

Fiche de stage

Sujet du stage : Méta-apprentissage pour l'exploration divergente avec récompenses rares

Encadrant·e : Salehi, A. , Doncieux, S et Coninx, A.

Durée du stage : 6 mois

Niveau d'études souhaité : Master 2, école d'ingénieur 3eme année

Laboratoire d'accueil : ISIR (*Institut des Systèmes Intelligents et de Robotique*), Campus Pierre et Marie Curie, 4 place Jussieu, 75005 Paris.

Personne à contacter

Prénom Nom : Achkan Salehi

Email : salehi@isir.upmc.fr

Envoyer votre candidature par mail, avec [*sujet du stage*] en objet, un CV et une lettre de motivation.

Description du stage (en français)

Sujet du stage : Méta-apprentissage pour l'exploration divergente avec récompenses rares

Résumé :

Les méthodes de qualité-diversité (QD) [2, 3], dont l'objectif est de découvrir des ensembles de solutions diverses sur le plan comportemental, se sont révélées être des outils efficaces pour traiter les récompenses rares dans un contexte d'apprentissage par renforcement. Cependant, elles ont du mal à généraliser à des tâches inconnues, pour lesquelles des politiques doivent souvent être apprises à partir de zéro. L'objectif de ce stage sera d'utiliser le résultat d'apprentissages antérieurs pour améliorer le processus d'apprentissage lui-même, d'une manière qui est étroitement liée aux approches de méta-apprentissage basées sur l'optimisation telles que [4].

Ce stage aborde des problèmes qui sont pertinents pour de nombreuses applications industrielles telles que la manipulation robotique et l'apprentissage ouvert. Il s'inscrira notamment dans le cadre du projet européen Veridream [1] qui implique de multiples partenaires industriels.

Objectifs du stage :

Les questions qui seront étudiées au cours de ce stage sont dans la continuité directe de [6, 5]. Le candidat retenu :

- Sera responsable de l'évaluation des approches existantes, en particulier en ce qui concerne la dynamique des populations d'agents.
- Devra contribuer à la conception et à l'évaluation de méthodes permettant de quantifier et de traiter la variabilité dans le processus d'optimisation.

Sous la co-tutelle de :

En fonction des progrès réalisés sur ces sujets ainsi que des intérêts du candidat, d'autres directions de recherche en relation avec la robustesse et la généralisation seront considérées.

Profil recherché :

Le candidat doit avoir un intérêt marqué pour l'apprentissage machine et être master ou école d'ingénieurs en informatique, en apprentissage machine ou dans des domaines connexes.

Compétences requises :

Solide bagage mathématique, de bonnes compétences en développement et la maîtrise du langage de programmation Python sont obligatoires. Une connaissance pratique de l'anglais est requise, la connaissance du français est appréciée mais pas nécessaire.

Références :

- [1] <https://www.veridream.eu/>.
- [2] A. Cully and Y. Demiris. Quality and diversity optimization: A unifying modular framework. *IEEE Transactions on Evolutionary Computation*, 22(2):245–259, 2017.
- [3] S. Doncieux, A. Laflaquière, and A. Coninx. Novelty search: a theoretical perspective. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 99–106, 2019.
- [4] C. Finn, P. Abbeel, and S. Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pages 1126–1135. PMLR, 2017.
- [5] B. Lim, L. Grillotti, L. Bernasconi, and A. Cully. Dynamics-aware quality-diversity for efficient learning of skill repertoires. *arXiv preprint arXiv:2109.08522*, 2021.
- [6] A. Salehi, A. Coninx, and S. Doncieux. Few-shot quality-diversity optimisation. *arXiv preprint arXiv:2109.06826*, 2021.

Description du stage (en anglais)

Subject: Meta-learning for divergent exploration with sparse rewards

Abstract:

Quality-Diversity (QD) methods [2, 3], whose aim is to discover *behaviorally diverse* sets of solutions to problems, have proven to be effective tools in dealing with sparse rewards in Reinforcement Learning. However, they struggle with generalization to unseen tasks, for which they often need to be trained from scratch. The goal of this internship will be to investigate the use of prior learning experience to improve the learning process itself, in a manner that is closely related to optimization-based meta-learning approaches such as [4].

This internship addresses problems that are relevant to many industrial applications such as robotics manipulation and open-ended learning. In particular, it will be part of the European project Veridream [1] which involves multiple industrial partners.

Internship Objectives:

The questions that will be investigated during this internship are in direct continuity with [6, 5]. The successful candidate will

- be responsible for the evaluation of existing approaches, in particular with regard to the dynamics of populations of agents.

Sous la co-tutelle de :

- contribute to the design and evaluation of methods that quantify and address variability in the optimization process

Depending on the progress made on those subjects as well as the candidate's interests, other research directions in relation to robustness and generalization will be considered.

Required Profile:

The candidate should have a strong interest in Machine Learning and be enrolled in a MSc or engineering school program in Computer Science, Machine Learning or related fields.

Required skills :

A strong mathematical background, good development skills and proficiency in the Python programming language are mandatory. A working knowledge of English is required, knowledge of French is appreciated but not necessary.

References :

[1] <https://www.veridream.eu/>.

[2] A. Cully and Y. Demiris. Quality and diversity optimization: A unifying modular framework. *IEEE Transactions on Evolutionary Computation*, 22(2):245–259, 2017.

[3] S. Doncieux, A. Laflaquière, and A. Coninx. Novelty search: a theoretical perspective. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 99–106, 2019.

[4] C. Finn, P. Abbeel, and S. Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pages 1126–1135. PMLR, 2017.

[5] B. Lim, L. Grillotti, L. Bernasconi, and A. Cully. Dynamics-aware quality-diversity for efficient learning of skill repertoires. *arXiv preprint arXiv:2109.08522*, 2021.

[6] A. Salehi, A. Coninx, and S. Doncieux. Few-shot quality-diversity optimisation. *arXiv preprint arXiv:2109.06826*, 2021.