



Offre n°2022-05495

Continuous Time Reinforcement Learning. From Theory to Practice.

Le descriptif de l'offre ci-dessous est en Anglais

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Stagiaire de la recherche

A propos du centre ou de la direction fonctionnelle

The Inria University of Lille centre, created in 2008, employs 360 people including 305 scientists in 15 research teams. Recognised for its strong involvement in the socio-economic development of the Hauts-De-France region, the Inria University of Lille centre pursues a close relationship with large companies and SMEs. By promoting synergies between researchers and industrialists, Inria participates in the transfer of skills and expertise in digital technologies and provides access to the best European and international research for the benefit of innovation and companies, particularly in the region.

For more than 10 years, the Inria University of Lille centre has been located at the heart of Lille's university and scientific ecosystem, as well as at the heart of Frenchtech, with a technology showroom based on Avenue de Bretagne in Lille, on the EuraTechnologies site of economic excellence dedicated to information and communication technologies (ICT).

Contexte et atouts du poste

- Level: Master Level Research Internship (M2) or equivalent (stage fin étude ingénieur)
- Where: Villeneuve d'Ascq
- When: Flexible, somewhere 2022-2023, 4 months minimum
- Financial support: about 500 euros/month (gratifications de stage)
- Team: [Scool](#)
- Advisers: Alena Shilova (alena.shilova@inria.fr), Philippe Preux (philippe.preux@inria.fr), Bruno Raffin (bruno.raffin@inria.fr)

Mission confiée

Principales activités

Reinforcement Learning in the recent years has attracted a lot of attention. Deep RL managed to beat human or even expert performance in such tasks as atari games and GO. Unlike classical machine learning, RL helps to train an agent capable of taking decisions based on the state of the environment the agent is in. It is an attractive function that leads to multiple applications in numerous domains.

One of the many applications of RL is control tasks. The simplest ones such as CartPole or Pendulum are the common testbeds for new RL algorithms, while more complex ones are of interest for robot learning. Those tasks are usually considered in the discrete time settings, which on the one hand simplifies the problem so that state-of-the-art RL algorithms can be applicable and on the other hand leads to suboptimal control related to the regularity of decision making process. Nevertheless, there are some problems for which it is necessary to be able to take decisions at the arbitrary moments of time or at high frequency, e.g. high frequency stock trading, autonomous driving and snowboard riding.

Continuous Time Reinforcement Learning (CTRL), compared to Discrete Time Reinforcement Learning (DTRL), deals with the continuity of the problem. In this context, the dynamics of the system are expressed as a PDE (Partial Derivative Equation) for deterministic environments and SDE (Stochastic Derivative Equation) for stochastic environments. The value function (a useful measure to estimate the quality of a policy of actions) can be found from Hamiltonian-Jacobi-Bellman equation that replaces Bellman equation in discrete time. Despite promising performance on simple use cases [1,2,3,4,5], CTRL methods do not match the performance of DTRL algorithms in general case. There are several challenges that prevent CTRL from further scaling:

- increased computational complexity of algorithms
- Require a model for dynamics for training
- exploration becomes even more difficult

But the emerging trend of SciML that intends to combine neural networks and PDE/SDE, like physics informed neural networks [6,7] or Neural ODEs [8] are bringing new tools to address CTRL.

The objective of this internship is to develop a basic environment for CTRL with a few classical use-cases (CartPole, Pendulum, Acrobot, Swimmer), test some promising strategies for CTRL and test some possible improvements.

Work:

- Understanding and implementing different algorithms from CTRL [3,4,5]
- Testing them on the continuous time environments, such as CartPole, Pendulum, Acrobot, Swimmer
- Improving existing strategies in different directions: learning of value function, exploration techniques, stability of the methods

References

1. Doya, Kenji. "Reinforcement learning in continuous time and space." Neural computation 12.1 (2000): 219-245. <https://direct.mit.edu/neco/article-abstract/12/1/219/6324/Reinforcement-Learning-in-Continuous-Time-and>
2. Munos, Rémi. "A study of reinforcement learning in the continuous case by the means of viscosity solutions." Machine Learning 40.3 (2000): 265-299. <https://link.springer.com/content/pdf/10.1023/A:1007686309208.pdf>
3. Yildiz, Cagatay, Markus Heinonen, and Harri Lähdesmäki. "Continuous-time model-based reinforcement learning." International Conference on Machine Learning. PMLR, 2021. <http://proceedings.mlr.press/v139/yildiz21a/yildiz21a.pdf>
4. Lutter, Michael, et al. "HJB optimal feedback control with deep differential value functions and action constraints." Conference on Robot Learning. PMLR, 2020. <http://proceedings.mlr.press/v100/lutter20a/lutter20a.pdf>
5. Lutter, Michael, et al. "Value iteration in continuous actions, states and time." arXiv preprint arXiv:2105.04682 (2021). <https://arxiv.org/pdf/2105.04682.pdf>
6. NVIDIA Modulus - Physics Informed Neural Networks https://docs.nvidia.com/deeplearning/modulus/user_guide/theory/phys_informed.html

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.)
- 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

about 590€ gross per month (internship allowance)

Informations générales

- **Thème/Domaine** : Optimisation, apprentissage et méthodes statistiques
Calcul Scientifique (BAP E)
- **Ville** : Lille
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** : 2023-02-01
- **Durée de contrat** : 6 mois
- **Date limite pour postuler** : 2022-12-22

Contacts

- **Équipe Inria** : [SC00L](#)
- **Recruteur** :
Shilova Alena / alena.shilova@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de

nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

CV + cover letter

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.