TensorFlow: https://www.tensorflow.org/
Gorilla: https://arxiv.org/pdf/1507.04296
A3C: https://arxiv.org/abs/1602.01783
Rainbow: https://arxiv.org/abs/1710.02298
Impala: https://arxiv.org/abs/1802.05681
Melissa: https://hal.inria.fr/hal-0167479v1

Principales activités
We are looking for a candidate with a PhD either in deep learning, reinforcement learning or high performance computing (a combination of these expertise would be ideal) for a 24 month contract at INRIA. The candidate will have the possibility to join either the Sequel team at Lille or the Grenoble Team at Grenoble.

The postdoc will have access to large supercomputers equipped with multiple GPUs for experiments. We expect this work to lead to international publications sustained by advanced software prototypes.

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération
Salary: 2 653 € gross/month.

Monthly salary after taxes: around 2 136,39 € (medical insurance included, income tax excluded).