Apprentissage et adaptation pour des ensembles de robots cooperants

Philippe Lucidarme

▶ **To cite this version:**

Philippe Lucidarme. Apprentissage et adaptation pour des ensembles de robots cooperants. JDOC'03 : Journées des Doctorants, Montpellier, France. pp.P nd. lirmm-00269488

HAL Id: lirmm-00269488

https://hal-lirmm.ccsd.cnrs.fr/lirmm-00269488

Submitted on 3 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

APPRENTISSAGE ET ADAPTATION POUR DES ENSEMBLES DE ROBOTS REACTIFS COOPERANTS

Philippe Lucidarme
Directeur de thèse : Alain Liégeois
LIRMM (UMR 55060)
161, rue Ada, 34392 Montpellier cedex

E-mail : .lucidarm@lirmm.fr

Résumé

Lorsque des robots évoluent en milieu hostile, il n'est pas toujours possible de les téléopérer, et les méthodes tradionnelles de planification sont souvent gourmandes en temps de calcul. Leurs facultés d'apprentissage et d'adaptation doivent donc être de plus en plus importantes. L'objectif est de faire apprendre à des robots une tâche collective complexe sans l'intervention d'un opérateur. Cet article présente deux méthodes d'apprentissage basées sur les algorithmes évolutionnistes et le recuit simulé. Nous avons implémenté ces deux méthodes en simulation et aussi sur une plate-forme composée de 4 robots mobiles présentée dans la troisième partie. La quatrième partie présente l'état actuel des résultats.

Abstract

When a group of robots manoeuvres in an unfriendly environment, teleoperation is not always possible, and traditional methods of planning are generally time consuming. Their learning and adaptation capabilities are constantly increasing. Our purpose is to make robots learn a complex collective task without the need for human intervention. This paper describes two methods based on evolutionist algorithms and simulated annealing. Both ways have been implemented firstly by simulation and secondly on a real platform composed of 4 mobile robots presented in the third part. The fourth part describes the actual results.

1. Introduction

Pour l'agriculture, l'exploration des planètes, ou les applications industrielles et domestiques, il peut être nécessaire de faire coopérer plusieurs robots. Pour coordonner ces groupes de robots, diverses approches peuvent être envisagées. Depuis une dizaine d'année, une nouvelle approche a vu le jour, basée sur une démarche plus proche des systèmes naturels. Les travaux sont basés sur l'émergence de comportements collectifs intelligents à partir d'individus aux capacités limitées [1], [2] et [3]. Ces méthodes sont inspirées du vivant, par exemple une colonie de fourmis est capable de réaliser des tâches complexes, sans qu'aucune d'entre-elles n'ait conscience de la tâche globale. Ce type de systèmes est appelé multi-agent. La réactivité de ces agents leur permet de réagir directement aux variations de l'environnement. Il nous est apparu naturel de s'intéresser à l'apprentissage du comportement de ces agents. Cet apprentissage se fait de façon non supervisée, c'est à dire qu'il n'y a pas de hiérarchie au sein du système. Les robots apprennent par leurs propres moyens, ils sont totalement autonomes, et ne nécessitent pas l'intervention d'un opérateur pendant l'apprentissage.

2. Les algorithmes d'apprentissage

2.1 Description de la tâche

Afin de confronter les performances des méthodes le plus objectivement possible, nous avons décidé de travailler dans un premier temps sur une tâche générique. Pour pouvoir également comparer nos résultats avec des travaux antérieurs, nous avons choisi une tâche d'évitement d'obstacles. Il s'agit, pour un ensemble de robots mobiles à deux roues motrices et équipés de capteurs infrarouges de distances d'apprendre à explorer toute la surface d'un univers physique de manière sûre, c'est à dire en évitant les collisions avec les murs, les obstacles et les autres robots. Pour permettre l'apprentissage autonome de

ces systèmes, les agents doivent être capables d'estimer leur propre performance. Nous avons choisi de maximiser la vitesse moyenne du robot.

2.2 Les algorithmes évolutionnistes

Inspirés de l'évolution darwinienne, les algorithmes évolutionnistes semblent naturellement disposés à apprendre ce genre de tâche. Pour utiliser ces algorithmes, il faut disposer d'une population d'individus. Chaque individu dispose d'une chaîne chromosomique qui régit son comportement. Pour accentuer l'analogie, la structure s'apparente à ce qui nous est inné, son contenu représente ce que nous avons appris depuis notre naissance. Cette chaîne s'apparente à l'ADN dans les organismes vivants. Comme dans les systèmes naturels, des croisements sont réalisés périodiquement et permettent à l'algorithme de créer la génération suivante d'individus. Contrairement à la version proposée par D.E. Goldberg [4] qui rassemble l'ensemble de la population pour créer la génération suivante, nous effectuons les croisements avec des paires d'individus, chaque paire produit deux nouvelles chaînes. Lorsque deux robots se rencontrent, où sont à une distance suffisamment proche pour communiquer, les deux agents s'échangent leurs chaînes chromosomiques et leurs estimations de performance. Deux nouveaux individus logiciels sont ainsi créés par croisement [5].

0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0

Figure 1. Un exemple de chaîne chromosomique.

Tout comme dans les organismes vivants, des mutations sont aussi effectuées. Ces mutations évitent à l'ensemble de la population de converger vers une solution qui ne serait pas optimale. Dans notre cas, la probabilité de mutation est liée à la récompense qu'obtient chaque individu de la part de son univers. Si la récompense est faible, la probabilité de mutation est grande, ce qui permet au système d'explorer un grand nombre de solutions. A l'inverse, lorsque la récompense est grande, les chances de muter sont faibles et le système applique la stratégie qu'il a apprise. Le contenu de la chaîne chromosomique représente la capacité de l'agent à réagir face à différentes situations. Ces situations sont classées et une réaction est associé à chaque cas. L'objectif de l'apprentissage est de créer cette table des liens états-actions. Pour pouvoir affiner le comportement des agents, il peut être intéressant d'augmenter le nombre d'états pour avoir des réactions proportionnelles aux perceptions. Toutefois, la durée d'apprentissage croit avec la complexité du problème. C'est pour cette raison que nous avons étudié une autre méthode d'apprentissage : le recuit simulé.

2.3 le recuit simulé

Pour obtenir des comportements plus fins, et des réactions proportionnelles aux perceptions, nous avons réalisé l'apprentissage à partir de méthodes d'optimisation. Le recuit simulé est basé sur le principe du refroidissement lent d'un matériau en fusion. Il a été généralisé pour trouver des solutions à des problèmes complexes, dont la solution analytique n'existe pas. Par exemple pour trouver le maximum d'une fonction qui est inconnue ou partiellement inconnue. Dans notre cas, nous allons utiliser le recuit simulé pour déterminer les paramètres du système afin de maximiser la récompense obtenue. La fonction de transfert liant les capteurs du robot à ses actionneurs est un réseau de neurone. Comme dans le cerveau humain, l'objectif est de faire évoluer les synapses, c'est à dire les connexions entre les neurones. Dans les réseaux de neurones artificiels, le renforcement de ces connexions correspond à la modification des poids du réseau.

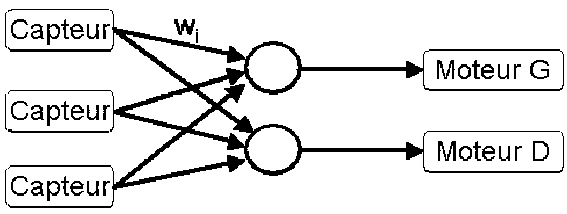

Figure 2. Architecture réactive à base de réseau de neurones.

La figure 2 montre un exemple de réseau réactif utilisé pour l'évitement d'obstacles. Les informations des capteurs influent directement sur les moteurs, et permettent au robot de corriger rapidement sa trajectoire en fonction des obstacles perçus.

3. Plate-forme expérimentale

Afin de valider et de comparer les méthodes proposées, nous avons fabriqué une plate-forme expérimentale composée de 4 robots mobiles conçus et réalisés au LIRMM [6] et [7] :

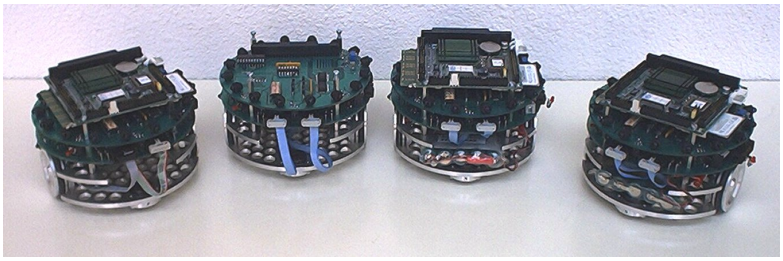

Figure 3. Plate-forme expérimentale.

Leur forme est un cylindre d'environ 13 cm de diamètre et 12 cm de haut. 16 émetteurs et 8 récepteurs infrarouges sont régulièrement espacés sur la périphérie. La portée des communications infrarouges peut être ajustée. Elle est typiquement de 0,5 m. La vitesse maximum a été limitée à 30 cm/s par sécurité pour le matériel. Un PC embarqué (80486 DX avec horloge à 66 MHz) assure le traitement de l'information.

4. Résultats

4.1 Algorithme évolutionniste

La figure 4 montre un exemple de simulation. Les tracés blancs montrent la surface explorée à la fin de l'apprentissage (tous les robots ont la bonne association perception-action). Les petits cercles noirs sont les robots, et les gros sont les obstacles. Nous avons étudié la durée de l'apprentissage en fonction du nombre de robots. Elle décroît jusqu'à 10 robots puis ne change plus. Nous n'avons donc pas intérêt à multiplier le nombre d'individus. Les résultats expérimentaux (figure 5) ont confirmé les résultats de simulation. La convergence est obtenue après seulement quelques minutes.

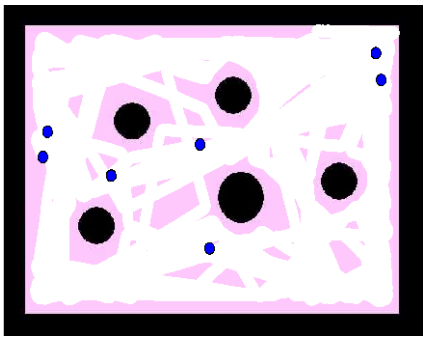

Figure 4. Surface explorée à la fin d'un apprentissage.

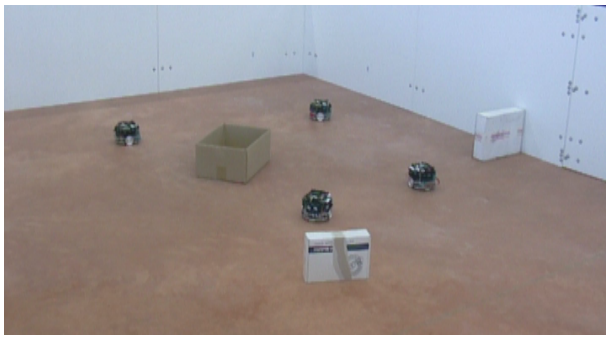

Figure 5. Exemple d'expérimentation.

4.2 Recuit simulé

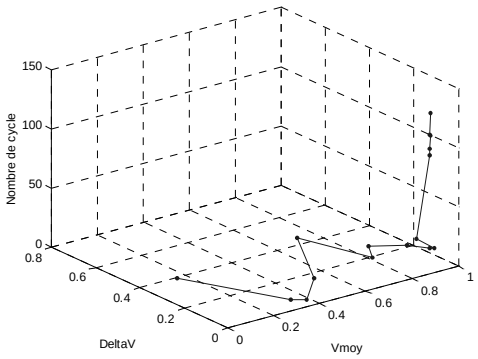

Figure 6: évolution de l'apprentissage.

Les résultats de simulations avec un robot ont montré la convergence de la méthode. La figure 6 représente l'évolution des trajectoires en fonction du temps. Au début des expériences, la vitesse moyenne (Vmoy) est faible, et la différence des deux roues (DeltaV) est grande, cela signifie que le robot tourne en rond sur place. Puis l'effet inverse se produit, la vitesse moyenne augmente et le DeltaV diminue, le robot avance tout en évitant les obstacles.

La méthode basée sur le recuit simulé est actuellement en cours d'expérimentation sur les robots mobiles. Les résultats de simulation montrent que le système est capable de s'extraire des minimums locaux et d'assurer la convergence. On voit sur la figure 6 qu'après une vingtaine de cycles le système n'évolue plus car la solution optimale a été trouvée. En cas de panne le système n'est pas capable de s'adapter, c'est à dire que le recuit simulé n'évoluera plus et les poids du réseau resteront figés. Dans le cas des algorithmes génétiques le système sera capable de s'adapter : la récompense diminuera, la probabilité de mutation deviendra grande, et les cycles d'apprentissage recommenceront.

5. Conclusion et perspectives

Nous avons proposé et expérimenté deux méthodes d'apprentissage pour robots mobiles. Nous avons proposé une version distribuée des algorithmes évolutionnistes et avons montré qu'un nombre d'agents trop important n'apporte plus d'efficacité à la population. La méthode basée sur le recuit simulé est actuellement en cours d'expérimentation sur les robots mobiles, et les résultats de simulation nous encouragent déjà à poursuivre dans cette voie. Pour permettre l'adaptation du système aux pannes et aux changement d'environnement, il peut-être intéressant d'étudier une version qui est capable de recommencer une phase d'apprentissage lorsque les performances sont trop faibles. A terme, proposer une méthode combinant les algorithmes génétiques et le recuit simulé pourrait permettre de tirer avantage du système distribué, et de la rapidité de convergence du recuit simulé.

Références

[1] T. Balch and R. Arkin, "Motor schema-based formation control of multiagent robot teams", Int. Conf. On Multiagent Systems, p. 10-16, 1995.
[2] P. Lucidarme, O. Simonin et A. Liégeois, “Implementation and Evaluation of a Satisfaction/Altruism Based Architecture for Multi-Robot Systems”, Proc. ICRA’02, Washington, p. 1007-1012, 2002.
[3] D. Floreano and F. Mondada, Automatic Creation of an Autonomous Agent : Genetic Evolution of a Neural-Network Driven Mobile Robot, SAB-3, Brighton, p. 421-430, 1994.
[4] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.
[5] P. Lucidarme, A. Liégeois,J-L. Vercher, R. Bootsma, "Un algorithme évolutionniste pour l’auto-apprentissage de groupes de robots mobiles autonomes", Proc. NSI 2002 (CD ROM), La Londe des Maures, article 56, Session Modèles 2, 2002.
[6] P. Lucidarme, P. Rongier et A. Liégeois, “Implementation and Evaluation of a Reactive Multi-Robot System”, IEEE Int. Conf. AIM, Côme, p. 165-170, 2001.
[7] P. Rongier et P. Lucidarme, “A sizing method for a multi-robot system”, IEEE Int. Conf. IROS, Hawaï, p 387-392, 2001.

Remerciements

Une partie de cette étude a été financée par le programme STISS de l'UM2 et par le programme RobEA du CNRS.