

# Un système coopératif pour la simulation comportementale. Application au contrôle d'un bras mobile

Samuel Delepouille<sup>(1),(2)</sup>      Philippe Preux<sup>(2)</sup>

Jean-Claude Darcheville<sup>(1)</sup>

19 mai 2000

## Résumé

Nous présentons MAABAC un système multi-agents coopératif. Les agents du système implémentent l'apprentissage par renforcement, c'est-à-dire la sélection du comportement par ses conséquences. L'organisation des agents du système est coopérative et possède une structure qui permet l'émergence de synchronisation et de coordination inter-agents.

Ce système artificiel — très simple — montre des comportements adaptatifs complexes classiquement observés lors de l'apprentissage naturel.

<sup>(1)</sup> URECA  
UPRES-EA 1059  
Université de Lille 3  
B.P. 149  
59653 Villeneuve d'Ascq Cedex  
France  
delepouille@univ-lille3.fr  
darcheville@univ-lille3.fr

<sup>(2)</sup> LIL  
Université du Littoral  
Côte d'Opale  
B.P. 719  
62228 Calais Cedex  
France  
preux@lil.univ-littoral.fr

# 1 Quels systèmes coopératifs distribués ?

Afin de réduire la complexité des systèmes artificiels, on tente souvent de décomposer ces systèmes en un ensemble de sous systèmes autonomes. C'est sur cette idée qu'est basée l'approche multi-agents (Ferber, 1995). Cette méthode permet de résoudre certains problèmes en utilisant des agents relativement simples qui traitent des informations locales. Cela pose entre autres le problème de la coordination inter-agents.

Pour arriver à un comportement global organisé, il est nécessaire que les actions des différents agents soient organisées et coordonnées. Une solution «classique» pour résoudre ce problème est d'utiliser une architecture centralisée. On définit alors deux types d'agents : des agents périphériques d'une part dont la fonction est essentiellement perceptive ou motrice et des agents centraux qui assurent la communication inter-agents pour aboutir à leur coordination.

Une autre solution est de définir une architecture totalement parallèle. Dans ce cas, il faut connaître les conditions nécessaires à la coordination non supervisée. Ainsi, nous avons mené des recherches sur des dyades de sujets humains (Delepouille, Preux, & Darcheville, 2000). Ces recherches avaient pour but de comprendre et de formaliser l'apparition de phénomènes coopératifs. Nous nous sommes notamment intéressés à l'apparition de la synchronisation entre deux agents en l'absence de toute communication. C'est sur cette organisation, définie en terme de co-dépendance, que nous avons développé MAABAC (Multi-Agents Animat for Behavioral Arm Control ou Animats Multi-Agent pour le Contrôle Comportemental du Bras). Ce système repose sur la sélection des comportements par leurs conséquences. Nous avons utilisé MAABAC pour simuler la dynamique d'un bras basé sur l'activité de ses muscles. L'architecture de MAABAC est totalement décentralisée ; chaque agent musculaire reçoit peu de stimulus de l'environnement et décide de son action essentiellement sur la base d'informations locales.

## 2 Apport de l'étude du comportement

Afin de comprendre l'apparition de comportements coopératifs structurés, nous avons mené des études de laboratoire. Aux cours de ces expériences, des dyades de sujets étaient amenées à agir conjointement dans un environnement partagé. Nous avons montré qu'une organisation complexe des comportements peut apparaître, même si les sujets ne communiquent pas. La condition de l'organisation du comportement est essentiellement la co-dépendance des sujets, c'est-à-dire une organisation coopérative.

L'ensemble des résultats que nous avons obtenus ainsi que de nombreux autres travaux (Mataric, 1995; Mataric, Williamson, Demiris, & Mohan, 1998) laissent penser qu'il est possible de produire des comportements coordonnés et adaptatifs en prenant pour base des agents relativement simples et autonomes les uns des autres, sans forcément recourir à une planification centrale.

### 3 L'architecture MAABAC

Pour réaliser une architecture de contrôle distribuée pour un bras mobile, il nous faut donc respecter deux conditions :

- Au niveau des agents, il faut qu'ils puissent sélectionner leurs comportements en fonction de leurs conséquences. Nous avons donc utilisé l'algorithme du Q-Learning (Watkins, 1989; Watkins & Dayan, 1992), algorithme d'apprentissage par renforcement. Cet algorithme est une bonne implémentation de la sélection du comportement.
- Au niveau de l'organisation, il faut que le renforcement de chacun des agents dépende du comportement global du système. C'est cette organisation coopérative que nous avons retenue.

MAABAC consiste en un bras contrôlé par des muscles et un système visuel. Les systèmes musculaires fonctionnent en parallèle sans aucune commande centrale ni communication. Les agents musculaires utilisent l'algorithme du Q-Learning et chacun tente d'optimiser la commande globale du système.

### 4 Résultats

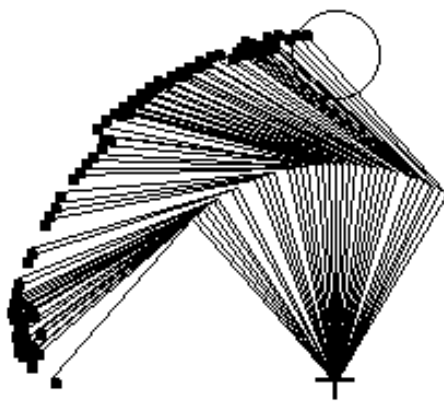


FIG. 1: *Mouvement coordonné. Après apprentissage, un mouvement coordonné du bras est obtenu. On a ici représenté les deux segments du bras et son extrémité.*

Nous portons un grand intérêt à la comparaison entre la dynamique démontrée par MAABAC et le développement des mouvements d'atteinte chez le jeune enfant. Malgré sa simplicité, MAABAC démontre des comportements complexes tels que la coordination du système musculaire pour atteindre une cible (figure 1). Nous enregistrons également des phénomènes émergents très proches de ceux tels observés

classiquement lors de l'apprentissage animal (Staddon, 1984) tels que la généralisation, l'extinction, l'apprentissage par façonnage, . . . .

Nous observons également une évolution dans l'organisation du mouvement. Au départ, des atteintes approximatives sont effectuées avec un haut degré de co-activation musculaire. Au fur et à mesure de l'apprentissage, cette co-activation disparaît et les muscles antagonistes ne sont plus actifs simultanément.

Une conséquence de notre travail est qu'il donne des arguments à l'étude de comportements complexes basée sur des principes sélectionnistes. MAABAC apprend des mouvements d'atteinte sans utiliser aucun système cognitif. Un autre point clé de ce travail est qu'il rend compte du développement du comportement qui commence par des mouvements non organisés. Puis, les comportements s'organisent pour apparaître finalement coordonnés.

Il est important de noter que du fait de notre approche, un mouvement acquis n'est jamais figé mais toujours capable d'évoluer en fonction de l'environnement du système. D'un point de vue «visuel», nous obtenons des mouvements assez «naturels» du fait de la fluidité du mouvement avec ses imperfections qui se gomment progressivement. Notons que ces imperfections jouent un rôle important dans les capacités d'adaptation de MAABAC, tout comme c'est le cas pour les organismes vivants.

## Références

- Delepoulle, S., Preux, P., & Darcheville, J.-C. (2000). Evolution of cooperation within a behavior-based perspective: confronting nature and animats. In *Evolution artificielle 1999, dunkerque*. Springer-Verlag.
- Ferber, J. (1995). *Systèmes multi-agents*. InterEdition, Paris.
- Mataric, M. J. (1995). Designing and understanding adaptive group behavior. *Adaptive Behavior*, 4, 51–80.
- Mataric, M. J., Williamson, M., Demiris, J., & Mohan, A. (1998). Behavior-based primitives for articulated control. In R. Pfeifer, B. Blumberg, J.-A. Meyer, & S. W. Wilson (éditeurs), *From animals to animats 5: Proceedings of the fifth international conference on simulation of adaptive behavior* (pp. 165–170). Cambridge, MA: The MIT Press/Bradford Books.
- Staddon, J. (1984). *Adaptive behavior and learning*. Cambridge: Cambridge University Press.
- Watkins, C. (1989). *Learning from delayed rewards*. Thèse non publiée, King's college, Cambridge, UK.
- Watkins, C., & Dayan, P. (1992). Q-learning, technical note. *Machine Learning*, 8, 279–292.