**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 (play 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 is ATARI games. The most visible success of DLR is probably AlphaZero 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 AlphaZero trained during more than 70 hours using 64 GPU workers and19 CPU parameter servers for playing 4.9 million games of generated self-play, using 1,600 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 actors’ policies are updated by the learned neural network. This workflow has evolved through various research works combining parallelisation, asynchronism, replay buffers and learning strategies (GORILA, A3C, IMPALA, ...).

Latest developments have shown that massive parallelism is a key enabler to address 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 how to combine massive parallelisation and 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). This postdoc is very flexible on the directions it can take. We expect that the candidate bring its own experience and view on these topics. Focus can address (not limited): 1) addressing middleware and system issues in deploying and running very large scale DRL 2) developing very novel parallelisation strategies for some of the DRL components (replay buffer, model/data parallel training) 3) application of DRL as an adaptive strategy for smart parametric space exploration for ensemble run based scenarios like data assimilation, hyperparameter search, uncertainty quantification 4) developing improved or novel learning rules specifically designed for large scale where lossening synchronisation requirements are critical

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

The SCOLL (formerly Sequel) team is leading research group on reinforcement learning, either deep or not, ranging from theoretical aspects to applications. For instance SCOLL organised the international Summer School on RL in 2019 ([https://rlss.inria.fr](https://rlss.inria.fr)). Among other projects, SCOLL has collaborated with Mila (Montréal) to design and develop the Guesswhat?! experiment ([https://guesswhat.ai](https://guesswhat.ai)). As
early as 2006, SCOLL 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 has 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/
Impala https://arxiv.org/abs/1802.01561
Rllib: https://ray.readthedocs.io/en/latest/lib.html
Melissa: https://hal.inria.fr/hal-01607479v1

Principales activités

We are looking for a candidate with a PhD either related to high performance computing, deep learning, reinforcement learning (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 SCOLL 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 up to two days per week 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).