Mission confiée

Reinforcement learning goal is to self-learn a task trying to maximize a reward (a game score for instance) interacting with simulations. Recently, researchers have successfully introduced deep neural networks enabling to address more complex problems. This is often referred as Deep Reinforcement Learning (DRL). DRL managed for instance to play many Atari games. The most visible success of DRL 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 and 192 CPU parameter servers for playing 4.9 million games of generated self-play, and 3,600 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 actor neural networks are being updated by the learned neural network. This workflow has evolved through various research works combining parallelization, asynchronism and novel learning strategies (Gorila, A3C, IMPALA, ...

The goal of this postdoc is to push forward the scalability of these approaches, and to proposing novel learning strategies to learn more rapidly and more complex tasks (multiple heterogeneous tasks, non-deterministic games, simulations of complex industrial or living systems).

This work will be performed in close collaboration in between the Sequel INRIA team specialized in DRL (https://team.inria.fr/sequel/) and the DataMove team specialized in HPC (https://team.inria.fr/datamove). DataMove has developed the Melissa (https://guesswhat.ai/) solution to manage large ensembles of parallel simulations and aggregate their data on-line in a parallel server. Melissa enabled to run thousands of simulation on up to 30 000 cores. So far Melissa was used to compute advanced statistics. But we expect this framework to be a sound base for a DRL workflow. The Sequel team has strong activities in reinforcement learning, either deep or not, ranging from theoretical aspects to applications. Among other projects, Sequel has collaborated with Mila (Montreal) to design and develop the GuessWhat? experiment (https://guesswhat.ai/). As early as 2006, Sequel worked on go and designed the first go program (Crazy Stone) able to challenge a human expert player.

Informations générales

- **Theme/Domaine**: Calcul distribué et à haute performance
- **Located in**: Grenoble or Lille
- **Industry/Field**: Engineering,

References

- **AlphaZero**: https://deepmindblog.com/blog/alphago-zero-learning-scratch/
- **TensorFlow**: https://www.tensorflow.org/
- **Gorila**: https://arxiv.org/pdf/1802.04215
- **Impala**: https://arxiv.org/abs/1602.01561
- **Elf**: https://arxiv.org/abs/1707.00787
- **Mikela**: https://hal.inria.fr/hal-01607479v1

Principales activités

- **Requirements**: PhD in computer science.
- **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.PreuX@inria.fr
- **Period**: To start somewhere in 2019
- **Duration**: 24 months

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).