

Proposition de stage 2014

Apprentissage par renforcement développemental

RESPONSABLE: Michèle Sebag

LABORATOIRE: Laboratoire de Recherche en Informatique CNRS UMR 8623 & INRIA Futurs

ADRESSE: Université Paris-Sud, 91405 Orsay Cedex

MAIL: Michele.Sebag@lri.fr

URL: http://www.lri.fr/~sebag/Stages/Dev_RL.pdf

RÉSUMÉ :

L'apprentissage par renforcement (RL) s'intéresse à l'apprentissage de comportements (appelés politiques ou stratégies) atteignant un but donné. Dans le cadre de la robotique, le RL rencontre la difficulté suivante: si l'apprentissage est effectué sur un simulateur, les politiques optimales sur le simulateur sont très dégradées quand elles sont portées sur le robot physique – ce qui est connu sous le nom de *reality gap*, rupture de réalité. Si l'apprentissage est effectué sur le robot physique, il est long et l'exploration de comportements peut mettre en danger le robot.

Objectif du stage

L'objectif est d'étudier une approche développementale de l'apprentissage par renforcement, considérant une séquence de robots de capacités croissantes, et de coût croissant. La question devient ainsi de transférer et d'adapter la politique apprise sur un robot "petit", sur un robot plus grand. L'objectif est double : accélérer l'apprentissage sur le robot grand; en diminuer le risque.

Approche

Une approche possible consiste à contraindre les actions autorisées et à relaxer graduellement les contraintes. Pour le robot petit les actions ont un impact limité (e.g. le robot va lentement) ; pour le robot plus grand, les actions sont légèrement différentes (le robot va plus vite, il a plus de force) ou le modèle de transition est différent (le robot est plus lourd, etc). En pratique, le stage commencera par considérer une unique plate-forme robotique ; le robot petit est émulé en bridant les actions possibles (par exemple en limitant la vitesse). Dans ce cadre une politique initiale est apprise. Les étapes consistent alors à :

- identifier les états sûrs (ceux qui sont visités par la politique optimale pour le petit robot) ;
- en déduire une estimation de la nouveauté/danger des états ;
- étendre la politique initiale sous contrainte de visiter des états peu dangereux ;
- adapter l'estimation de la dangerosité des états.

Ce stage peut être traité sous l'angle théorique ou sous l'angle empirique. L'étude théorique demande de très bonnes capacités d'abstraction et d'analyse. L'étude théorique demande d'excellentes capacités de programmation (C++).

Référence

- Reinforcement Learning, An introduction. Richard S. Sutton and Andrew G. Barto. MIT Press, Cambridge, MA, 1998.
- Transfer Learning for Reinforcement Learning Domains: A Survey. M.E. Taylor and P. Stone, Journal of Machine Learning Research, 2009.