

Proposition de stage : Meta reinforcement learning

Nathanaël Fijalkow

Charles Grellois

Encadrants : Nathanaël Fijalkow (CNRS, LaBRI, Bordeaux) et Charles Grellois (Aix-Marseille Université)

Localisation : Laboratoire d'Informatique et Systèmes (LIS), Aix-Marseille

Sujet du stage : Ce stage propose d'étudier le problème du choix du modèle dans le cadre du reinforcement learning (apprentissage par récompense). La technique dite de "meta-learning" fait le choix du modèle d'apprentissage : par exemple s'il s'agit d'un réseau de neurones, de son architecture, nombre de couches, etc. L'approche de ce stage est de développer des techniques de meta-learning en apprentissage par récompense en intégrant le choix du modèle dans l'apprentissage lui-même, afin de faire à la fois le choix du modèle et son entraînement.

Thèmes : apprentissage par récompense, meta-learning

Contexte scientifique et état de l'art

Une des principales difficultés d'utilisation des techniques d'apprentissage profond, que ce soit pour l'apprentissage par récompense (reinforcement learning, voir le livre de référence [SB98]) ou pour d'autres scénarios d'apprentissage, est le choix du modèle et son architecture. Le "meta-learning" (développé en particulier par Yoshua et Bengio dans les années 90) désigne un ensemble de techniques pour "apprendre à apprendre", dont une des applications les plus réussies est l'automatisation du choix de modèle. En plus de l'article de référence sur le meta-learning [FAL17], qui servira de base pour le stage, nous recommandons la lecture de cette note de blog :

<https://bair.berkeley.edu/blog/2017/07/18/learning-to-learn/>

La littérature sur le meta-learning et son application aux choix de modèles est très vaste, et s'étend chaque année depuis une demi-dizaine d'années à une vitesse exponentielle. Une approche très répandue en ce moment utilise l'apprentissage par renforcement, voire par exemple pour un des articles fondateurs [BGNR17].

Objectifs du stage

Dans sa grande majorité, les techniques actuelles étudient séparément les deux étapes, choix du modèle et entraînement. L'approche que nous souhaitons développer pour l'apprentissage par récompense intègre les deux étapes en permettant l'évolution – ou mutation – du modèle au cours de son entraînement. Le développement de cette approche pose de nombreuses questions, à commencer par l'évaluation d'un modèle, le transfert d'informations d'un modèle à un autre, et l'existence de métriques pertinentes sur des classes de modèles. Pendant ce stage nous développerons de nouvelles techniques de meta-learning pour l'apprentissage par récompense. Elles seront définies et étudiées mathématiquement, puis mises en pratique et comparées expérimentalement sur des plateformes existantes (OpenAI Gym).

Prérequis

Le candidat devra idéalement avoir des connaissances en programmation et en particulier en programmation Python. Des bases en apprentissage automatique, couplées à une utilisation de bibliothèques Python pour l'apprentissage, seraient appréciables.

References

- [BGNR17] Bowen Baker, Otkrist Gupta, Nikhil Naik, and Ramesh Raskar. Designing neural network architectures using reinforcement learning. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017.
- [FAL17] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70, ICML'17*, page 1126–1135. JMLR.org, 2017.
- [SB98] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning - an introduction*. Adaptive computation and machine learning. MIT Press, 1998.