



Adaptation créative par apprentissage

Antoine Cully¹

Accessit du prix de thèse Gilles Kahn 2016

Antoine Cully a soutenu sa thèse² en décembre 2015 à l'université Pierre et Marie Curie (Paris-VI), sous la direction de Jean-Baptiste Mouret et Stéphane Doncieux, au sein de l'Institut des systèmes intelligents et de robotique (ISIR). Il occupe actuellement un poste d'Associate Researcher au sein du Personal Robotics Laboratory à l'Imperial College London au Royaume-Uni. Ses nouvelles recherches visent à appliquer des méthodes d'apprentissage pour personnaliser et adapter le comportement de robots aux préférences ou spécificités de leurs utilisateurs.



Les robots ont le potentiel d'aider notre société dans de nombreuses situations, par exemple pour chercher des survivants après des catastrophes naturelles ou pour alerter les pompiers en cas de feu de forêt. Au-delà des profondes transformations que la robotique a entraînées dans l'industrie manufacturière, les robots sont susceptibles d'apporter de grands bénéfices pour la société, par exemple en intervenant sur des lieux de catastrophes naturelles, lors de secours à la personne ou dans le cadre de la santé et des transports. Cependant, l'un des obstacles majeurs à leur utilisation en dehors des environnements parfaitement contrôlés des usines, est leur fragilité [MTN⁺08]. Les robots actuels ont des difficultés à faire preuve de créativité

1. <http://www.antoinecully.com/>

2. Manuscrit consultable à l'adresse <https://hal.archives-ouvertes.fr/tel-01265957/>

lorsqu'ils doivent surmonter un problème inattendu : ils sont limités aux capteurs qu'ils embarquent et ne peuvent diagnostiquer que les situations qui ont été anticipées par leurs concepteurs.

Contrairement aux robots actuels, les êtres vivants ont une impressionnante capacité d'adaptation aux blessures. Ainsi, la plupart des chiens amputés d'une patte sont capables de jouer, sauter, et courir ; et un enfant avec une cheville foulée n'a besoin que de quelques minutes pour trouver une manière de boiter. L'objectif de mes travaux de thèse est de fournir aux robots les capacités d'apprentissage nécessaires afin de leur fournir les mêmes capacités d'adaptation. Nous sommes partis de l'observation que les animaux ne partent pas de zéro pour s'en sortir, ils connaissent leur corps, ils ont de l'expérience et possèdent de bonnes intuitions sur les différentes manières de réagir. L'ensemble de ces connaissances leur permet de choisir intelligemment quelques comportements à essayer et, après quelques tests, ils arrivent à en trouver un qui fonctionne malgré la blessure. Les algorithmes que nous avons développés permettent à nos robots de faire la même chose.

Avant d'être envoyé en mission, le robot utilise une simulation de son corps pour créer une « carte » détaillée des milliers de manières différentes de réaliser sa tâche : cette carte représente les intuitions du robot concernant les comportements intéressants et leur potentiel. Si le robot est endommagé, il utilise ses intuitions pour guider un algorithme d'apprentissage qui réalise des expériences afin de découvrir rapidement un comportement de compensation. Le nouvel algorithme a été baptisé « *Intelligent Trial & Error* » (essai-erreur intelligent).

Lorsqu'il est endommagé, notre robot se comporte comme un scientifique. Il a des *a priori* à propos des différentes actions qui pourraient fonctionner et il commence par les tester. Cependant, ces *a priori* viennent de la simulation du robot intact. Il doit donc trouver celles qui fonctionnent toujours malgré le/les dommages. Chaque action qu'il essaie est comme une expérience qui confirme ou infirme ses hypothèses. Si une action ne fonctionne pas, l'algorithme élimine de manière intelligente des catégories entières d'action pour essayer des choses complètement différentes. Par exemple, si marcher en s'appuyant essentiellement sur les pattes arrières ne fonctionne pas correctement, le robot essaiera alors de marcher en mettant son poids sur les pattes avant.

Nous avons testé notre algorithme sur un robot à six pattes qui réapprend à marcher après plusieurs scénarios de panne (une ou deux pattes arrachées, brisées ou déconnectées), ainsi que sur un bras robot apprenant à placer correctement un objet malgré plusieurs moteurs coincés (14 scénarios différents ont été testés avec ce robot)³. Les résultats montrent que notre algorithme permet à nos robots de s'adapter à toutes les situations de panne en moins de deux minutes, même lorsque la

3. Vidéo disponible à l'adresse : <https://youtu.be/T-c17RKh3uE>

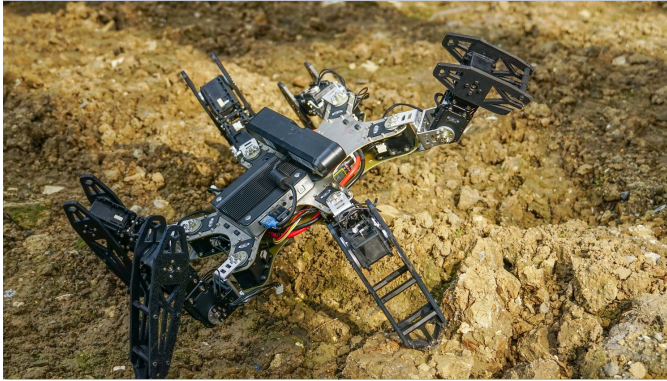


FIGURE 1. Notre robot hexapode apprend de manière autonome comment faire face à des situations de pannes mécaniques grâce à notre algorithme « *Intelligent Trial & Error* ». La photo montre le cas où l'une des pattes du robot est cassée. Pour continuer à avancer malgré la panne, le robot utilise une carte contenant différentes façons de marcher et leur performance respective qui a été générée *a priori* à l'aide d'une simulation physique. Pour s'adapter, le robot sélectionne un comportement qu'il prédit comme étant efficace et le teste en réalité. Si le comportement sélectionné ne fonctionne pas, le robot continue son exploration de la carte en sélectionnant un nouveau comportement qui est à la fois suffisamment différent du précédent et prédit comme étant performant. Ce nouvel algorithme permet à des robots endommagés de continuer leur mission après seulement une minute de temps d'apprentissage et moins d'une dizaine d'essais.

panne est particulièrement sévère, comme lorsque l'on arrache deux pattes à notre robot [CCTM15].

D'un point de vue technique, l'algorithme est divisé en deux étapes : (1) créer la carte de l'espace comportement-performance, et (2) l'adaptation à la nouvelle situation. La carte de la première étape est créée avec un nouveau type d'algorithme évolutionniste appelé MAP-Elites [MC15]. Cet algorithme s'inspire de l'évolution darwinienne pour trouver une large collection de solutions à la fois diverses et performantes. Cet algorithme, ainsi que d'autres travaux de cette thèse ont conduit à l'émergence d'un nouveau domaine de recherche nommé « Optimisation qualité et diversité » [CM13, CM16, PSSS15]. L'adaptation dans la seconde partie s'appuie

sur un algorithme d'optimisation bayésienne [BCDF10], qui exploite les connaissances *a priori* fournies par la carte pour trouver rapidement un nouveau comportement. Les résultats expérimentaux montrent que le point clé de cet algorithme est dans la création et l'exploitation des *a priori*, ce qui permet à notre approche d'être au moins un ordre de grandeur plus rapide que les méthodes d'apprentissage classiques [CCTM15].

Remerciements. Ce travail a reçu l'aide de l'Agence nationale pour la recherche (Creadapt, ANR-12-JS03-0009), de l'European Research Commission (ResiBots, grant agreement No 637972), et de la Direction générale de l'armement (thèse de A. Cully).

Références

- [BCDF10] Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv :1012.2599*, 2010.
- [CCTM15] Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. *Nature*, 521(7553) :503–507, 2015.
- [CM13] Antoine Cully and Jean-Baptiste Mouret. Behavioral repertoire learning in robotics. In *Proceedings of the 15th annual conference on Genetic and evolutionary computation*, pages 175–182. ACM, 2013.
- [CM16] Antoine Cully and J-B Mouret. Evolving a behavioral repertoire for a walking robot. *Evolutionary computation*, 24(1) :59–88, 2016.
- [MC15] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv :1504.04909*, 2015.
- [MTN⁺08] Robin R Murphy, Satoshi Tadokoro, Daniele Nardi, Adam Jacoff, Paolo Fiorini, Howie Choset, and Aydan M Erkmen. Search and rescue robotics. In *Springer Handbook of Robotics*, pages 1151–1173. Springer, 2008.
- [PSSS15] Justin K Pugh, LB Soros, Paul A Szerlip, and Kenneth O Stanley. Confronting the challenge of quality diversity. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 967–974. ACM, 2015.