



**HAL**  
open science

# Modèles de comportements sociaux pour les collectivités d'agents et de robots

Sébastien Picault

► **To cite this version:**

Sébastien Picault. Modèles de comportements sociaux pour les collectivités d'agents et de robots. Système multi-agents [cs.MA]. Université Pierre et Marie Curie - Paris VI, 2001. Français. NNT : . tel-00851693

**HAL Id: tel-00851693**

**<https://theses.hal.science/tel-00851693v1>**

Submitted on 16 Aug 2013

**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.

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS 6

SPÉCIALITÉ : INFORMATIQUE

présentée par

Sébastien PICAULT

pour obtenir le titre de  
Docteur de l'Université Paris 6

---

MODÈLES DE  
COMPORTEMENTS SOCIAUX  
POUR LES COLLECTIVITÉS  
D'AGENTS ET DE ROBOTS

---

Soutenue le 1<sup>er</sup> octobre 2001 devant le jury composé de :

M.	Jean-François PERROT	(Président)
M.	Dario FLOREANO	(Rapporteur)
M.	Joël QUINQUETON	(Rapporteur)
Mme	Kerstin DAUTENHAHN	(Examinatrice)
M.	François MICHAUD	(Examineur)
M.	Alexis DROGOUL	(Directeur)



---

# Remerciements

---

*Je tiens tout d'abord à exprimer mes vifs remerciements aux chercheurs et professeurs qui ont accepté de participer à ce jury :*

*Dario Floreano a accepté le rôle de rapporteur sur cette thèse. C'est pour moi un grand honneur d'avoir éveillé l'intérêt d'un spécialiste de la Robotique Évolutionniste, dont les recherches m'ont beaucoup influencé.*

*Joël Quinqueton, dont j'ai eu plusieurs fois l'occasion d'apprécier les idées originales et la grande ouverture d'esprit, a également rapporté sur mes travaux ; je lui en suis particulièrement reconnaissant.*

*J'exprime également toute ma gratitude à Kerstin Dautenhahn et François Michaud, auxquels je dois, à travers articles et discussions, une partie de la réflexion sur la robotique sociale présentée ici.*

*Enfin, je ne sais trop comment témoigner ma reconnaissance à Jean-François Perrot, qui durant ces quatre ans n'a cessé de me prodiguer ses encouragements et des conseils avisés, m'incitant à toujours plus de rigueur.*

*Cette thèse n'aurait pu se dérouler sans l'initiative d'Anne Collinot, qui m'a plongé avec enthousiasme dans le monde fascinant des primates et des primatologues. Elle a su d'emblée me communiquer son enthousiasme pour la recherche et ses préoccupations méthodologiques.*

*C'est ensuite avec Alexis Drogoul que j'ai poursuivi ce chemin, en passant des primates aux robots. Je lui dois la découverte des mystères de l'intelligence distribuée et un fourmillement de concepts sur lesquels reposent mes travaux. Je ne lui suis pas redevable que de ses idées, mais également des ses qualités humaines, notamment son optimisme contagieux et entier, et de l'autonomie qu'il a su m'accorder dans mes recherches.*

*J'ai trouvé auprès de Jean-Pierre Briot une grande capacité d'écoute à tous points de vue et un soutien moral permanent. Quant à Jean-Daniel Zucker, je ne compte plus le temps que nous avons passé à débattre dans la bonne humeur des sujets les plus divers, que ce soit devant la machine à café ou à Royaumont.*

*J'aimerais également mentionner les chercheurs du LIP6 ou d'ailleurs qui m'ont beaucoup apporté durant cette thèse, que ce soit à travers des exposés ou des discussions de couloir : Jean-Louis Dessalles, Jean-Arcady Meyer, Agnès Guillot, Zahia Guessoum, Dominique Lestel. Merci également à Olivier Sigaud pour sa relecture minutieuse de ce mémoire.*

*Depuis deux ans, des sociologues (Fabrice Rousseau et Albert Piette) ont infiltré l'équipe de MICRobES : je leur dois un questionnement original et un regard extérieur sur ce projet – ce qui m'a forcé à expliciter une grosse partie de ma démarche.*

*L'ambiance chaleureuse qui a régné entre les doctorants et stagiaires de l'équipe MIRIAD/MICRobES m'a permis de surmonter les découragements de la robotique et de bénéficier d'un échange d'idées ininterrompu. Merci donc tout particulièrement à Diane Vanbergue, Angelica Muñoz, Louis Hugues, François Sempé, Guillaume Hutzler, Valérie Renault, Thomas Meurisse, Nicolas Bredèche, Pierre-Emmanuel Viel.*

*Je ne saurais trop remercier Alain Cardon pour son soutien indéfectible et profondément humain, et ses exhortations à toujours « garder foi » en la Recherche. Ses idées, quelquefois incompréhensibles mais toujours stimulantes, ont également beaucoup compté dans l'évolution de mes réflexions.*

*Durant ces années, c'est également auprès de mes vieux compagnons, Frédéric Kaplan et David Servat, que j'ai développé, tout au long de nos discussions commencées entre Lannion et Heidelberg, quelques-unes des idées qui sous-tendent cette thèse, mais aussi bien d'autres qui n'y apparaîtront pas. Notre trio avait conçu EDEN ; pour aboutir à ATNoSFERES, il m'aura fallu quatre ans et une coopération étroite avec Samuel Landau, dont la gentillesse n'a d'égale que le foisonnement de ses idées.*

*Un gros merci également à Dimitri, dont l'amitié a bien souvent allégé mes soucis, sans compter quelques traits d'une plume trempée au VITRIOL.*

*C'est à mes parents que je dois d'avoir pu entamer ces premiers pas dans un parcours de recherche, tant ils ont su éveiller et encourager ma curiosité et mon goût de la réflexion. Ils n'ont cessé de m'apporter leur appui et leurs encouragements, bien que je leur aie parfois semblé planer dans des abstractions bien étranges !*

*Enfin, à Anne, qui m'a supporté avec patience et amour, et sans qui rien n'aurait de sens, merci pour tout.*

# Résumé

Les travaux présentés ici, dans le cadre des Systèmes Multi-Agents (SMA) et de l'Intelligence Artificielle Distribuée (IAD), s'intéressent au problème de l'organisation dans les « systèmes ouverts ».

Dans ce cadre, nos recherches visent à définir des modèles de comportements sociaux permettant aux agents de s'organiser pour s'adapter à leur environnement. Notre démarche fait appel, entre autres principes méthodologiques, au concept de « cercle vertueux » qui se propose d'emprunter des métaphores à d'autres disciplines scientifiques pour concevoir des modèles informatiques.

Dans un premier temps, nous recherchons dans les sociétés animales (en l'occurrence chez les primates) des métaphores de comportements proches de nos besoins. Nous définissons alors des modèles d'agent qui permettent, en simulation, de reproduire une des caractéristiques sociales observées chez les primates, la reconnaissance des relations de dominance. Ces modèles font ensuite l'objet d'une transposition à un domaine différent, la construction collective d'un lexique, pour estimer plus finement les dynamiques collectives sous-jacentes.

Dans un second temps, nous nous intéressons à une expérimentation de « Robotique Collective Ouverte » dans laquelle un groupe de robots doit s'adapter à un environnement où travaillent des humains (projet MICRobES). Nous montrons que dans ces conditions, une simple transposition n'est plus possible et qu'il faut prendre en compte la corporéité des robots. Nous proposons alors pour cela des principes de conception de comportements d'agents faisant appel à la sélection naturelle (l'Éthogénétique) et nous présentons les résultats obtenus avec un framework implémentant ces concepts (ATNoS-FERES).

Nous montrons ainsi comment élargir les principes de départ en conciliant approche multi-agent et algorithmes évolutionnistes, en empruntant des concepts issus de l'éthologie.

# Abstract

## Social behavior models for agents and robots groups

This thesis takes place in the field of Multi-Agent Systems (MAS) and Distributed Artificial Intelligence (DAI). We particularly focus on the issue of the organization in “open systems”.

Thus our research aims at defining social behavior models, making the agents organize and adapt themselves to their environment. Our approach implies (among other methodological principles) the use of a « vertuous circle » that proposes to use metaphors from other scientific fields to design models in computer science.

In the first part, we seek metaphors of behaviors that are close to our needs in animal societies (especially primates ones). We define agent models that are able to produce, in simulation, one of the social features observed in primates, i.e. the ability to recognize dominance relations. We then transpose those models into a different application, the collective building of a lexicon, in order to get indications on the collective dynamics involved in the models.

In the second part, we present an experiment of “Open Collective Robotics” in which a group of robots has to adapt to an human-inhabited environment (MICRobES Project). We show that a simple transposition is not possible any more under these conditions, and that the embodiment of the robots has to be taken into account. Thus, we propose new design principles for agent behaviors involving natural selection (“Ethogenetics”) and present the results we have obtained with a framework implementing these concepts (ATNoSFERES).

We therefore generalize the initial principles by conciliating the specific features of both a multi-agent approach and evolutionary algorithms through ethological concepts.

---

# Table des matières

---

<b>INTRODUCTION</b>	<b>1</b>
<b>1 Présentation de la problématique</b>	<b>3</b>
1.1 Systèmes Multi-Agents et IA Distribuée . . . . .	3
1.2 L'organisation dans un SMA . . . . .	6
1.3 Plan de la thèse . . . . .	9
<b>2 Démarche et outils</b>	<b>11</b>
2.1 Construire pour comprendre . . . . .	11
2.1.1 Modèles et agents . . . . .	11
2.1.2 La simulation multi-agent . . . . .	13
2.2 Éthologie et IAD . . . . .	17
2.2.1 Qu'est-ce que l'éthologie? . . . . .	17
2.2.2 L'éthologie comme référence épistémologique en IAD . . . . .	20
2.2.3 L'IAD comme « éthologie synthétique » . . . . .	23
2.3 « Cercle vertueux » et conception de SMA . . . . .	24
 <b>PARTIE I.</b>	
<b>Modèles de construction de la socialité en simulation</b>	<b>27</b>
<b>3 Les sociétés de primates</b>	<b>29</b>
3.1 Les diverses formes de la socialité animale . . . . .	29
3.2 Les compétences sociales des primates . . . . .	33
3.2.1 Phylogénie des primates . . . . .	33
3.2.2 Spécificités des comportements sociaux des primates . . . . .	33
3.2.3 La reconnaissance des relations sociales . . . . .	36
3.2.4 Les réseaux d'alliance . . . . .	36
3.2.5 Les leurres de puissance . . . . .	37
3.2.6 Tâches collectives et coopération . . . . .	38
3.2.7 La tromperie . . . . .	39

3.3	Débats autour de l'intelligence des primates . . . . .	44
3.3.1	Du primate à l'homme . . . . .	44
3.3.2	Épistémologie et primatologie . . . . .	48
<b>4</b>	<b>Travaux existants</b>	<b>51</b>
4.1	Les approches inspirées de l'IA classique . . . . .	51
4.1.1	Les systèmes de production . . . . .	52
4.1.2	Les scripts . . . . .	53
4.2	Les approches inspirées de l'éthologie . . . . .	56
4.3	Point de départ : MACACA . . . . .	58
4.4	Bilan . . . . .	63
<b>5</b>	<b>Modèles de représentations sociales</b>	<b>65</b>
5.1	La représentation sociale . . . . .	65
5.2	Cadre méthodologique . . . . .	68
5.2.1	CASSIOPEE . . . . .	68
5.2.2	Quelques définitions . . . . .	71
5.2.3	Outils informatiques . . . . .	73
5.3	Modèle proposé . . . . .	77
5.3.1	Représentation de la hiérarchie . . . . .	77
5.3.2	Modèle de perception . . . . .	81
5.4	Expériences . . . . .	82
5.5	Conclusion sur les primates . . . . .	87
5.5.1	Les résultats du modèle et son extension . . . . .	87
5.5.2	Les hypothèses formulées et la méthodologie . . . . .	88
<b>6</b>	<b>La transposition des modèles</b>	<b>91</b>
6.1	Comprendre les dynamiques sociales . . . . .	91
6.1.1	L'étape suivante du « cercle vertueux » . . . . .	91
6.1.2	Approches voisines . . . . .	92
6.2	Langage et socialité . . . . .	94
6.2.1	Point de départ . . . . .	94
6.2.2	Influence de la socialité sur le lexique . . . . .	97
6.2.3	Modèle de départ . . . . .	97
6.2.4	Rétroaction du lexique sur la socialité . . . . .	102
6.2.5	Conclusions . . . . .	105

## PARTIE II.

<b>Modèles de construction de la socialité pour les systèmes situés</b>	<b>111</b>
<b>7 L'émergence de la socialité robotique</b>	<b>113</b>
7.1 Le robot non fonctionnel . . . . .	113
7.2 La robotique collective . . . . .	115
7.3 La robotique « sociable » . . . . .	118
7.4 La « Robotique Collective Ouverte » . . . . .	124
<b>8 Le projet MICRobES</b>	<b>131</b>
8.1 Un projet de Robotique Collective Ouverte . . . . .	131
8.1.1 Objectifs scientifiques . . . . .	131
8.1.2 Sujets de recherche . . . . .	133
8.2 Acteurs et méthodes . . . . .	140
8.3 Quelques détails techniques . . . . .	142
8.3.1 Les robots . . . . .	142
8.3.2 Système embarqué . . . . .	143
8.3.3 Outils communs . . . . .	144
8.4 Difficultés rencontrées . . . . .	144
8.4.1 Problèmes d'origine techniques et implications . . . . .	144
8.4.2 Transposition, simulation et réel . . . . .	146
<b>9 Construire une socialité située</b>	<b>149</b>
9.1 Corporéité et socialité . . . . .	150
9.1.1 De la mise en situation à la corporéité . . . . .	150
9.1.2 Comment construire un corps? . . . . .	154
9.2 Évolution et socialité . . . . .	157
9.2.1 L'Évolution par sélection darwinienne . . . . .	157
9.2.2 Évolution et SMA . . . . .	159
9.3 Conséquences méthodologiques . . . . .	161
<b>10 Modèles d'agents évolutionnistes</b>	<b>163</b>
10.1 La dualité ontogenèse–phylogenèse . . . . .	163
10.2 Vers une « Éthogénétique » . . . . .	165
10.2.1 Principes . . . . .	165
10.2.2 Approches évolutionnistes existantes . . . . .	169
10.3 ATNoSFERES, un framework pour l'Éthogénétique . . . . .	174

10.3.1	Principes d'ATNoSFERES . . . . .	174
10.3.2	Produire un comportement : les ATN . . . . .	176
10.3.3	La construction de l'ATN : traduction et interprétation . . . . .	178
10.3.4	Les classes d'agents . . . . .	182
<b>11</b>	<b>Expériences et extensions</b>	<b>185</b>
11.1	Une expérience élémentaire . . . . .	185
11.1.1	Description de l'expérience . . . . .	185
11.1.2	Premiers résultats et analyse . . . . .	187
11.1.3	Optimalité et adaptation . . . . .	189
11.2	Extensions élémentaires du modèle . . . . .	198
11.2.1	La classe <code>FuzzyATNAgent</code> . . . . .	198
11.2.2	La classe <code>CompositeAgent</code> . . . . .	198
11.2.3	Encapsulation du traducteur . . . . .	199
11.2.4	Encapsulation de l'interpréteur . . . . .	200
11.2.5	La classe <code>BuilderAgent</code> . . . . .	201
11.3	L'Éthogénétique : perspectives . . . . .	204
11.3.1	Expériences en cours . . . . .	204
11.3.2	Vers une plus grande autonomie . . . . .	205
	<b>CONCLUSION</b>	<b>209</b>
	<b>BIBLIOGRAPHIE</b>	<b>217</b>
	<b>INDEX</b>	<b>239</b>

---

# Liste des figures

---

1.1	Organisation et agents dans un SMA . . . . .	6
2.1	Définition de la simulation (1) . . . . .	14
2.2	Définition de la simulation (2) . . . . .	15
2.3	Le « cercle vertueux » en IAD . . . . .	25
3.1	Comportement de toilettage . . . . .	35
3.2	Exemple de tromperie sociale (1) . . . . .	41
3.3	Exemple de tromperie sociale (2) . . . . .	42
3.4	Exemple de tromperie sociale (3) . . . . .	43
3.5	Chimpanzé utilisant un outil . . . . .	45
3.6	Le « test du miroir » de GALLUP . . . . .	48
4.1	Un exemple de script factuel . . . . .	54
4.2	Un exemple de script de règle . . . . .	55
4.3	Un exemple de tâche EMF . . . . .	59
4.4	Mémoire des interactions d'un agent dans MACACA . . . . .	62
5.1	Les liens entre rôles dans CASSIOPÉE . . . . .	70
5.2	Dépendances entre rôles . . . . .	71
5.3	L'outil <i>Task Browser</i> . . . . .	75
5.4	L'outil d'analyse statistique des simulations . . . . .	76
5.5	Représentation de la relation de dominance . . . . .	79
5.6	Évolution des seuils en cas de victoire . . . . .	80
5.7	Évolution des seuils en cas de défaite . . . . .	81
5.8	Linéarité des dominances réelle et formelle dans MACACA . . . . .	83
5.9	Simulation A. Évolution de la linéarité de la dominance réelle et de la dominance formelle . . . . .	83
5.10	Simulation A. Évolution du nombre d'interactions affiliatives et de combats, ainsi que de menaces et de soumissions . . . . .	84
5.11	Simulation B. Évolution de la linéarité réelle . . . . .	84

5.12	Simulation B. Évolution du nombre de combats et de soumissions . . . . .	85
6.1	Temps de convergence en fonction de $M$ , sans socialité . . . . .	100
6.2	Temps de convergence en fonction de $N$ , sans socialité . . . . .	101
6.3	Temps de convergence en fonction de $M$ , avec une hiérarchie . . . . .	102
6.4	Temps de convergence en fonction de $N$ , avec une hiérarchie . . . . .	103
6.5	Temps de convergence en fonction de $\gamma$ (hiérarchie) . . . . .	106
6.6	Temps de convergence en fonction de $M$ , avec une rétroaction . . . . .	107
6.7	Temps de convergence en fonction de $N$ . . . . .	107
6.8	Temps de convergence en fonction de $M$ , dans un système social individualisé . . . . .	108
6.9	Temps de convergence en fonction de $N$ , dans un système social individualisé . . . . .	108
6.10	Comparaison de $T = f(N)$ pour les divers modèles . . . . .	109
6.11	Comparaison de $T = f(M)$ pour les divers modèles . . . . .	109
6.12	Mesure expérimentale de $\mu = f(N)$ pour les divers modèles . . . . .	110
6.13	Mesure expérimentale de $\nu = f(M)$ pour les divers modèles . . . . .	110
7.1	Un match de RoboCup . . . . .	117
7.2	Expressions du robot KISMET . . . . .	121
8.1	Les robots MICRobES dans un couloir . . . . .	132
8.2	Configuration spatiale du site d'expérimentation . . . . .	133
8.3	Carte topologique de l'environnement . . . . .	135
8.4	Représentation spatiale de l'environnement . . . . .	136
8.5	Les robots PIONEER en situation . . . . .	137
8.6	Vues rapprochées d'un robot PIONEER . . . . .	142
8.7	Perceptions visuelles du robot . . . . .	143
10.1	Mécanisme de la synthèse des protéines . . . . .	175
10.2	Un exemple d'ATN . . . . .	177
10.3	Mécanisme d'expression génétique « à pile » . . . . .	178
10.4	Exemple d'interprétation par pile . . . . .	184
11.1	Fitness moyenne (mutations, $r = 1$ %, 50 à 100 bits) . . . . .	188
11.2	Fitness moyenne (insertions/délétions de codons, $r = 1$ %, $p = 20$ %, 50 à 100 bits) . . . . .	189

11.3	L'ATN « optimal » (donnant la fitness maximale avec la structure la plus simple) . . . . .	190
11.4	Fitness moyenne (mutations, $r = 1\%$ , 200 à 300 bits) . . . . .	191
11.5	Fitness moyenne (insertions/délétions, $r = 1\%$ , $p = 20\%$ , 200 à 300 bits) . . . . .	191
11.6	Fitness moyenne (mutations, $r = 1\%$ , 500 à 700 bits) . . . . .	192
11.7	Fitness moyenne (insertions/délétions, $r = 1\%$ , $p = 20\%$ , 500 à 700 bits) . . . . .	192
11.8	Évolution de la fitness (insertions/délétions, $r = 1\%$ , $p = 20\%$ , 200 à 300 bits) . . . . .	193
11.9	Un ATN construit par sélection naturelle . . . . .	195
11.10	Variation de la taille des chromosomes (initialement de 50 à 100 bits) . . . . .	196
11.11	Variation de la taille des chromosomes (initialement de 200 à 300 bits) . . . . .	196
11.12	Variation de la taille des chromosomes (initialement de 500 à 700 bits) . . . . .	197
11.13	Variation de la taille des chromosomes au cours d'une expérience, montrant un cas de diminution (taille initiale de 500 à 700 bits) . . . . .	197
11.14	L'ATN de la classe <code>TranslatorAgent</code> . . . . .	199
11.15	L'ATN de la classe <code>InterpreterAgent</code> . . . . .	201
11.16	L'ATN de la classe <code>BuilderAgent</code> . . . . .	203



---

# Liste des tableaux

---

2.1	Les quatre schémas explicatifs utilisés en éthologie . . . . .	19
3.1	Quelques exemples de groupes animaux . . . . .	30
5.1	Les préconditions d'activation des tâches . . . . .	73
5.2	Les tâches agonistiques du modèle . . . . .	74
10.1	Les catégories de lexèmes . . . . .	181
10.2	Le langage de construction des ATN . . . . .	181
11.1	Lexèmes d'actions et de conditions de la classe <code>Pieton</code> . . . . .	186
11.2	Attributs de la classe <code>TranslatorAgent</code> . . . . .	199
11.3	Lexèmes de la classe <code>TranslatorAgent</code> . . . . .	200
11.4	Attributs de la classe <code>InterpreterAgent</code> . . . . .	201
11.5	Lexèmes de la classe <code>InterpreterAgent</code> . . . . .	202
11.6	Attributs de la classe <code>BuilderAgent</code> . . . . .	203
11.7	Lexèmes de la classe <code>BuilderAgent</code> . . . . .	207



---

# INTRODUCTION

---

*« Il étudia sur le plan notre itinéraire, et nous fûmes perdus »*

L. KLOETZER in (MOUTON 2001)



---

# Chapitre premier.

## Présentation de la problématique

---

*Dans ce premier chapitre, nous allons préciser les concepts clés du domaine dans lequel s'inscrivent nos travaux, celui des systèmes multi-agents (SMA) et de l'Intelligence Artificielle Distribuée (IAD). Nous définirons aussi, dans ce cadre, la problématique scientifique que notre thèse entend aborder, à savoir les mécanismes d'organisation sociale au sein d'un SMA placé dans un environnement dynamique et ouvert à des partenaires humains.*

### 1.1 Systèmes multi-agents et Intelligence Artificielle Distribuée

Le domaine de recherches dans lequel s'inscrit cette thèse est celui de l'Intelligence Artificielle Distribuée (IAD) et des Systèmes Multi-Agents (SMA).

#### **Agents et Systèmes multi-agents**

Un **agent** peut être défini comme une entité informatique *autonome*, située dans un *environnement* qu'elle peut *percevoir* et sur lequel elle peut *agir* (FERBER 1995; DROGOUL 1993). L'ensemble des actions réalisées par un agent dans son environnement constitue son **comportement**.

L'agent a surtout vocation à constituer, avec d'autres, un **système multi-agent** (SMA), c'est-à-dire un *groupe organisé d'agents en interaction*. On attend d'un tel système qu'il manifeste un *comportement collectif*, autrement dit que l'ensemble des actions individuelles des agents dans leur environnement et de leurs interactions produise une structure particulière ou une action globale sur l'environnement.

Il peut s'agir d'une certaine configuration spatiale des agents ou de l'environnement, de la résolution distribuée d'un problème ou d'un ensemble de

## ◇1. PRÉSENTATION DE LA PROBLÉMATIQUE

contraintes (l'agencement des composants électroniques dans une voiture par exemple (FOISEL et al. 1999)), de l'accomplissement d'une tâche qui ne pourrait être effectuée par un seul agent, ou encore de la modélisation d'un système physique ou biologique...

Pour que cette organisation se mette en place, il n'est pas toujours nécessaire que les agents disposent d'une vision globale de l'ensemble du système : une partie de l'approche multi-agent s'attache au contraire à montrer comment des agents au comportement simple, ne possédant de capacités de perception et d'action que locales, sont capables de créer des structures spatio-temporelles complexes – telles que la production en simulation de nids de guêpes (BONABEAU et THERAULAZ 1994) ou d'embouteillages (RESNICK 1994) – ou de résoudre des problèmes difficiles, comme c'est le cas dans l'ÉcoRésolution (DROGOUL et DUBREUIL 1992; DROGOUL et al. 1991).

### Les formes de l'IAD

L'étude et la conception de telles organisations d'agents artificiels, quel que soit leur degré de complexité individuelle, sont l'objet de recherche de l'**Intelligence Artificielle Distribuée** (IAD) (ERCEAU et FERBER 1991), discipline dans laquelle on peut distinguer deux courants extrêmes :

- L'IAD « cognitive », héritière en partie de l'IA classique, s'intéresse à des architectures d'agents utilisant des bases de connaissances symboliques reposant sur des logiques modales, des mécanismes inférentiels, de la planification, des protocoles explicites de négociation, de coordination entre les agents, etc. (RAO et GEORGEFF 1991; JENNINGS 1993).
- L'IAD « réactive » met en œuvre des agents ne disposant que de capacités de perception et d'action locales, sans connaissances explicites *a priori* sur leur environnement.

Cette opposition vient du fait que l'IAD s'enracine dans une *métaphore sociale* bicéphale. La métaphore la plus utilisée actuellement est celle de la *société humaine*, distribuant aux agents l'approche cognitive de l'IA classique, qui vise à reproduire des facultés posées comme spécifiquement humaines (langage, raisonnement, etc.). La socialité, de même, est le produit de normes sociales, de négociations, de la prise en compte explicite de dépendances entre les agents. Elle est seconde par rapport à l'intelligence des agents et elle en procède.

## 1.1. SYSTÈMES MULTI-AGENTS ET IA DISTRIBUÉE

L'autre approche s'appuie au contraire sur le modèle des sociétés animales, et à plus forte raison sur celles dont les membres, pris séparément, semblent les plus éloignés des capacités cognitives humaines (insectes sociaux tels que les fourmis par exemple). La forme d'intelligence à laquelle on s'intéresse se distingue de la précédente en ce qu'elle n'a de sens *qu'au niveau du groupe* d'une part, et qu'elle est d'autre part le résultat d'interactions locales entre les agents.

Ces deux courants mettent l'accent sur des organisations très différentes, qui font chacune appel à des capacités cognitives spécifiques. Les mécanismes par lesquels l'organisation d'agents s'instaure et se manifeste dans un SMA trouvent donc leur réalisation informatique à travers des *architectures* qui se répartissent entre ces deux voies.

On aurait tort de croire que certains problèmes peuvent être « plus naturellement » abordés par l'une ou l'autre méthode. Un même problème, tel que la classique compétition proies/prédateurs (KORF 1991; BOURON 1993), peut être traité avec autant de succès au moyen d'approches fort différentes.

Ceci dit, à l'opposition initiale entre deux types d'architecture d'agent, issues de deux métaphores sociales, s'est superposé un second clivage, fondé quant à lui sur le champ d'application des systèmes considérés.

- Dans l'un des pôles, les SMA doivent *répondre à des spécifications* qui peuvent être établies *a priori* et avec précision, dans le but d'obtenir une fonctionnalité bien identifiable. En règle générale cela n'est possible que dans des cadres suffisamment balisés, dans des « environnements clos » où rien de vraiment imprévisible ne peut se produire. Les SMA sont alors une *façon de programmer une application*, et les problématiques de recherche suscitées par cette approche se rattachent à celles du Génie Logiciel (PARUNAK 1997).
- Dans l'autre, les SMA servent à *construire des modèles* d'interaction entre des entités auxquelles on peut attribuer, au moins de manière métaphorique, les caractéristiques d'un agent (telles que définies plus haut : perception et action dans un environnement). Les buts visés peuvent être la simulation de phénomènes physiques (SERVAT 2000), biologiques ou sociaux réels, l'exploration de dynamiques de groupes dans des « mondes » complètement artificiels (EPSTEIN et AXTELL 1996), ou encore l'immersion distribuée d'artefacts (par exemple des robots ou des « objets communicants ») dans le monde réel.

## ◇1. PRÉSENTATION DE LA PROBLÉMATIQUE

Les travaux que nous présenterons ici se rattachent essentiellement à ce dernier axe, comme nous le montrerons au chapitre suivant.

### 1.2 L'organisation dans un SMA

Dans le cadre général des SMA, l'**organisation** des agents est un concept clef mais encore mal défini. Elle constitue la spécificité d'un SMA par rapport à un système parallèle, en *ajoutant des contraintes* aux comportements individuels des agents, en restreignant leurs possibilités d'action en vue d'une certaine forme de coordination, donnant lieu le cas échéant à des structures spatio-temporelles particulières ou à un comportement collectif intéressant. Ces contraintes sont issues elles-mêmes de fonctionnalités qui apparaissent dans le système du fait de l'activité des agents (cf. figure 1.1).

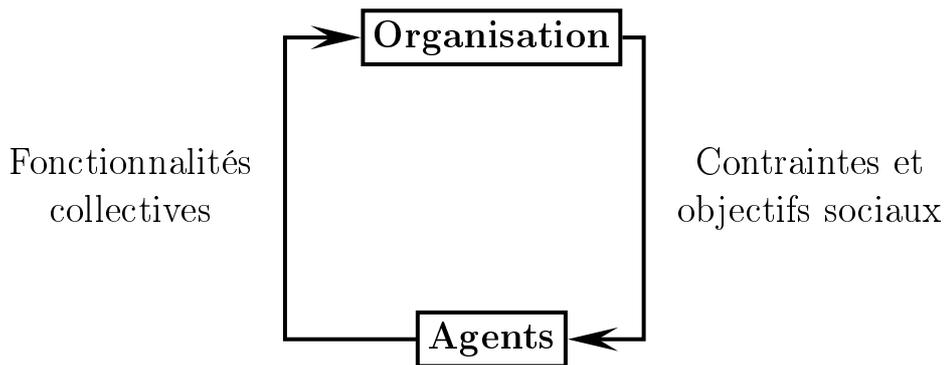


FIG. 1.1 – *L'organisation d'un SMA est le produit de l'interaction des agents, mais elle rétroagit sur les comportements individuels de ceux-ci en imposant un certain nombre de contraintes.* — d'après (FERBER 1995)

Il faut garder à l'esprit qu'un SMA n'est pas qu'un système informatique parallèle : dans ces derniers, les conflits qui peuvent survenir au sein d'un ensemble de processus accédant à des ressources partagées sont d'ordinaire régis par un mécanisme centralisé de synchronisation (comme le font par exemple les systèmes d'exploitation). Dans un SMA, les agents sont des entités *autonomes*, qui doivent gérer de façon décentralisée les conflits pouvant survenir entre eux. C'est le rôle premier de l'organisation, que de *prévenir* ou à défaut *résoudre* les conflits d'accès aux ressources. Par ailleurs, les agents peuvent avoir des objectifs communs et se coordonner ou coopérer pour les atteindre. C'est également dans ce processus qu'intervient l'organisation.

## 1.2. L'ORGANISATION DANS UN SMA

Le problème de l'organisation d'un SMA est généralement abordé au moyen des métaphores sociales correspondant à l'un des deux paradigmes cités ci-dessus. En effet, dans les groupes humains ou animaux, l'organisation du groupe est d'abord une organisation sociale, c'est-à-dire un ensemble de règles partagées par tous, s'appliquant aux interactions entre individus. Les comportements qui sont utilisés par les agents pour établir (explicitement ou non) une organisation dans le système sont donc naturellement appelés « comportements sociaux ».

Ces règles peuvent être fixées explicitement, avec une valeur normative (métaphore sociologique, fortement anthropomorphique), en tant qu'ensemble de connaissances et de contraintes *a priori* qui vont fournir un cadre aux actions des agents, ou au contraire être inscrites implicitement dans les comportements individuels (métaphore biologique) et dans les modes de communication entre agents (QUINQUETON et HAMADI 1999). Dans le second cas, on parlera en général d'organisation « émergente » (JEAN 1997; VISETTI 1996), au sens où le comportement collectif qui peut être observé n'est pas immédiatement déductible de l'ensemble des comportements individuels (il peut dépendre de façon critique du nombre d'agents, de leur environnement, etc.).

Les organisations émergentes, si elles permettent de représenter par exemple le fonctionnement de systèmes complexes existant dans la nature, posent en revanche de sérieux problèmes quant à la conception d'un SMA *en vue* d'une application précise, en dépit des efforts méthodologiques amorcés dans ce sens (MÜLLER 1998). Ceci dit, si l'on est capable de donner au système le moyen de s'organiser lui-même, la forme de cette organisation sera en général particulièrement adaptée à la fonction imposée au système, l'organisation ne faisant alors que refléter des contraintes imposées aux comportements des agents par la fonction souhaitée (BONABEAU et THERAULAZ 1994).

En revanche, les organisations normatives (sauf lorsqu'elles font elles-mêmes l'objet de recherches quant aux mécanismes par lesquels elles s'instaurent et sont construites collectivement par un groupe d'agents), ne peuvent être employées que dans un cadre bien spécifique, prévisible, peu dynamique : car en permettant l'introduction de règles explicites sur les interactions entre agents, elles facilitent sans doute la conception du système, mais rien n'assure que la forme imposée aux relations sociales soit *adaptée fonctionnellement* aux événements pouvant survenir ultérieurement dans l'environnement (AGRE et CHAPMAN 1987).

## ◇1. PRÉSENTATION DE LA PROBLÉMATIQUE

Le débat qui oppose tenants de l'IAD cognitive et de l'IAD réactive, celui qui met en regard l'utilisation d'agents pour la programmation d'une application ou pour la construction de modèles, se retrouve donc également dans la question des comportements sociaux.

### **Position de nos recherches**

Le problème que nous nous proposons d'aborder dans cette thèse est la recherche de modèles d'organisation sociale permettant à un groupe d'agents d'opérer dans un environnement « ouvert ».

Les *systèmes ouverts*, tels qu'ils sont définis par HEWITT (HEWITT et INMAN 1991), sont ceux :

- 1° dans lesquels l'environnement des agents est fortement dynamique et imprévisible ;
- 2° où le groupe d'agents lui-même est ouvert (au sens où des agents peuvent y entrer et en sortir), voire hétérogène (composé d'agents de natures différentes) ;
- 3° en interaction avec l'homme.

Les systèmes ouverts sont amenés dans un futur proche à se développer et à s'insérer dans de nombreuses activités professionnelles ou domestiques. Ce phénomène est déjà patent en ce qui concerne les agents logiciels ; ainsi, un projet comme les *Electric Elves* de M. TAMBE (PYNADATH et al. 2000; SCERRI et al. 2000) vise à intégrer des organisations d'agents « intelligents » (tels que des assistants personnels) dans un groupe humain, les obligeant à une allocation distribuée des tâches à réaliser, et surtout à *ajuster l'autonomie de leurs comportements* vis-à-vis des utilisateurs humains (pour éviter de prendre des décisions jugées intempestives).

Les projets de ce genre s'accompagnent de plus en plus fréquemment d'une distribution des « agents intelligents » sur des objets très différents : ordinateurs personnels classiques, assistants de poche, matériel électro-ménager, etc. Il semble bien que l'intégration d'organisations d'agents dans les sociétés humaines passe dans les prochaines années par une « Intelligence Diffuse » (DROGOUL et SERVAT 2001) qui « ancre » les agents logiciels auparavant désincarnés dans des dispositifs techniques de plus en plus variés.

Plus visibles, les robots, d'abord limités aux machines-outils industrielles, puis devenus un support d'expérimentation de nombreuses techniques informa-

tiques (comme nous le verrons au chapitre 7), commencent déjà à envahir eux aussi notre environnement quotidien.

Ce mouvement a été notamment amorcé par le biais de la « robotique de compagnie », dont le robot-chien AIBO de SONY est une illustration fameuse, et dans de nombreux travaux portant sur la robotique collective (ARKIN 1998; ASADA et al. 1998).

Ce phénomène va conduire à court terme (quelques années au plus) à l'émergence d'une « Robotique Collective Ouverte » dans laquelle des équipes de robots auront à opérer au contact quotidien d'un groupe social humain, non pas en tant que « machines » mais en qualité de *partenaires*.

Il apparaît alors nécessaire de les doter de capacités qui leur permettront de s'insérer dans le groupe humain sans imposer à celui-ci des efforts d'adaptation insupportables. En particulier, il leur faudra manifester des comportements « sociaux » ou « sociables » voisins de ceux qui sont à l'œuvre pour maintenir la cohésion entre individus dans toute société humaine.

Ces contraintes pesant sur les comportements sociaux humains sont telles que nous les projetons volontiers sur toute entité animée, lui attribuant par exemple des intentions (DENNETT 1990), une individualité et un comportement social anthropomorphique. Nous sommes en effet habitués à interagir avec des individus, et nous ne faisons rien d'autre en baptisant les animaux de compagnie que renforcer leur statut d'individu.

Il sera possible sans doute de stimuler et d'encourager cette projection spontanée d'intelligence en dotant les robots de comportements bien choisis. L'individualisation des robots doit par ailleurs leur permettre de construire une organisation sociale qui ressemble plus à celle d'un groupe humain que d'une colonie d'insectes.

Nos recherches visent donc à étudier comment *construire des modèles* de comportements sociaux pour des systèmes multi-agents ouverts et situés dans un environnement social humain. Le projet MICRobES (décrit au chapitre 8), qui vise à immerger un groupe de robots au sein du LIP6, constituera une application particulièrement représentative de ce problème.

## 1.3 Plan de la thèse

Nous commencerons dès le chapitre suivant par un examen des outils conceptuels dont nous ferons usage dans l'ensemble de nos travaux.

## ◇1. PRÉSENTATION DE LA PROBLÉMATIQUE

L'exposé de nos recherches se développera alors selon l'un de ces paradigmes, celui du « cercle vertueux » qui se propose de puiser dans des *phénomènes naturels* (biologiques ou physiques) des *métaphores* susceptibles de conduire à des modèles fructueux.

Dans un premier temps, nous allons donc chercher dans des sociétés animales et notamment celles des primates un certain nombre de phénomènes sociaux jugés intéressants au regard de notre problématique d'IAD (chapitres 3 et 4). Nous proposerons, pour en rendre compte, de modèles qui feront l'objet de simulations au moyen de SMA, pour établir leur capacité à reproduire des mécanismes sociaux naturels (chapitre 5).

Il sera ensuite nécessaire, pour transposer ces modèles à d'autres domaines informatiques, d'en dégager la « substantifique moelle », c'est-à-dire de déterminer les dynamiques qu'ils peuvent produire hors de leur contexte initial, dans d'autres types d'applications, pour d'autres espèces d'agents, toujours en simulation. Nous examinerons pour cela les relations entre représentations sociales et construction d'un langage dans une population d'agents (chapitre 6).

Dans un second temps, nous nous intéresserons au passage à une application ancrée dans le monde réel (tout comme le phénomène qui a été modélisé et simulé) : l'immersion d'un groupe de robots autonomes dans notre laboratoire (« Robotique Collective Ouverte », chapitre 7 et projet MICRobES, chapitre 8).

La poursuite du « cercle vertueux » conduirait à essayer de transposer aux robots les modèles issus des simulations précédentes. Toutefois, nous serons amenés à réviser un certain nombre de nos hypothèses pour tenir compte des spécificités radicales de ces objets ancrés dans le monde physique que sont des robots (chapitre 9).

Par suite nous proposerons d'autres types de modèles pour l'étude de la socialité, inspirés également de phénomènes naturels, en l'occurrence l'évolution darwinienne (il s'agira de « l'Éthogénétique » et du framework ATNoS-FERES, chapitres 10 et 11). Nous montrerons alors comment les limitations du paradigme initial peuvent être surmontées pour conduire à une approche qui, tout en s'appuyant sur la même démarche scientifique, permet d'aborder des problèmes dans un cadre plus large.

---

# Chapitre deuxième.

## D é m a r c h e   e t   o u t i l s

---

*« Ne demande jamais ton chemin à quelqu'un qui le connaît, car tu ne pourrais pas t'égarer ! »*

RABBI NAHMAN DE BRASLAW.

*Nous allons maintenant exposer le cadre méthodologique au sein duquel se sont initialement déroulés nos travaux, la démarche que nous avons choisie pour étudier des modèles de construction sociale, et les outils que nous avons exploités pour cela. En particulier, nous rappellerons les liens étroits qui unissent Éthologie et IAD, liens méthodologiques autant qu'épistémologiques, sur lesquels repose une part importante des principes de modélisation, de simulation et de validation.*

## 2.1 Construire pour comprendre

### 2.1.1 Modèles et agents

Nos travaux de recherche visent à étudier comment des groupes d'agents peuvent s'organiser pour adapter leurs actions à un environnement ouvert (c'est-à-dire essentiellement socialisé et dynamique).

Pour cela, il est nécessaire de faire appel à des **modèles**. Un modèle est « une représentation théorique qui ne prétend pas décrire fidèlement l'objet d'étude, mais qui revendique au contraire son caractère *délibérément schématique*, en même temps que sa *fécondité* eu égard à un objectif spécifié » (SOLER 2000).

## ◇2. DÉMARCHE ET OUTILS

Au sein du paradigme général de l'agent et du système multi-agent, un *modèle d'agent* est une description calculable du comportement des agents dans leur environnement (incluant le cas échéant des représentations, des buts, etc.), à distinguer en principe de la notion d'*architecture d'agent*, qui est le choix de telle ou telle technique informatique pour implanter sur un ordinateur un modèle d'agent donné.

Ces modèles présenteront naturellement un certain nombre de caractéristiques intrinsèques (de nature mathématique par exemple), mais ce n'est généralement pas selon celles-ci qu'ils seront évalués : car dans la mesure où ils représentent les actions possibles d'*entités agissantes*, ils ne prennent tout leur sens, et ne tirent leur valeur, que d'une *mise en œuvre dynamique* (il est d'usage en informatique de parler de *simulation*, mais nous veillerons à préciser ce que recouvre exactement ce terme dans la section 2.1.2).

En cela, un modèle d'agent s'oppose assez radicalement à un modèle mathématique, essentiellement atemporel, dont toutes les propriétés (universelles et nécessaires) peuvent être immédiatement déduites de lois dès qu'on dispose de sa description formelle. C'est d'ailleurs ce qui fait l'un des intérêts d'une mathématisation des modèles, mais aussi ses limites : théorèmes non prouvables (HOFSTADTER 1985; DENNETT 1996, pour un aperçu des conséquences), systèmes d'équations non intégrables (POINCARÉ 1893).

Le modèle d'agent, au contraire, ne fait que décrire des potentialités, des liens plus ou moins sophistiqués entre perceptions et actions, mais le comportement réel, lui, ne peut être observé que dans un *environnement* (un espace dans lequel surviennent des événements, et qui contient d'autres agents). Dès lors, il n'y a guère que dans des situations simplifiées à l'extrême que l'on peut prédire de façon certaine l'évolution du système – dans pareil cas, on peut d'ailleurs douter de la pertinence de l'utilisation d'agents.

La compréhension scientifique des phénomènes existants à l'état naturel ou produits artificiellement par des moyens informatiques sera donc abordée ici au moyen de la *construction de modèles d'agents* et de leur *mise en œuvre expérimentale*.

Il nous faut donc préciser d'une part, *ce qu'est* cette forme d'expérimentation (ce qui sera l'objet des sections 2.1.2 et 2.2.2) ; d'autre part, *comment* construire ces modèles de façon pertinente par rapport aux phénomènes à étudier (cf. § 2.2.1 et § 2.3).

## 2.1.2 La simulation multi-agent

### Expérience et simulation

La fécondité d'un modèle, la validité d'une théorie scientifique reposent, dans le cas des sciences empiriques, sur une confrontation aux faits observables, sur l'« expérience ».

D'un point de vue épistémologique, il convient de distinguer en fait quatre notions :

1. **L'expérience**, en tant qu'ensemble de phénomènes et de données sensibles, autrement dit ce qui se produit dans le monde et que nous pouvons percevoir.
2. **L'expérience scientifique**, qui consiste à « modifier délibérément les conditions de l'enchaînement naturel des phénomènes afin de déterminer quels paramètres concourent effectivement à produire un effet donné » (SOLER 2000) : il s'agit donc ici d'une *méthode* servant à mettre les théories à l'épreuve de l'expérience au sens 1.
3. **La simulation**, qui consiste à « pratiquer des tests *sur un substitut* de la situation réelle [...] (le substitut étant supposé reproduire fidèlement les caractéristiques pertinentes essentielles de la situation à caractériser) » (*op. cit.*). Il peut s'agir de dispositifs physiques ou plus abstraits (informatiques).
4. **L'expérience de pensée**, enfin, qui est une expérience (au sens 2) « *uniquement conçue* (et non pas effectivement [mise] en œuvre), car irréalisable en fait ou en principe » (*op. cit.*).

### Simulation et agents

Le terme informatique de « simulation » est issu de l'utilisation de calculateurs électroniques pour résoudre des modèles équationnels (par exemple dans le cas de systèmes non intégrables où il faut procéder par approximations).

Par la suite, une « simulation », en informatique, désigne tout programme qui se substitue à une expérience, *y compris à une expérience de pensée*.

Prenons quelques exemples. Lorsqu'on modélise le comportement d'entités réelles, qu'elles soient biologiques (DROGOUL et FERBER 1992) ou physiques (SERVAT 2000), il s'agit bien d'une simulation au sens premier : on construit un substitut informatique de la situation réelle, qu'il est possible de paramétrer à volonté pour comprendre quels mécanismes sont impliqués dans

## ◇2. DÉMARCHE ET OUTILS

le phénomène réel. On peut schématiser la méthode scientifique employée au moyen de la figure 2.1.

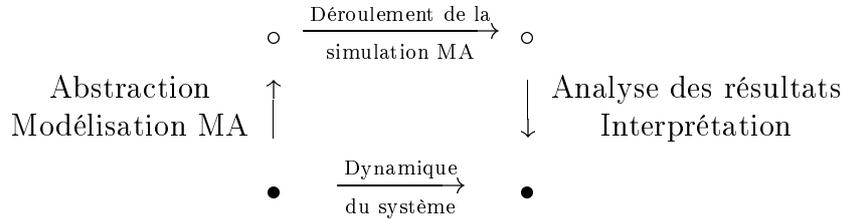


FIG. 2.1 – *Schéma habituellement invoqué pour définir la simulation. Le phénomène étudié est d’abord modélisé sous la formes d’entités abstraites, les agents, dont l’implémentation permet d’observer le fonctionnement dynamique. Les résultats obtenus en simulation doivent alors être confrontés à la dynamique du système réel.*

En revanche, lorsqu’il est question de « modéliser » le comportement de robots devant assurer une fonctionnalité collective (DROGOUL et FERBER 1993), il ne s’agit plus nécessairement de robots réels (ni réalisables) : la situation simulée est fictive, imaginaire, construite *ex nihilo* dans le système informatique. Pour autant, si l’on considère le « monde virtuel » que constitue l’environnement « simulé », c’est-à-dire ce monde artificiel et uniquement informatique, ce qui est réalisé par le chercheur est bien une *expérimentation*, puisqu’on peut faire varier séparément tous les éléments jugés pertinents pour comprendre leur rôle dans le fonctionnement du système.

Autrement dit, le terme de « simulation informatique » recouvre deux notions :

- la simulation au sens classique ;
- l’expérimentation dans un monde virtuel, purement informatique, qui n’est relié au réel que par des métaphores.

La frontière entre ces deux acceptions n’est pas hermétique, elle dépend en grande partie du *niveau d’abstraction* auquel on se place dans la définition du modèle. En effet, la plupart du temps une simulation même au sens originel ne correspond pas au schéma 2.1 mais à celui de la figure 2.2 : les modèles informatiques sont rarement construits directement à partir des phénomènes eux-mêmes mais sont la traduction calculable soit de modèles préexistants, définis par des scientifiques « du domaine », soit de généralisations abstraites à partir d’une classe de phénomènes considérés comme semblables (MEURISSE et VANBERGUE 2001; DORAN 2001).

## 2.1. CONSTRUIRE POUR COMPRENDRE

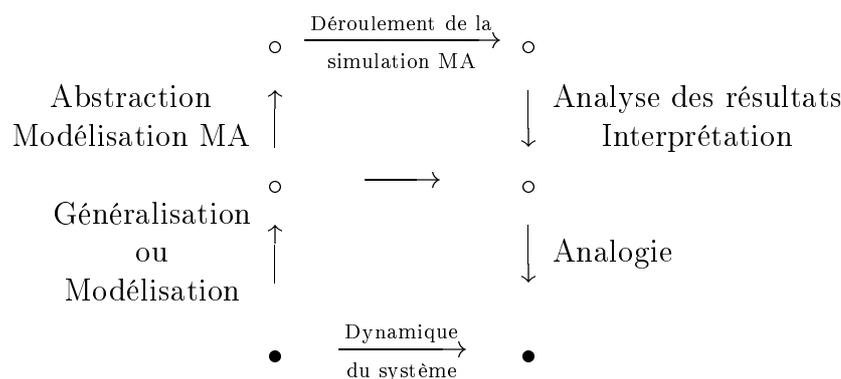


FIG. 2.2 – Schéma qui sous-tend fréquemment les simulations multi-agents. Ce n'est pas le phénomène lui-même qui est modélisé, mais un modèle de celui-ci, qui peut être un modèle construit par les chercheurs « du domaine » (physiciens, biologistes), ou une classe de phénomènes semblables (généralisation).

L'utilisation d'un SMA comme modèle formel pour la simulation n'est donc pas neutre. Elle pose des problèmes spécifiques mais apporte des solutions originales à la question de la modélisation.

Un de ces « problèmes », fréquemment invoqué pour critiquer la simulation par SMA, est celui de l'imprévisibilité des résultats et de la difficulté (voire de l'impossibilité) de leur formalisation. En effet, l'utilisation d'entités fortement couplées (CARDON 2001b) ou dont les actions comportent une part d'aléatoire introduit un indéterminisme dans l'évolution du système. Cela ne signifie pas, sauf cas très particuliers, que sa trajectoire devient *radicalement imprévisible*, mais plutôt qu'elle suit des lois probabilistes. Cela peut se traduire au niveau macroscopique par des lois déterministes<sup>1</sup>, par des attracteurs (PRIGOGINE 1996), par des régularités statistiques. La critique de l'indéterminisme a donc peu de portée : au contraire, elle ne fait que souligner le potentiel des SMA comme outil d'étude de phénomènes non déterministes.

Parfois en revanche, lorsque le but de la modélisation est de tenter une *explication* des phénomènes concernés, il peut être intéressant de s'attacher à une analyse fine des comportements qui provoquent l'évolution du système, plutôt qu'à des résultats globaux. Les SMA constituent là encore un modèle pratique, puisque les comportements individuels, les interactions entre les agents, tous

1. En physique, la cinétique des gaz compte de telles lois pourtant fondées sur une hypothèse d'agitation aléatoire des molécules.

## ◇2. DÉMARCHE ET OUTILS

les événements qui surviennent peuvent être mémorisés, analysés, *observés* soit directement, soit *a posteriori*.

Par ailleurs, l'utilisation d'agents permet de se rapprocher des entités et des interactions rencontrées dans le système réel, ou postulées par la théorie. Cet aspect était déjà amorcé avec la modélisation par objets, mais les agents permettent de tenir compte plus aisément des interactions qui ont lieu dans le phénomène étudié (COQUILLARD et HILL 1997).

Néanmoins, il y a un risque permanent de surestimer les résultats obtenus au moyen d'une simulation multi-agent, lorsqu'elle vise à explorer un certain nombre de propriétés d'un phénomène réel, car il est facile d'oublier le rôle joué par le niveau d'abstraction où se situe le modèle d'agent (EDMONDS 2000).

Plus celui-ci est proche du comportement précis des entités réelles, plus il est possible d'obtenir des simulations « réalistes » (i.e. dont les résultats sont d'une part comparables immédiatement au phénomène modélisé et d'autre part très proches des données empiriques). Cela ne se fait qu'aux dépens de la généralité du modèle qui risque d'être inapplicable à des phénomènes légèrement différents.

Inversement on peut choisir d'opérer directement à un niveau d'abstraction relativement élevé. La simulation peut alors fournir une aide précieuse, sur la base d'analogies dans les dynamiques observées, pour suggérer des explications quant au fonctionnement de systèmes réels dont elle s'inspire.

Enfin, à l'extrême, on trouvera le domaine de la Vie Artificielle (LANGTON 1988) qui se propose de reproduire informatiquement des mondes dans lesquels des systèmes manifestent des caractéristiques semblables à celles des êtres vivants. Les SMA, en mettant en avant la notion d'environnement – un espace doué de propriétés spécifiques, analogues aux lois physiques, contenant des objets et qui est le lieu de la perception et de l'action des agents – et celle de comportement, constituent des outils privilégiés pour la création de ces « mondes artificiels » et l'investigation autour de problématiques biologiques fondamentales pour lesquelles l'expérimentation n'est pas ou n'est plus possible.

L'approche Animat (MEYER et GUILLOT 1990; MEYER et WILSON 1991), en particulier, s'intéresse à une modélisation informatique des capacités adaptatives qui permettent à des animaux naturels ou « artificiels » (les *animats*) de survivre dans leur environnement. Elle regroupe ainsi des problématiques de simulation et de conception autour du concept de « comportement adaptatif »,

en des termes proches de la notion d'agent en IAD.

En retour, les mécanismes « inspirés du vivant<sup>2</sup> » peuvent à leur tour servir au développement d'applications informatiques originales.

## 2.2 Éthologie et IAD

L'IAD réactive, en s'appuyant métaphoriquement sur les sociétés d'insectes pour défendre le concept d'« intelligence collective », a recherché nombre de ses applications, de ses méthodes, de ses modèles dans une branche de la biologie : l'éthologie. Nous allons voir que l'éthologie, en fait, n'intéresse pas seulement l'IAD réactive, mais qu'elle apporte également un cadre méthodologique et épistémologique pour la construction de phénomènes complexes au moyen d'agents.

### 2.2.1 Qu'est-ce que l'éthologie ?

#### Définition et objet

L'éthologie est une branche de la zoologie qui se définit avant tout comme une *étude comparée du comportement* (LORENZ 1984), tout particulièrement dans une perspective phylogénétique.

Cette étude s'applique aussi bien aux diverses espèces animales qu'à l'homme, au moyen des mêmes méthodes, focalisées sur les comportements, donc sur un ensemble de réactions *objectivement observables*.

Elle s'appuie pour cela sur une méthodologie originale (que nous présentons ci-dessous) et s'inscrit dans le cadre de la théorie de l'évolution par sélection naturelle de Charles DARWIN (DARWIN 1859).

#### Deux approches

Konrad LORENZ, considéré comme le principal fondateur de l'éthologie, distingue deux « motivations instinctives » (*sic*) chez l'éthologiste, qui se traduisent par deux formes assez distinctes d'observation des comportements animaux.

La première est celle des « chasseurs » : leur amour des animaux les pousse à rejeter l'idée même d'étudier des animaux en captivité, et ils tentent

---

2. Au sens de l'expression anglaise *life-like*.

## ◇2. DÉMARCHE ET OUTILS

de « *justifier rationnellement ce refus purement affectif avec l'argument selon lequel on ne peut jamais savoir dans quelle mesure le comportement d'un animal est altéré par ses conditions de vie en captivité* » (LORENZ 1984). C'est en partie vrai, aussi l'observation des animaux dans leur milieu naturel est-elle nécessaire pour ne pas souffrir de tels biais. Mais ils omettent de mentionner son principal avantage, à savoir « *qu'elle permet une appréhension directe de l'adaptation écologique de l'espèce en question* ». (*op. cit.*).

L'autre courant (le « type paysan ») consiste à étudier les animaux en captivité, qui prête le flanc à des attaques de la part des naturalistes qui estiment que l'altération du comportement est préjudiciable. Mais l'observateur peut se livrer à de véritables expériences (au même titre que le physicien) et comparer plus aisément différentes espèces. De plus, « *un second avantage de l'observation des animaux en captivité réside paradoxalement dans les troubles comportementaux provoqués par les conditions anormales de vie. [...] C'est ainsi que l'observation des troubles comportementaux résultant de la détention en captivité fournit bien souvent une foule d'informations inattendues sur la nature et la formation de l'enchaînement comportemental en question* » (*op. cit.*).

### Principes méthodologiques

Si l'éthologie n'apparaît en tant que telle qu'avec les travaux de Konrad LORENZ et Nikolaas TINBERGEN, elle s'enracine dans divers courants de la biologie de la fin du XIX<sup>e</sup> siècle et du début du XX<sup>e</sup>.

En particulier, Lloyd MORGAN, psychologue comparatiste, a formulé l'un des principes méthodologiques fondamentaux de l'éthologie, le « canon de MORGAN » :

*« En aucun cas, nous ne pouvons interpréter une action comme la conséquence d'un exercice ou d'une faculté psychique plus haute, si elle peut être interprétée comme l'aboutissement d'une faculté qui est située plus bas dans l'échelle psychologique »* (MORGAN 1894).

Il s'agit en fait d'une application à l'interprétation des comportements animaux du « rasoir d'Ockham », selon lequel il ne faut retenir que les plus simples et le plus petit nombre d'hypothèses expliquant un même phénomène. Mais elle a une importance intrinsèque eu égard aux schémas explicatifs utilisés par l'éthologie.

L'éthologie en effet combine de manière originale explication *fonctionnelle* et explication *causale* (MCFARLAND 1990). Un même comportement peut être abordé selon deux axes complémentaires : l'un oppose causalité efficiente (ou cause proximale, ou encore mécanisme) et causalité finale (ou fonction), l'autre met en regard le comportement tel qu'il est actuellement et son « histoire », la façon dont il s'est constitué. Le tableau 2.1 page 19 donne une synthèse des quatre schémas explicatifs qui résultent de la combinaison de ces deux axes.

TAB. 2.1 – *Les quatre schémas explicatifs qui interviennent en éthologie.*

	<b>cause proximale</b>	<b>cause finale</b>
comportement actuel	causes <i>efficientes</i> du comportement (physiologie, réactions aux stimuli, interactions...): <i>comment</i> se déclenche et se manifeste le comportement	<i>fonction adaptative</i> du comportement: son rôle pour la survie ou le succès reproductif de l'animal
« histoire » du comportement	<i>ontogenèse</i> du comportement: <i>comment</i> il a évolué durant le développement de l'individu, depuis sa naissance, comment il a été acquis et s'est modifié	<i>phylogenèse</i> du comportement dans l'espèce, au cours de l'histoire évolutive: <i>pourquoi</i> le comportement a pris sa forme actuelle (évolution conjointe du mécanisme et de la fonction adaptative)

Ainsi, avec le « canon de MORGAN » d'une part pour régenter le choix des hypothèses concernant les causes efficientes des comportements, et la théorie darwinienne d'autre part, l'éthologie se donne les moyens d'interpréter un comportement d'une manière radicalement dépourvue de téléologie (c'est-à-dire de finalité), mais *qui peut pourtant être exprimée en termes intentionnels*.

Elle rejoint en cela la « posture intentionnelle<sup>3</sup> » défendue par le philosophe américain Daniel DENNETT (DENNETT 1990). Selon lui, il est parfaitement légitime d'interpréter un comportement en termes intentionnels indépendamment de l'existence effective d'« états mentaux » correspondants. L'attribution d'intentionnalité est une heuristique souvent fructueuse pour comprendre un phénomène, tout particulièrement chez les êtres vivants.

3. L'expression « posture intentionnelle » traduit ordinairement l'expression anglaise *intentional stance* utilisée par Daniel DENNETT.

## ◇2. DÉMARCHE ET OUTILS

Le pont entre les causes efficientes et la « raison d'être » d'un comportement s'appuie sur l'hypothèse, constitutive de l'éthologie, selon laquelle les comportements observés ont été *sélectionnés* en vertu de leur valeur adaptative. La théorie darwinienne de l'Évolution suppose un processus « aveugle », radicalement dépourvu de finalité, mais qui, parce qu'il agit en faveur d'une adaptation fonctionnelle des organismes à un environnement, peut *être interprété comme téléologique* (DENNETT 1996).

La question du « pourquoi ? » en éthologie a donc réussi à écarter les réponses finalistes, téléologiques, au profit de réponses non finalistes qui peuvent néanmoins être formulées en termes finalistes. L'éthologie fait appel pour cela à la notion de *téléonomie* proposée par Collin PITTENDRIGH qui permet d'expliquer un comportement *conduisant à un objectif* en termes d'adaptation favorable du point de vue de la sélection naturelle.

Nous allons voir à quel point ce débat sur la finalité et l'intentionnalité trouve des échos en IAD.

### 2.2.2 L'éthologie comme référence épistémologique en IAD

Les méthodes employées en éthologie, que ce soit pour la formulation d'hypothèses sur les causes efficientes, pour l'observation des phénomènes ou pour leur interprétation, ont fortement influencé l'IAD. Nous allons décrire ceux de ces principes que nous avons mis en application dans nos travaux (PICAULT et COLLINOT 1998b).

#### Le « réductionnisme méthodologique »

Il s'agit d'une forme de réductionnisme appliquée aux entités informatiques, et plus spécialement aux architectures et aux modèles d'agents, mises en œuvre dans les systèmes que nous allons construire.

Il s'illustre essentiellement à travers le « **principe de parcimonie** » défini par Alexis DROGOUL par analogie avec le canon de MORGAN :

*« Il n'est pas nécessaire de faire appel à un agent de granularité élevée pour résoudre un problème donné s'il s'avère qu'un agent doté d'une granularité inférieure est capable de lui trouver une solution »* (DROGOUL 1993).

La notion d'« échelle psychologique » invoquée par MORGAN fait place ici au concept de *granularité*, qui repose sur la taille, la complexité, les coûts

associés à telle ou telle architecture d'agent et à tel ou tel modèle d'agent.<sup>4</sup>

L'utilisation concrète du principe de parcimonie requiert une démarche spécifique, l'approche *incrémentale* dans la construction des modèles d'agents, à savoir commencer par tester la pertinence des modèles les plus simples avant d'examiner la nécessité de modèles plus compliqués.

### Parcimonie, construction et explication

Cette construction incrémentale des modèles peut sembler en partie paradoxale dans une démarche de simulation de phénomènes réels, dans la mesure où l'on pourrait chercher au contraire, au moyen d'une architecture cognitive, à représenter des entités considérées comme complexes dans toute leur complexité.

Mais il ne faut pas oublier qu'un modèle est d'abord une représentation abstraite, simplifiée volontairement, du phénomène à expliquer. En utilisant les schémas éthologiques, il tire souvent intérêt de sa simplicité.

Prenons un exemple, emprunté à M. RESNICK (RESNICK 1994).

1. On observe un système dans lequel des agents déplacent des objets.
2. On constate au fil du temps que les objets, initialement tous dispersés, se retrouvent au bout d'un moment *rassemblés en un tas unique*.

On peut donc légitimement donner une interprétation de leur activité collective *en termes intentionnels* : les agents « cherchent à » déposer les objets au même endroit. C'est la *fonction* du système.

C'est alors qu'intervient le schéma des causes efficientes : *comment* les agents parviennent-ils concrètement à ce résultat commun ? Plusieurs mécanismes peuvent être proposés, notamment celui d'une coordination explicite, planifiée, entre les agents. Mais dans une certaine mesure ce type d'hypothèse *n'explique rien*, elle peut être reformulée sous la forme d'une tautologie : « les agents rassemblent les objets au même endroit parce qu'ils cherchent à rassembler les objets au même endroit ».

Or on peut opposer à cette version téléologique une hypothèse reposant sur des comportements plus simples (ceux qui d'ailleurs sont effectivement implantés dans les agents) : chaque agent, *indépendamment des autres et en ne*

---

4. Ce principe peut en fait s'appliquer à deux niveaux : d'une part, essayer de construire les modèles les plus simples possibles, d'autre part, les implémenter au moyen des architectures les plus économiques.

## ◇2. DÉMARCHE ET OUTILS

*transportant qu'un objet à la fois*, prend un objet lorsqu'il en trouve un, puis le dépose auprès d'un autre objet rencontré sur son chemin<sup>5</sup>.

Le fonctionnement *apparemment téléologique* du système émerge donc en fait de comportements *téléonomiques* d'agents en action dans un même environnement – c'est-à-dire des comportements qui ne visent pas explicitement à réaliser la fonction globale observée, mais qui y concourent néanmoins par leur effet local et distribué. Non seulement cette hypothèse permet de comprendre *par quels mécanismes la fonction se constitue*, mais en plus, du point de vue de la *conception de systèmes informatiques*, elle permet de dissocier la *fonction réalisée* au niveau collectif des *moyens mis en œuvre* pour l'atteindre – contrairement à l'hypothèse téléologique.

S'il est possible d'interpréter une même situation en postulant ou non un comportement téléologique de la part des agents, le canon de MORGAN nous incitera à préférer une approche réactive qui ne fait appel, au plus, qu'à la téléonomie.

Resterait, dans un système vivant, à utiliser les schémas historiques : comment ces comportements se sont-ils développés, pour quelles raisons adaptatives ont-ils évolué jusqu'à leur forme actuelle ?

La réponse la plus fréquente dans le cas de systèmes informatiques est que la fonction du système est *imposée par le concepteur*, ce qui fonde un certain type de méthodologie (DROGOUL 2000).

Nous verrons dans la partie II (à partir du chapitre 9) comment ces schémas ontogénétiques et phylogénétiques peuvent être également pris en considération et quels sont leurs apports en termes de modèles, de concepts et de méthodes.

### **Simulation, expérience, interprétation**

Les SMA sont parfois considérés comme des systèmes « non fiables » sous prétexte qu'ils ne produisent pas des résultats de façon déterministe.

Or les schémas d'interprétation et d'explication issus de l'éthologie nous semblent parfaitement appropriés pour compenser l'indéterminisme possible des SMA. En effet, si un SMA « en général » n'est pas forcément prévisible, il n'en reste pas moins que les agents qui le composent agissent en fonction de

---

5. Ainsi, au gré de fluctuations aléatoires, les objets se concentrent d'abord en petits tas, dont les plus petits finissent par disparaître au profit de tas de plus en plus gros. Au bout d'une durée plus ou moins longue, il ne peut rester qu'un seul tas.

leur environnement. Autrement dit, le comportement global du système reflète la complexité de l'environnement. L'éthologie, précisément, fait intervenir dans ses interprétations fonctionnelles la notion d'adaptation au milieu.

Comprendre comment un agent ou un SMA se comporte dans un certain type d'environnement, c'est justement permettre d'explicitier les conditions d'obtention de telle ou telle fonction, ce qui obère largement l'argument du « manque de fiabilité ».

Nous avons évoqué précédemment (§ 2.1.2) la nature assez particulière de la simulation informatique, et plus spécialement de la simulation multi-agent, à savoir que la frontière entre simulation et expérience est assez floue.

Dans ce cadre également, l'éthologie apparaît comme *le* cadre épistémologique adéquat :

- Les méthodes expérimentales de la physique s'attachent à des expériences répétables, parce qu'elles supposent qu'« il existe des ensembles d'objets à certains égards tellement semblables qu'ils peuvent être jugés *interchangeables vis-à-vis d'une expérience donnée* » (SOLER 2000). Ce n'est le cas que rarement dans les SMA (SERVAT 2000, qui précisément s'intéresse à la simulation de phénomènes physiques).
- Les méthodes des sciences humaines conduisent trop facilement à des réductions drastiques de leur objet d'observation (théorie des jeux par exemple), ou à l'inverse à des explications en termes exclusivement intentionnels, téléologiques.
- L'éthologie au contraire permet de faire le lien entre une démarche de conception orientée vers une finalité (les spécifications du concepteur remplaçant la sélection naturelle) et des modèles explicatifs minimaux pour ce qui concerne les causes efficientes des comportements.

### 2.2.3 L'IAD comme « éthologie synthétique »

Ainsi que nous l'avons évoqué à propos de la Vie Artificielle, les SMA sont un support idéal pour construire des « mondes artificiels » plus ou moins réalistes.

Ils peuvent en particulier constituer les bases d'une « éthologie synthétique » (MACLENNAN 1991) en devenant de véritables « laboratoires virtuels » qui permettent, après une phase de validation des modèles, de faire réaliser par les « chercheurs du domaine » eux-mêmes, *in silicio*, des expériences qui non seulement ne sont pas réalisables sur les systèmes réels, mais qui en outre sont

en mesure de leur *suggérer des hypothèses* (DROGOUL 1993, pour MANTA par exemple). La simulation informatique permet alors en quelque sorte de « réaliser » des expériences de pensée (avec toutes les réserves que cela implique en termes de réalisme).

## 2.3 « Cercle vertueux » et conception de SMA

### De la simulation à la conception

L'utilisation de principes inspirés de l'éthologie a été fructueuse en IAD, et a conduit à l'élaboration d'un certain nombre d'approches cherchant à explorer le lien entre la simulation de systèmes naturels et la conception d'applications informatiques au moyen de mécanismes « semblables » (SMA réactifs notamment).

En particulier, Alexis DROGOUL a proposé le paradigme du « cercle vertueux » (DROGOUL 2000), dans lequel les modèles utilisés pour des simulations de phénomènes naturels sont réutilisés, au moins en tant que métaphores, pour la conception de systèmes artificiels dans d'autres domaines (par exemple pour de nouvelles méthodes de résolution de problèmes). Ces applications, à leur tour, permettent d'affiner les techniques de modélisation et de simulation multi-agent, donc fournissent des outils plus adaptés encore aux sciences humaines et naturelles (cf. figure 2.3).

Pour passer de la simulation à d'autres domaines d'application, deux approches opposées sont possibles (DROGOUL 2000) :

- L'approche par *transposition* consiste essentiellement à reproduire une *fonction collective* propre à un système naturel dans un SMA. Par exemple, le tri collectif effectué par une colonie de fourmis peut aisément être réutilisé tel quel pour des groupes de robots (DENEUBOURG et al. 1991).
- L'approche par *analogie*, quant à elle, s'attache à réutiliser la manière dont la fonction se construit à partir des dynamiques d'interactions entre les agents, ou entre les agents et leur environnement, et ce indépendamment de la nature de la fonction elle-même. Il s'agit donc de comprendre des *dynamiques d'auto-organisation* pour les appliquer à d'autres fonctions collectives. L'une des meilleures illustrations de cette approche est l'ÉCO-RÉSOLUTION (DROGOUL et DUBREUIL 1992; DROGOUL et DUBREUIL 1993; DROGOUL et al. 1991).

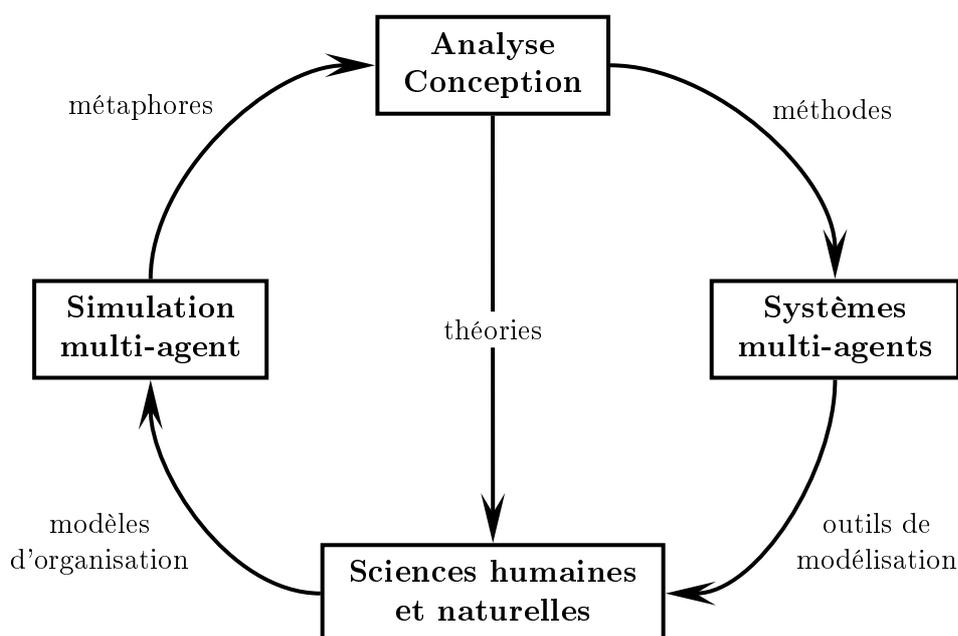


FIG. 2.3 – Le « cercle vertueux » qui lie simulation et conception en IAD. La simulation de systèmes naturels produit des métaphores à l’origine de méthodes de conception de SMA. Ceux-ci peuvent à leur tour être repris pour construire des simulateurs plus adéquats. Les modèles d’agents et les schémas explicatifs produits peuvent par ailleurs avoir un retour sur les domaines d’où sont tirées les modélisations.

### « Cercle vertueux » et socialité

Appliqué à la problématique de cette thèse, le paradigme du « cercle vertueux » va nous pousser, en vue de la conception de SMA capables de s’organiser socialement dans un environnement ouvert, à rechercher des métaphores de phénomènes sociaux dans des sociétés naturelles. Nous verrons notamment au chapitre 3 quels comportements animaux font écho aux problèmes posés par les SMA ouverts. Ces sociétés pourront alors faire l’objet de construction de modèles pour la simulation des phénomènes intéressants (chapitres 4 et 5).

Il nous faudra ensuite examiner les dynamiques sous-jacentes à ces modèles, pour voir dans quelles conditions et sous quelles formes ils peuvent être transposés à des domaines d’applications assez différents (chapitre 6).

Enfin, nous proposerons d’étudier des modèles d’organisation sociale pour un groupe de robots correspondant à la définition des groupes ouverts, puisque immergés dans un environnement « naturel » peuplé d’humains (chapitre 8).

## ◇2. DÉMARCHE ET OUTILS

Nous serons alors amenés à faire un bilan critique des principes et des méthodes que nous avons proposées dans ce chapitre et à proposer des modèles et des approches alternatives (chapitres 9 à 11).

---

Première partie

MODÈLES DE CONSTRUCTION  
DE LA SOCIALITÉ  
EN SIMULATION

nd

## Résumé

Cette première partie sera consacrée à l'étude de la socialité à travers la mise en œuvre des premières étapes du « cercle vertueux » : l'utilisation de la métaphore animale pour construire des modèles capables, en simulation, de produire des phénomènes sociaux transposables aux applications visées ; puis justement la transposition de ces modèles à d'autres applications.

Nous avons trouvé plus particulièrement chez les primates des comportements sociaux correspondant assez bien aux besoins de l'IAD en milieu ouvert. Dans un premier temps, nous présenterons donc les simulations réalisées pour rendre compte de représentations sociales simples existant chez les primates. Nous montrerons comment une hiérarchie peut se reconfigurer au moyen de modèles réactifs qui peuvent donner naissance à des dynamiques complexes.

Ces dynamiques seront dans un deuxième temps explorées dans le cadre d'autres situations, afin de préciser les propriétés de ces modèles et leur transposabilité. C'est ce que nous ferons à travers une application assez éloignée de la primatologie : la construction collective d'un lexique au sein d'un groupe d'agents.

---

# Chapitre troisième.

## Les sociétés de primates

---

« *Si ce sont des animaux, alors que sommes-nous ?* »

FRANS DE WAAL. *La politique du chimpanzé*

*Les sociétés animales, qui vont dans un premier temps servir de métaphores pour l'élaboration de modèles, présentent des formes variées de socialité. Nous allons plus particulièrement nous intéresser aux primates, dont la vie sociale, d'une complexité inégalée, se rapproche assez des objectifs assignés à des agents, logiciels ou robotiques, opérant dans un milieu ouvert et socialisé. Les compétences sociales des primates non humains sont en effet très proches de celles de l'homme et résolvent des problèmes organisationnels qu'on rencontre en IAD (hiérarchies flexibles, reconnaissance de relations entre congénères, alliances, coopération, allocation distribuée de tâches, tromperie, etc.).*

### 3.1 Les diverses formes de la socialité animale

La notion de « société animale » recouvre en pratique des réalités fort différentes d'une espèce à une autre, aussi en donner une définition synthétique n'est pas chose aisée. Notons d'abord qu'il ne suffit pas de vivre en groupe pour vivre en société. Le groupe n'est généralement considéré comme social qu'au vu d'un certain nombre de caractéristiques quant aux liens qui unissent ses membres. Ainsi, E. O. WILSON, fondateur de la sociobiologie, définit une dizaine de critères pour estimer le degré de socialité d'un groupe (WILSON 1975), parmi lesquels sa taille, sa distribution démographique, les flux d'informations, etc.

### ◇3. LES SOCIÉTÉS DE PRIMATES

#### Les types de groupes

Le premier stade social des groupes animaux est *l'agrégation*, qui rassemble des individus en un même lieu de façon plus ou moins durable, sans organisation particulière (c'est le cas d'animaux réunis autour d'un point d'eau pour se désaltérer).

On peut ensuite catégoriser les groupes organisés selon deux critères, l'ouverture du groupe d'une part (c'est-à-dire la capacité à distinguer entre membres du groupe et intrus) et la reconnaissance particulière des individus qui le composent d'autre part (cf. tableau 3.1).

TAB. 3.1 – *Quelques exemples de groupes animaux*

	Groupe ouvert	Groupe fermé
Groupe anonyme	essaims d'insectes, bancs de poissons, nuées d'oiseaux	insectes sociaux (fourmis, termites...)
Groupe non anonyme (individualisé)	colonie d'oiseaux (nidification)	rongeurs, canidés, primates

**N.B.** Cette distinction biologique entre groupes ouverts et groupes fermés n'a rien de commun avec la distinction informatique entre systèmes ouverts et systèmes fermés. Un groupe animal peut difficilement être « fermé » au sens informatique, sa fermeture consistera à reconnaître l'appartenance à un groupe qui s'est construit au cours du temps. Un système informatique « ouvert » pourra justement posséder cette propriété de « fermeture » au sens biologique, s'il lui est nécessaire de détecter d'éventuels intrus par rapport à un groupe préalablement constitué.

#### Les groupes ouverts

Les groupes anonymes ouverts, tels que les bancs de poissons ou les nuées d'oiseaux, manifestent une organisation essentiellement spatiale, dont des algorithmes très simples en Vie Artificielle ont su rendre compte avec un grand degré de vraisemblance (REYNOLDS 1987). De même, dans un groupe ouvert non anonyme, la reconnaissance de certains individus (typiquement, les voisins) sert avant tout à gérer l'espace, notamment à se localiser (dans le cas de la nidification des oiseaux, il s'agira de rejoindre son nid par exemple).

### 3.1. LES DIVERSES FORMES DE LA SOCIALITÉ ANIMALE

#### Les groupes fermés

Les groupes fermés sont d'un intérêt plus grand pour l'IAD, dans la mesure où ils soulèvent implicitement le problème de la constitution d'une identité de groupe (autrement dit de la réification dans le système d'une entité collective).

Les membres des sociétés d'insectes ne sont généralement pas individualisés. La colonie peut parfois être constituée de castes, c'est-à-dire de sous-groupes spécialisés fonctionnellement, ce qui s'accompagne d'un comportement propre (polyéthisme) et parfois d'une physiologie spécifique (polymorphisme). Un animal sait alors reconnaître l'appartenance d'un congénère à telle ou telle caste. Autrement dit, la société organise des *fonctions* assurées collectivement par des sous-groupes de la colonie.

On distingue tout particulièrement les espèces dites *eusociales* qui sont définies par la conjonction des trois critères suivants (DESOR 1999) :

- coopération entre les membres pour le soin aux jeunes ;
- division du travail entre individus stériles et fertiles ;
- coopération entre au moins deux générations.

L'écrasante majorité des espèces eusociales sont des insectes (toutes les fourmis, tous les termites, les abeilles et certaines guêpes). Toutefois on en a découvert récemment parmi les vertébrés et même les mammifères (les rats-taupes, *Heterocephalus glaber*), qui présentent exactement les trois prérequis ci-dessus (JARVIS 1981).

Les insectes eusociaux, qui représentent une forme d'« optimalité » dans l'efficacité qu'apporte l'organisation sociale au regard de certaines fonctionnalités collectives, ont fait l'objet au cours de la dernière décennie de nombreuses études en informatique, que ce soit dans un but de simulation (DROGOUL et FERBER 1992; THERAULAZ 1991)<sup>1</sup>, ou pour servir de métaphore dans l'élaboration de nouvelles méthodes de résolution de problèmes, comme l'éco-résolution (DROGOUL et DUBREUIL 1993) ou la notion d'« intelligence en essaim » (BONABEAU et al. 1999a), des algorithmes de routage de réseaux (SCHOONDERWOERD et al. 1997), ou encore pour défendre en IAD certaines

---

1. Signalons toutefois que certains des mécanismes connus pour leur utilisation massive par des espèces eusociales, comme la *stigmergie* ou communication par le biais de l'environnement (GRASSÉ 1959), sont également utilisés comme base de la coopération chez des espèces qui ne sont pas eusociales, par exemple par certaines araignées (BOURJOT et al. 1999).

### ◇3. LES SOCIÉTÉS DE PRIMATES

approches méthodologiques dérivées de l'éthologie (DROGOUL et COLLINOT 1997).

Par ailleurs les problématiques liées à la fermeture du groupe commencent à être modélisées dans une perspective multi-agent ; c'est le cas par exemple de la formation de l'*odeur coloniale* des fourmis (travaux de Paolo URBANO, LIP6, en collaboration avec l'équipe de Dominique FRESNEAU au Laboratoire d'Éthologie de Paris XIII).

Restent donc les groupes fermés individualisés, ceux dont les membres à la fois *se reconnaissent mutuellement* et se reconnaissent *comme membres du même groupe*. Le parallèle avec les besoins de groupes de robots devant opérer dans un environnement humain, que nous avons évoqués au chapitre 1, est assez immédiat.

C'est donc au sein de telles sociétés animales qu'il conviendra de chercher des modèles d'organisation pouvant donner lieu à une transposition en IAD.

La plupart des vertébrés ayant développé des organisations sociales complexes sont des mammifères, tels que les rongeurs (les rats notamment), les canidés (par exemple les loups), et les primates. C'est précisément sur ces derniers que nous allons nous focaliser, pour les raisons suivantes :

- En premier lieu, c'est l'ordre auquel appartient l'espèce humaine ; à ce titre, la socialité des primates ne peut manquer d'alimenter des théories et des débats sur les origines de l'Homme et de ses compétences spécifiques (ou supposées telles), comme le langage, la fabrication d'outils, la culture... Mais ces querelles s'enracinent de fait dans une grande proximité entre les comportements des primates non humains et ceux de leurs cousins *Homo sapiens*, tout particulièrement dans le domaine social. Simuler des comportements sociaux des primates pour ensuite les implanter sur des robots serait alors une piste pour donner à ceux-ci une « vraisemblance » aux yeux de leurs hôtes humains.
- En second lieu, le comportement social des primates présente des particularités qui le différencient de celui des autres mammifères, et nous intéresse tout spécialement dans notre recherche de modèles biologiques. Nous allons voir en effet que ces compétences font écho à de nombreuses problématiques de l'IAD, dont certaines pourront être mises à profit dans le projet de robotique sociale déjà évoqué.

### 3.2 Les compétences sociales des primates

#### 3.2.1 Phylogénie des primates

Les primates, selon les classifications habituelles, sont divisés en :

- prosimiens (par exemple lémuriens, loris, aye-aye) ;
- simiens (les « singes », en anglais *monkeys* ), ceux-ci eux-mêmes composés de :
  - platyrrhiniens (atèles, capucins, tamarins par exemple) ;
  - catarrhiniens (colobes, vervets, macaques, babouins) ;
  - hominiens.

Les hominiens forment un sous-ordre restreint comportant, outre l'Homme moderne (*Homo sapiens*) et ses ancêtres hominidés, les singes anthropoïdes, à savoir les gibbons et les « grands singes » (en anglais *apes*) : orang-outan, gorille, chimpanzés.

L'ancêtre de l'orang-outan actuel (*Pongo*), le *Sivapithecus*, est apparu il y a environ 16 millions d'années, tandis que les gorilles ne se sont différenciés qu'il y a 7 à 9 millions d'années. Les ancêtres des hominidés se sont séparés des chimpanzés il y a environ 6,5 millions d'années ; enfin, plus récemment (2,5 millions d'années), les chimpanzés se sont scindés en deux espèces : les chimpanzés communs (*Pan troglodytes*) et les chimpanzés pygmées ou bonobos (*Pan paniscus*).

Du fait des liens phylogénétiques, on observe une proximité génétique plus forte entre chimpanzés et humains qu'entre chimpanzés et orang-outans. L'Homme partage plus de 99% de ses gènes avec les chimpanzés, et une grande partie des différences n'a pas de fonction clairement établie.<sup>2</sup>

#### 3.2.2 Spécificités des comportements sociaux des primates

##### La dominance dans le monde animal

La *dominance* est « une caractéristique de l'organisation sociale dans laquelle certains individus acquièrent un statut élevé (généralement à la suite

---

2. La majorité des variations génétiques affectent des parties non codantes de l'ADN (KING et WILSON 1975). L'explication des spécificités humaines serait plutôt à rechercher dans un réarrangement chromosomique qui a modifié la position, donc vraisemblablement l'ordre d'expression, de certains gènes (KUPIEC et SONIGO 2000).

### ◇3. LES SOCIÉTÉS DE PRIMATES

*d'une agression) tandis que d'autres conservent un statut inférieur* » (MCFARLAND 1990) ; elle a été mise en évidence pour la première fois dans des études sur les poules domestiques (SCHJELDERUP-EBBE 1922; SCHJELDERUP-EBBE 1935).

Ce statut supérieur permet en général d'obtenir un accès privilégié aux ressources (nourriture, territoires, femelles) mais impose souvent une vigilance constante pour protéger le groupe des prédateurs et y maintenir sa position contre des rivaux. Inversement, les animaux dominés reconnaissent la structure sociale et ne cherchent donc que rarement des confrontations violentes avec les dominants : l'agression est donc maîtrisée, régulée (HUTZLER 1995).

La relation de dominance forme une *hiérarchie* souvent *linéaire* (i.e. la relation de dominance est transitive : si A domine B et B domine C, alors A domine également C). L'intérêt majeur de la hiérarchie chez les animaux sociaux est de diminuer considérablement les confrontations violentes, en permettant une forte *ritualisation* des rencontres : la plupart du temps, celles-ci ne débouchent pas sur un véritable combat mais sur un « combat simulé » ou un échange de signaux (cris, mimiques faciales...) qui produit un renforcement des statuts hiérarchiques déjà en place (MCFARLAND 1990).

Chez la plupart des espèces sociales, les individus dominants se différencient des dominés par leur comportement ou leur apparence (taille du corps notamment). La relation de dominance s'appuie essentiellement sur la force physique des adversaires et le résultat des interactions antérieures.

#### **La dominance chez les primates**

Toutefois, les primates sont les rares animaux parmi lesquels la relation de dominance met en jeu des facteurs autres que la seule puissance physique. Le rang d'un individu dans la hiérarchie peut dépendre de celui de ses parents : un enfant de dominant aura plus de chances de devenir dominant. Ainsi, les femelles chimpanzés de rang élevé consacrent plus de temps à leurs enfants mâles (qui resteront au sein du groupe) qu'aux femelles (qui devront quitter le groupe).

Par ailleurs, les relations d'*affiliation* entre individus jouent un rôle crucial. Les singes passent une grande partie de leur temps à se toiletter mutuellement (cf. figure 3.1), ce qui ne constitue pas tant une mesure prophylactique qu'un comportement ritualisé destiné à renforcer les liens sociaux (i.e. l'*affiliation* entre individus), apaiser des tensions, etc. Selon R. DUNBAR, ce

### 3.2. LES COMPÉTENCES SOCIALES DES PRIMATES

comportement jouerait le même rôle que le « bavardage » ou le « commérage » (*gossiping*) chez l'homme (DUNBAR 1996). L'utilisation du langage à la place du toilettage pour cette activité présente en outre l'avantage de pouvoir s'adresser à plusieurs partenaires simultanément.

Les relations d'affiliation entre individus mènent à la constitution de coalitions lors des combats : lorsque deux individus se battent, un troisième peut entrer en jeu pour défendre l'un des deux protagonistes, et ainsi de suite. L'ensemble des coalitions effectives et des liens de parenté ou d'affiliation constituent des réseaux d'alliance qui contribuent fortement au rang des individus.

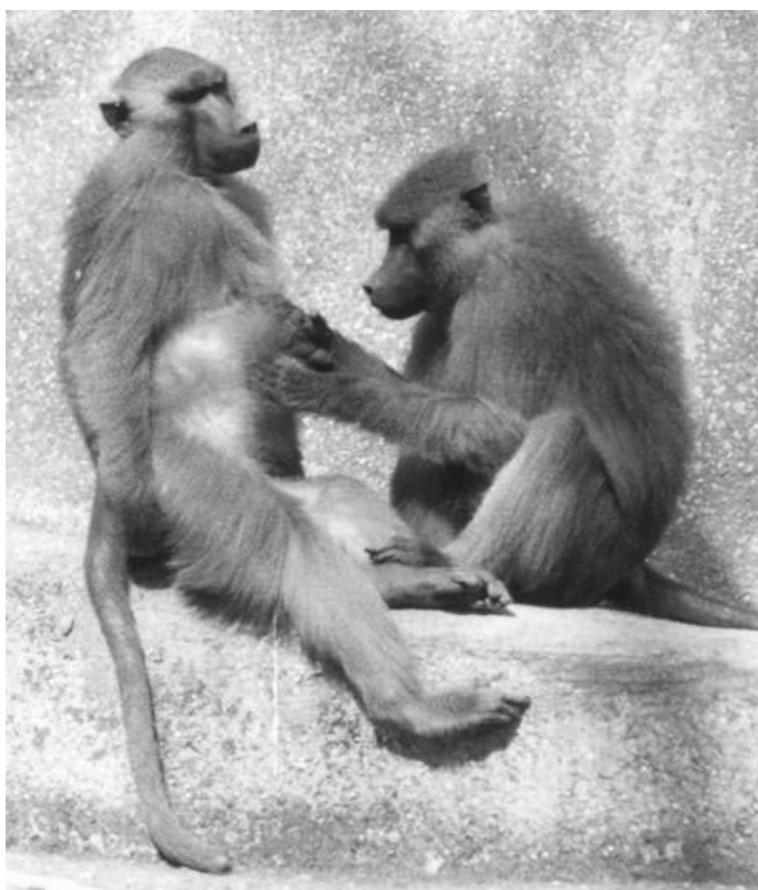


FIG. 3.1 – *Comportement de toilettage (babouins). Ces comportements ont une fonction moins hygiénique que sociale (renforcer les liens entre individus).*

L'organisation sociale des primates, et plus spécialement des simiens, s'appuie également sur une gamme étendue de relations et de comportements, que nous allons approfondir.

### 3.2.3 La reconnaissance des relations sociales

L'une des aptitudes les plus répandues au sein des primates est la capacité à reconnaître, outre les individus particuliers du groupe, les relations sociales qui existent entre eux. Ce point a été mis en évidence par des expériences de laboratoire comme de terrain. Ainsi, en utilisant une méthode d'appariement de diapositives, les études menées par V. DASSER (DASSER 1987; DASSER 1988) ont montré que les macaques disposaient d'un « proto-concept » de dominance et de filiation ; ils sont de plus capables d'inférences simples sur ces concepts, comme la transitivité. Autrement dit, les individus ne répondent pas seulement à un comportement social mais aussi à une *relation sociale*, ce qui requiert un degré d'abstraction supérieur.

De même, D. CHENEY et R. SEYFARTH ont réalisé des expériences en milieu naturel sur les capacités de reconnaissance vocale des vervets ou singes verts (*Cercopithecus aethiops*) (CHENEY et SEYFARTH 1980). Lorsqu'on diffuse à un groupe de mères le cri enregistré d'un enfant, 1° toutes réagissent à la vocalisation en se tournant vers le haut-parleur ; 2° la mère du jeune dont on diffuse le cri a un temps de latence plus court ; 3° les autres mères présentes observent de façon soutenue la mère concernée après diffusion de l'enregistrement. L'analyse de cas où la mère du jeune n'a pas réagi montre la même réaction de la part des autres femelles, ce qui indique que leur comportement est dicté non seulement par la reconnaissance du jeune, mais aussi par sa relation de filiation avec une femelle du groupe.

### 3.2.4 Les réseaux d'alliance

Le degré d'affiliation entre les individus, qui dépend pour une part de leurs liens de parenté (part plus ou moins forte selon les espèces) et pour une autre du temps passé en toilettage mutuel, sert à cimenter les liens entre individus. Ce comportement fait l'objet d'une rétroaction positive, dans la mesure où un individu choisira préférentiellement comme partenaire un animal avec lequel il est déjà fortement affilié. Comme l'agressivité est tempérée par l'affiliation, les dominants font l'objet de toilettages nombreux.

Les individus affiliés s'entraident lors des conflits, ce qui permet de défaire un adversaire trop fort pour chacun des participants. Ces coalitions occasionnelles peuvent se stabiliser en réseaux d'alliances qui se maintiennent durant d'assez longues périodes (des semaines, des mois, voire des années).

Selon les espèces, ces réseaux peuvent être seulement « défensifs » (deux

## 3.2. LES COMPÉTENCES SOCIALES DES PRIMATES

dominants s'entendent pour écarter leurs rivaux), comme chez les macaques et les babouins, ou parfois également « offensifs » (deux individus assez forts, mais pas assez cependant pour atteindre à eux seuls le statut de dominant, s'allient pour renverser le dominant actuel), chez les chimpanzés par exemple.

Ce dernier cas a été étudié très en détail par F. DE WAAL au zoo d'Arnhem (Pays-Bas) (DE WAAL 1982). Les alliances supposent une réciprocité sans laquelle elles ne subsistent pas, l'individu exploité pouvant se retourner contre son ancien allié en pactisant avec l'ennemi d'hier. Ces stratégies de type « donnant-donnant » rappellent les expériences menées par AXELROD et HAMILTON en théorie des jeux sur le dilemme itéré du prisonnier (AXELROD et HAMILTON 1981; AXELROD 1996).

### 3.2.5 Les leurre de puissance

Un certain nombre d'observations montrent que les primates peuvent être induits en erreur par des apparences physiques de la puissance. Pour appréhender le monde social, ils se fient à des éléments « subjectifs<sup>3</sup> ». Ils peuvent être déroutés par l'introduction d'éléments nouveaux dans leur environnement, mais font preuve d'une grande capacité d'adaptation en attribuant un rôle social aux nouveautés. En voici trois exemples, cités par B. CYRULNIK (CYRULNIK 1983).

1° Les chimpanzés acquièrent en vieillissant des poils blancs sur les épaules. Dans la nature, les jeunes chimpanzés attribuent un statut élevé à ces « augustes vieillards », car ils apprennent très vite que ceux-ci ont accumulé une expérience, notamment sociale, qui peut être mise à profit.

Cela est vrai à tel point que si on badigeonne les épaules d'un jeune chimpanzé inexpérimenté de peinture blanche, on change en même temps son statut social : les autres individus vont se comporter vis-à-vis de lui comme ils le feraient pour un vieux mâle dominant.

Cela montre en premier lieu que les primates sont sensibles à des leurre relativement simple même dans leurs comportements sociaux.

2° Un autre chimpanzé, de rang moyen, avait réussi à voler un bidon

---

3. C'est-à-dire des caractéristiques observables qui sont ordinairement corrélées à un statut social élevé, mais qui n'en sont qu'un indice indirect. On peut parler de « subjectivité » par opposition à l'utilisation de perceptions reflétant « objectivement » un statut social (comme c'est le cas par exemple chez les guêpes qui diffusent des molécules en proportion de leur rang).

### ◇3. LES SOCIÉTÉS DE PRIMATES

vide aux primatologues qui étudiaient son groupe. Il ne s'en séparait plus et s'en servait pour intimider ses congénères, très impressionnés. Il acquit ainsi rapidement un statut de dominant.

Une fois le bidon récupéré par ses propriétaires légitimes, le chimpanzé est resté dominant : les habitudes sociales du groupe avaient été durablement modifiées par un simple leurre, par un objet au possesseur duquel les autres chimpanzés ont accordé une forme de « prestige » sans commune mesure avec l'usage qu'on pouvait effectivement faire du bidon.

3° J. B. WOLFE et J. T. COWLES ont étudié la notion de « valeur » en introduisant un système de récompenses dans lequel des jetons colorés permettent d'obtenir 0, 1 ou 2 grains de raisins selon la couleur qu'ils portent (WOLFE 1936). Les singes sont très vite capables de comprendre la règle sous-jacente et tentent d'obtenir les meilleurs jetons.

Ce système monétaire est immédiatement investi par la socialité dans le groupe : les individus fortement dominés passent leur temps à adopter des postures de soumission vis-à-vis des dominants, et ceux-ci en profitent pour leur voler les quelques jetons de valeur qu'ils peuvent avoir, en leur laissant ceux dépourvus d'intérêt.

Or, si l'on dote l'un de ces dominés d'un gros sac de jetons les plus convoités, il subit une ascension sociale assez rapide, les autres membres du groupe lui adressant des signaux de soumission en mendiant quelques piécettes...

On voit donc à quel point la socialité des primates est éloignée de celle de la plupart des autres animaux, pour lesquels elle ne repose que sur la force des individus (y compris lorsque cette force physique dépend elle-même du rang). Chez les primates, les rapports sociaux s'enracinent dans une part de « subjectivité », dans une représentation approximative de ce que font et ce que sont les autres.

#### **3.2.6 Tâches collectives et coopération**

Les singes sont peu enclins à la coopération. La plupart du temps, chacun s'alimente sans se soucier du voisin, voire égoïstement lorsque les ressources sont limitées. À l'exception du soin aux jeunes ou de la défense contre un prédateur, il existe peu de tâches requérant la participation de plusieurs membres du groupe.

On en trouve toutefois quelques exemples chez les chimpanzés.

## 3.2. LES COMPÉTENCES SOCIALES DES PRIMATES

### La chasse

Les chimpanzés mangent principalement des fruits, accompagnés de noix, de fourmis et autres insectes. Mais ils ne dédaignent pas la viande, qu'ils trouvent par exemple sous la forme d'autres singes, les colobes. Les chimpanzés vivant en savane peuvent chasser de façon opportuniste, en tuant de très jeunes colobes. Chacun agit alors pour soi, dans la mesure où la proie peut être confisquée par le dominant (NISHIDA et al. 1979).

C. BOESCH rapporte qu'en revanche, dans la forêt équatoriale, les chimpanzés chassent *collectivement* (BOESCH et BOESCH 1989). Il faut, au sol, encercler les groupes de colobes (qui vivent à la cime des arbres), puis grimper silencieusement pour les surprendre, anticiper leurs mouvements, les rabattre, et en attraper. Il semblerait que chacun ait un rôle fonctionnel précis durant la chasse, rôle qui dépendrait de son « savoir-faire ». De plus, la coordination entre les chasseurs doit s'effectuer de façon complètement distribuée. Enfin, la répartition de la nourriture se fait non en fonction du rang hiérarchique, mais du rôle joué par chaque chasseur. Si certains primatologues se montrent sceptiques quant à la réalité de ces rôles, au demeurant difficiles à attribuer clairement en raison de l'extrême confusion qui règne durant la chasse, il n'en reste pas moins que cette activité est l'une des rares situations de coopération et de coordination entre singes en milieu naturel, qui implique de plus une allocation distribuée des tâches nécessaires pour mener à bien la capture des proies.

### Contrôle du territoire

On a pu également observer des « rondes d'inspection » effectuées par les mâles aux frontières du territoire occupé par leur groupe. Ils se déplacent silencieusement et n'hésitent pas à attaquer très violemment et même mortellement un individu d'un autre groupe qui se trouverait sur leur chemin (GOODALL 1979). Il ne s'agit plus dans ce cas de combats ritualisés mais de véritables agressions.

### 3.2.7 La tromperie

Depuis quelques années, une nouvelle gamme de comportements sociaux a été mise en évidence : la *tromperie sociale* (BYRNE et WHITEN 1988b). Celle-ci a d'abord été repérée dans des anecdotes éparses rapportées par les prima-

### ◇3. LES SOCIÉTÉS DE PRIMATES

tologues, puis ces observations ont fait l'objet de compilations méthodiques et détaillées pour finalement donner naissance à une nouvelle théorie sur les origines de l'intelligence chez les primates : l'hypothèse de *l'intelligence machiavélique*<sup>4</sup> (BYRNE et WHITEN 1988a).

#### La tromperie sociale

BYRNE et WHITEN donnent une définition fonctionnelle de la « tromperie sociale » : il s'agit de l'utilisation d'un *comportement usuel* par un individu, dans un contexte tel qu'il provoque une erreur d'interprétation de la part de ses congénères, pour le plus grand profit de cet individu (BYRNE 1995).

La tromperie implique donc :

- un agent : le manipulateur ;
- une cible : la (ou les) victime(s) de la tromperie ;
- un éventuel outil social (c'est-à-dire un objet ou un autre individu instrumentalisé) que l'agent manipule contre la cible ;
- un objectif qui n'est pas forcément intentionnel mais qui peut être décrit comme tel par un observateur, et qui correspond au bénéfice obtenu par l'agent.

En voici trois exemples, rapportés par (BYRNE 1995).

1° Paul, un jeune babouin, s'approche de Mel, une femelle, occupée à déterrer un tubercule. C'est une opération pénible mais fructueuse. Lorsqu'elle a terminé, Paul regarde autour de lui, ne voit aucun autre babouin, et crie. Sa mère, de rang plus élevé que Mel, arrive en courant et en grognant, et s'en prend aussitôt à Mel. Celle-ci s'enfuit sans demander son reste. Dès que Mel et la mère de Paul sont hors de vue, Paul mange le tubercule. Ici, Paul est l'agent (A), Mel la cible (T), la mère de Paul l'outil (*tool*) et le tubercule l'objectif (cf. figure 3.2).

2° Un mâle, Melton, a énervé des adultes qui le poursuivent. Mais au lieu de continuer à courir, celui-ci se dresse soudain sur ses pattes postérieures et fixe l'horizon, comme le fait tout babouin qui a perçu un danger. Ses poursuivants s'arrêtent et regardent dans la même direction. Il n'y a rien de suspect<sup>5</sup>, mais la poursuite interrompue n'est pas reprise (cf. figure 3.3).

---

4. Pour traduire l'anglais *machiavellian intelligence*, je préfère utiliser le terme d'intelligence *machiavélique* plutôt que *machiavélique*, dont la connotation morale négative pourrait nuire à l'idée sous-jacente.

5. BYRNE a pris soin de vérifier très attentivement.

### 3.2. LES COMPÉTENCES SOCIALES DES PRIMATES

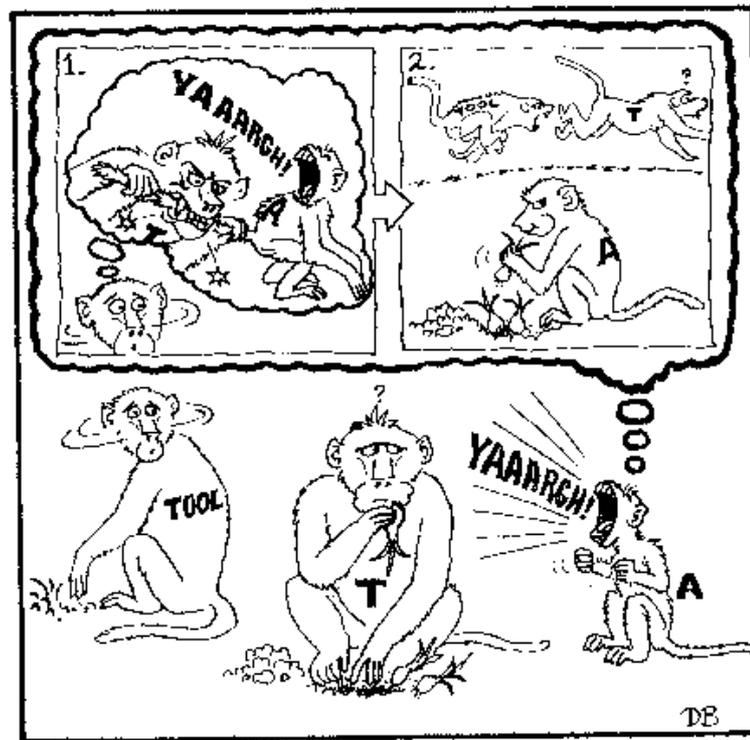


FIG. 3.2 – *Un exemple de tromperie sociale (voir texte). Les intentions attribuables aux agents font intervenir une représentation de ce que « pensent » les congénères et de la façon dont ils vont interpréter la situation.* — d'après (BYRNE 1995)

3° La tromperie peut donner lieu à des comportements de contre-tromperie. Ainsi, les chimpanzés qui savent où est cachée de la nourriture sont capables, en présence d'un dominant, d'inhiber leur appétit pour pouvoir manger tranquillement après son départ, sans quoi celui-ci pourrait s'approprier le repas de son subordonné (tromperie). Mais le psychologue F. PLOOIJ a noté que le dominant fait parfois mine de s'en aller, pour en fait se dissimuler derrière un arbre. Il revient rapidement sur les lieux au moment où le trompeur récupère la nourriture, et s'en empare (contre-tromperie, cf. figure 3.4).

Ces capacités de tromperie et de contre-tromperie ont un intérêt pour l'IAD, par exemple pour des applications d'agents distribués sur un réseau, pour lesquelles il peut être nécessaire de savoir détecter des comportements indésirables voire malveillants (MARSH 1992).

### ◇3. LES SOCIÉTÉS DE PRIMATES

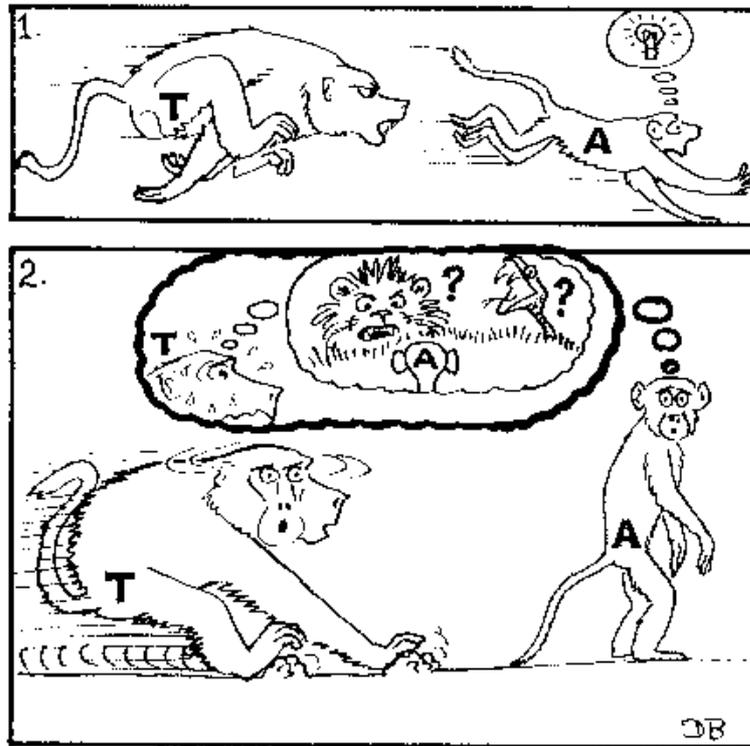


FIG. 3.3 – *Un autre exemple de tromperie sociale (voir texte).*  
— d'après (BYRNE 1995)

#### L'hypothèse de l'intelligence machiavélique

L'hypothèse de l'intelligence machiavélique pose que les aptitudes cognitives sophistiquées des primates (dont les limites sont présentées § 3.3) sont essentiellement *d'origine sociale*.

L'origine sociale de l'intelligence serait née d'un difficile compromis entre, d'une part, une hostilité individualiste envers la promiscuité induite par la vie en groupe et, d'autre part, la nécessité de se rassembler pour lutter contre des prédateurs nombreux. Les primates sont en effet fondamentalement individualistes dans leurs comportements (notamment vis-à-vis de la nourriture qu'ils ne partagent que très rarement).

Dans un groupe composé d'agents guidés par leur intérêt économique propre, la théorie des jeux montre qu'une stratégie équilibrée consiste à pratiquer la réciprocité (AXELROD et HAMILTON 1981; DAWKINS 1980), ce qui suppose une mémoire des interactions avec chacun des congénères. Mais surtout, il reste la possibilité de fausser le jeu à son avantage si l'on parvient à développer des comportements de tricherie *qui sont rarement décelés*.

### 3.2. LES COMPÉTENCES SOCIALES DES PRIMATES

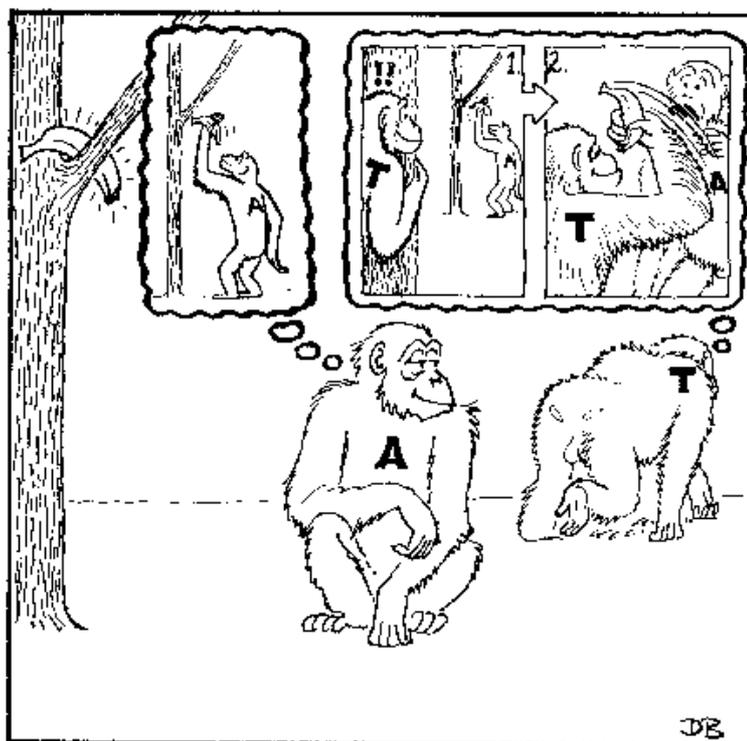


FIG. 3.4 – *Un exemple de contre-tromperie (voir texte) : A pense pouvoir s'emparer de la banane après le départ de T, mais ce dernier subodore la tromperie et s'apprête à revenir.* — d'après (BYRNE 1995)

BYRNE discute assez longuement les conditions nécessaires pour qu'un comportement de tromperie sociale puisse conférer un avantage évolutif et être sélectionné. Il faut en particulier que le bénéfice obtenu par le trompeur soit très grand par rapport aux dommages subis par les victimes, ou encore mieux, qu'elles ne se rendent compte de rien, ce qui évite des représailles. Dans ce cas, les actes de tromperie ne doivent pas être trop fréquents, car alors un apprentissage simple permettrait de les déjouer aisément. Or ces caractéristiques sont celles des phénomènes effectivement observés.

À l'appui de cette théorie, DUNBAR a étudié les corrélations entre taille du cerveau, taille des groupes et le taux de prédation. Il a remarqué que le quotient néocortical (rapport entre la masse du néocortex et celle du reste du cerveau) augmente très nettement, chez toutes les espèces de primates, avec la taille des groupes (DUNBAR 1992).

De plus, observant que, chez l'homme, les conversations quotidiennes ont un contenu essentiellement social (plus précisément, elles contiennent une forte

### ◇3. LES SOCIÉTÉS DE PRIMATES

proportion de « bavardage » ou de « commérage »), il soutient que le langage joue un rôle équivalent au toilettage chez les primates non humains, c'est-à-dire qu'il sert en grande partie à consolider les liens sociaux (DUNBAR 1996). Cette fonction affiliative du langage aurait pu faciliter l'accroissement de la taille des groupes humains en permettant d'avoir plusieurs partenaires simultanément.

Autrement dit, pour maintenir la cohésion d'un groupe, il serait crucial de disposer d'une vaste capacité à traiter les informations (reconnaître les relations sociales, les exploiter à son profit, voire les manipuler).

## 3.3 Débats autour de l'intelligence des primates

### 3.3.1 Du primate à l'homme

La proximité phylogénétique des primates et de l'homme pose le problème du « propre de l'homme ». Les outils, le langage, l'intentionnalité, la culture, l'intelligence sociale ont régulièrement été mentionnés comme candidats (PICQ et COPPENS 2001). Nous avons vu que la complexité sociale des primates est assez proche de celle qu'on peut rencontrer dans les groupes humains, langage et organisation normative mis à part.

De plus, l'hypothèse de l'« intelligence machiavélienne » laisse à penser que les compétences cognitives des primates sont en grande partie issues des interactions sociales et des capacités nécessaires pour s'en sortir à bon compte. On trouve effectivement des traces de socialité dans la plupart des activités complexes des primates<sup>6</sup>.

#### Les outils

Depuis quelques dizaines d'années déjà, le rôle prépondérant de l'outil dans le processus d'homínisation a été remis en question. De nombreuses espèces animales utilisent des objets pour une fonction précise, mais on peut montrer qu'il s'agit le plus souvent de comportements innés.

En revanche, les grands singes sont capables d'évaluer très finement l'adéquation entre un « outil », c'est-à-dire un objet d'une certaine forme, et une

---

6. Dans un exposé récent à l'EHESS, toutefois, BYRNE est en partie revenu sur l'hypothèse de l'intelligence machiavélienne en précisant que chez les grands singes, les techniques très élaborées liées à l'obtention de la nourriture pouvaient expliquer leurs capacités, qui sont nettement supérieures à celles des autres primates.

### 3.3. DÉBATS AUTOUR DE L'INTELLIGENCE DES PRIMATES

tâche. C'est le cas par exemple des brindilles utilisées pour la « pêche aux termites », mais surtout des lourds « marteaux » de bois dont ils se servent pour casser les noix. Les noix en question sont très dures, difficiles à casser et cet art requiert une dizaine d'années d'entraînement. Le morceau de bois doit être choisi soigneusement, pris en main de façon correcte, etc. On observerait même une « industrie » du cassage de noix, avec des sites spécialisés (BOESCH 1983; BOESCH 1990).

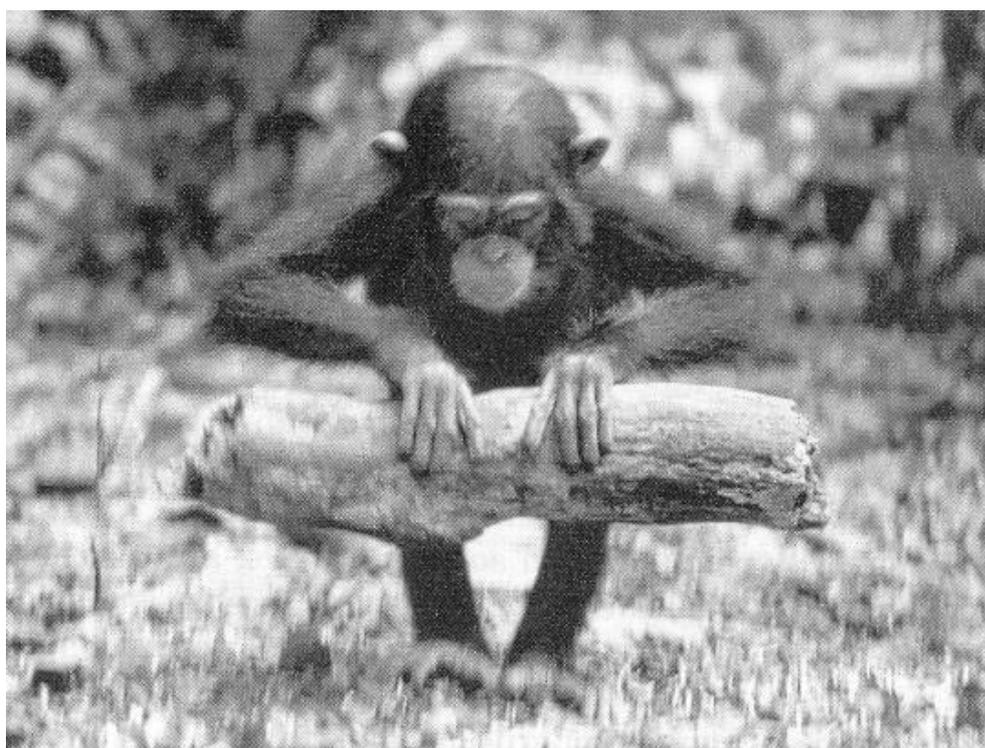


FIG. 3.5 – *Un jeune chimpanzé tente d'utiliser un marteau de bois pour casser des noix. L'acquisition de la technique correcte demande de nombreuses années d'apprentissage.* — d'après (BYRNE 1995)

Il semble que chez les chimpanzés, l'usage d'outils soit en partie social, notamment de mère à enfant. C. BOESCH rapporte en effet des comportements de « facilitation » dans lesquels la mère laisse un marteau approprié à disposition de son enfant, voire (dans un seul cas observé) repositionne l'outil dans la main de l'enfant (BOESCH 1991).

### ◇3. LES SOCIÉTÉS DE PRIMATES

#### **Le langage et la culture**

Le langage a longtemps été considéré comme l'apanage exclusif de l'homme, de même que la culture. Les tentatives réalisées pour tenter d'enseigner le langage humain aux primates ont souvent été vaines, ou entachées de manque de rigueur dans le protocole, jusqu'aux travaux de Susan SAVAGE-RUMBAUGH (SAVAGE-RUMBAUGH et al. 1993).

Pour ses expériences, elle utilise un lexigramme : une console permettant, en appuyant sur des touches portant des symboles non figuratifs, de prononcer un mot préenregistré. Lors des essais menés sur une femelle bonobo, l'enfant de celle-ci, nommé Kanzi, s'est mis spontanément à utiliser le lexigramme pour communiquer (SAVAGE-RUMBAUGH et al. 1986). Depuis, il a appris à maîtriser un millier de mots environ et vit au contact permanent des expérimentateurs et de ses semblables.

Il manipule un « langage » sans vraie syntaxe, associant quelques mots les uns aux autres. C'est cependant assez pour établir une communication efficace et étudier la manière dont il se développe dans un environnement humain. De plus, les études comparatives menées entre Kanzi et une petite fille de deux ans, Alia, ont montré un léger avantage en faveur du chimpanzé<sup>7</sup> (SAVAGE-RUMBAUGH et al. 1993).

Enfin, on parle depuis quelque temps de phénomènes « culturels » chez les primates, plus spécialement chez les grands singes (BOESCH 1994; FOUTS et al. 1982, entre autres). Cette culture désigne la plupart du temps une variabilité dans les comportements (notamment dans ceux qui jouent un rôle social) qui ne trouve d'explication ni dans l'hérédité, ni dans une adaptation à un environnement particulier. Ainsi, dans deux groupes apparentés et habitant des milieux similaires, une même fonction (attirer l'attention, parader) peut être assurée par des gestuelles radicalement différentes ; inversement, une même gestuelle peut dans chacun des groupes correspondre à des fonctions séparées.

Les individus qui intègrent un groupe constitué sont capables de changer leurs habitudes comportementales pour adopter celles de leurs nouveaux congénères. Il s'agit en fait de situations dans lesquelles la fonction visée sous-détermine la forme que peut prendre le comportement, celui-ci étant alors soumis à un choix conventionnel et en partie normatif par le groupe.

---

7. La seule différence étant que quelques mois plus tard, la maîtrise du langage « décolla » chez Alia, pas chez Kanzi...

#### **Intentionnalité et conscience**

La seule étude expérimentale menée sur l'intentionnalité des comportements des primates est celle de WOODRUFF et PREMACK (WOODRUFF et PREMACK 1979), qui place un chimpanzé face à des partenaires humains et à deux boîtes, l'une vide, l'autre contenant de la nourriture. Le chimpanzé sait où se trouve la nourriture mais ne peut y accéder qu'avec l'aide de son partenaire humain.

L'un des partenaires est amical et coopératif : lorsque le chimpanzé lui indique où est la nourriture, il la partage. L'autre garde égoïstement tout pour lui. L'animal réagit très rapidement en donnant les bonnes indications au partenaire coopératif et en induisant l'autre en erreur.

VAUCLAIR critique très judicieusement cette expérience (VAUCLAIR 1992) et propose à la suite de DENNETT une contre-expérience dans laquelle les boîtes seraient transparentes, pour évaluer la capacité du chimpanzé à tromper intentionnellement son partenaire.

L'autre expérience fameuse est le test de reconnaissance dans le miroir inventé par GALLUP qui consiste à peindre sous anesthésie une tache sur le front de l'animal à tester (GALLUP 1970). À son réveil, il est placé devant un miroir et on observe ses réactions. La plupart des espèces se comportent comme face à un congénère (dont la conduite est de plus particulièrement inhabituelle). Seuls les chimpanzés et les orangs-outans « explorent » la tache sur leur front, et observent des parties de leur corps auxquelles ils n'ont d'ordinaire pas accès (cf. figure 3.6).

Des interprétations enthousiastes en termes de « conscience de soi » ont été rapidement avancées, mais l'hypothèse la plus raisonnable actuellement semble plutôt celle d'une représentation mentale du corps.

Il n'en reste pas moins que la plupart des schémas proposés pour expliquer les comportements sociaux les plus complexes – relations « politiques » chez les chimpanzés (DE WAAL 1982) ou tromperie sociale (BYRNE 1995; BYRNE et WHITEN 1988a) – reposent sur l'hypothèse extrêmement forte selon laquelle les primates disposent d'une *théorie de l'esprit*, et ce en dépit d'un manque de preuves expérimentales solides, voire d'études tendant à montrer qu'il y a déjà des différences importantes entre les jeunes enfants et les primates dans le traitement des mécanismes intentionnels (TOMASELLO 2000).

### ◇3. LES SOCIÉTÉS DE PRIMATES

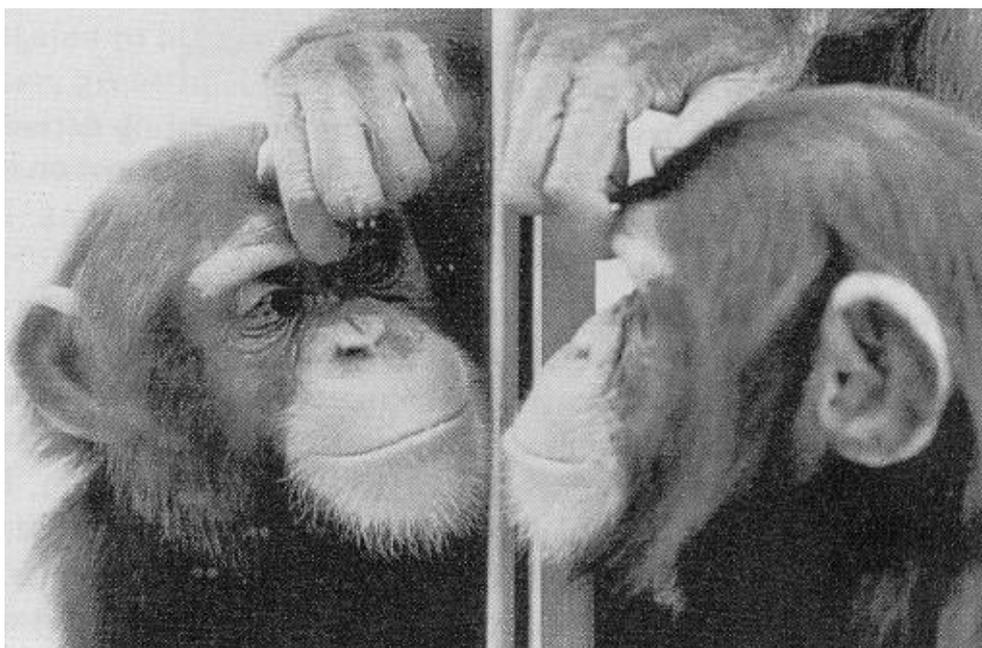


FIG. 3.6 – Le « test du miroir » de GALLUP. Le chimpanzé possède une représentation de son propre corps, voire une « conscience de soi ». — d'après (BYRNE 1995)

#### 3.3.2 Épistémologie et primatologie

##### L'humanité de l'animal

On peut suspecter qu'il ne s'agit pas d'une simple controverse scientifique, mais qu'au contraire ce débat s'ancre dans des options philosophiques fortes, parce qu'elles engagent une certaine conception de l'homme et de ses éventuelles spécificités. On trouve alors deux positions antagonistes (entre lesquelles le moyen terme du continuum est peu représenté) : d'un côté, les tenants de la radicale spécificité humaine, de l'autre les défenseurs d'une « humanité » des primates.<sup>8</sup>

---

8. La querelle déborde même du seul cadre des capacités cognitives des primates pour empiéter sur celui de la classification des espèces : « *La phylogénie des primates est très étudiée, et en même temps le théâtre d'oppositions méthodologiques fortes. [...] C'est parmi les primatologues que l'on trouve de fortes résistances au classement strictement fondé sur la parenté, quand il ne s'agit pas de classifications empruntées de mysticisme, réminiscentes d'une paléontologie teilhardienne. Il convient encore, pour beaucoup de primatologues ou de paléanthropologues très en vue, de réserver une boîte à part pour notre espèce, boîte qui n'a pour effet que de masquer sa parenté avec son espèce sœur. [...] Ainsi, le nom scientifique du chimpanzé deviendrait Homo troglodytes. Cette proposition n'est sans doute pas prête d'être retenue ! L'un de nos collègues remarquait pour plaisanter qu'il serait plus amusant de faire*

### 3.3. DÉBATS AUTOUR DE L'INTELLIGENCE DES PRIMATES

Ces problèmes sont soulignés également par Dominique LESTEL : ainsi par exemple, au-delà des tergiversations sur la qualité de « langage » de la forme de communication maîtrisée par Kanzi, il reste le problème de l'asymétrie des relations entre l'homme et l'animal : « *Que fait vraiment le singe avec ce nouvel outil que l'homme lui a fait acquérir? À ce stade, plus que d'une éthologie au sens strict, peut-être avons-nous besoin d'une ethnologie du primate. [...] Ces études [sur Kanzi] sont présentées un peu vite comme étant de la psychologie animale, alors qu'elles relèvent tout autant de la psychologie humaine. [...] Il en résulte, paradoxalement, que la question éthique est plus volontiers posée, que celle de fond sur quoi portent ces recherches. Car si le primate a des capacités intellectuelles très supérieures à celles des autres animaux, ne doit-on pas lui accorder un statut éthique spécial, entre celui de l'humain et celui de l'animal?* » (LESTEL 1998b)

Les enjeux anthropologiques, philosophiques et sociaux<sup>9</sup> ont tendance à influencer sur la nature et le contenu des théories des primatologues.

En cela, la primatologie est moins « objective » que l'éthologie en général, comme le souligne Rémy CHAUVIN : « *C'est à propos des singes supérieurs que se pose avec une particulière acuité le problème de l'attitude dite « objective » en éthologie. Jusqu'à présent, la fiction de l'organisme en tant que boîte noire s'est montrée efficace en éthologie* », mais ce parti pris pourrait pour les grands singes avoir « un moindre rôle prédictif que l'attitude qui ne refuse pas d'envisager le contenu et l'organisation du monde perceptif à l'intérieur de la boîte noire » (CHAUVIN 1982).

#### Méthodologies

Enfin, nous devons mentionner les conséquences épistémologiques du débat, classique en éthologie, de l'opposition entre l'expérimentation en laboratoire et l'observation sur le terrain (y compris en captivité). Dans le cas de la primatologie, elle se double des enjeux évoqués et suscite des polémiques assez vives.

Les résultats obtenus en laboratoire sont répétables, effectués dans des conditions contrôlées, ils mettent en lumière telle ou telle capacité cogni-

---

*de l'homme un Pan sapiens !* » (LECOINTRE et LE GUYADER 2001).

9. Lancé il y a quelques années, le *Great Ape Project* vise à accorder un certain nombre de droits civiques aux grands singes. Il existe également des courants de « primatologie féministe » (BOBBÉ 1998)...

### ◇3. LES SOCIÉTÉS DE PRIMATES

tive dans une activité ciblée. Ils semblent illustrer idéalement la démarche « **conjectures et réfutations** » (POPPER 1963)

À l'extrême inverse, les données de terrain regorgent d'imprécisions, elles sont sujettes à interprétations, elles s'appuient même parfois sur des anecdotes et des phénomènes singuliers. Mais ce travail de terrain (outre la passion qu'il suscite) conduit à des hypothèses brillantes, stimulantes. Il correspondrait plutôt à un empirisme à la BACON: « **observer et induire** » (BACON 1620).

Nous verrons dans la partie II qu'il ne manque pas de parallèles entre la primatologie et la robotique sociale, jusque dans ces oppositions méthodologiques.

Nous proposons une approche alternative, qui consiste à « **construire et simuler** », au sens évoqué dans l'introduction (cf. § 2.1.2 et § 2.3). Ce sera l'objet des prochains chapitres.

---

# Chapitre quatrième.

## T r a v a u x   e x i s t a n t s

---

*Les relations sociales des primates ont fait l'objet de travaux de simulation encore peu nombreux. Nous allons présenter ici quelques-unes des approches existantes. Certaines ne permettent pas de s'attaquer au problème de la construction des liens sociaux ; d'autres visent à expliquer certains phénomènes sociaux avec des hypothèses minimales, et en cela se rapprochent de nos propres objectifs.*

### 4.1 Les approches inspirées de l'IA classique

L'hypothèse la plus forte posée par certains primatologues, à savoir l'existence d'une théorie de l'esprit chez les primates, leur fournit un cadre de travail particulièrement fécond pour interpréter les comportements complexes qui peuvent être observés.

En contrepartie, ils sont dans l'ensemble peu sensibles à l'utilisation de simulations, que ce soit pour comprendre *comment* ces comportements se construisent, s'intègrent aux activités du groupe, aux facultés cognitives des individus, etc., ou pour comprendre *pourquoi* ils se sont développés jusqu'à leur forme actuelle (alors que d'un point de vue éthologique, ils correspondent à une adaptation évolutive).

Certains travaux toutefois cherchent à formaliser et unifier les descriptions des anecdotes recensées, de façon à dégager les linéaments des comportements généraux, un peu plus abstraits – soit pour prouver la validité de l'hypothèse de la théorie de l'esprit, soit plus modestement pour rapprocher les diverses observations, les classer, et leur donner un statut épistémologique plus « honorable ».

### 4.1.1 Les systèmes de production

Ainsi, BYRNE a proposé un formalisme computationnel de description et d'explication des comportements de tromperie sociale, reposant sur l'utilisation de *règles de production* (BYRNE 1993), utilisant le principe de parcimonie connu en éthologie. Ces règles sont de la forme suivante :

**pattern**  $\longrightarrow$  **procedure**

Lorsqu'une situation d'ensemble est reconnue (*pattern*), une certaine *procédure* est suivie. Dans l'exemple cité précédemment, où un jeune crie comme pour faire croire à une agression, la formalisation est la suivante :

$$\left. \begin{array}{l} \text{besoin d'éliminer } X \\ \text{mère domine } X \\ \neg(\text{mère en vue}) \end{array} \right\} \longrightarrow \text{crier}$$

Ces règles peuvent être construites incrémentalement de façon à refléter les données empiriques. Des conditions peuvent être ajoutées au fur et à mesure pour affiner la description, ainsi que des méta-prémises portant par exemple sur la fréquence d'utilisation des règles. BYRNE estime par ailleurs que si deux règles peuvent s'appliquer simultanément, il faut choisir de préférence celle ayant les prémisses les plus complexes (il s'agit en effet d'un événement *a priori* plus rare). Il est bien sûr possible d'utiliser ces règles pour prédire un certain nombre de comportements ou planifier une séquence d'actions.

L'argument de BYRNE en faveur d'un formalisme cognitif relève du problème de « continuité » entre l'animal et l'homme : « *to be fully adequate, a formalism for animal behaviour should also be extensible to human behaviour. [...] Indeed, the whole point of "cognitive ethology" is surely to allow artificial barriers between species to overcome, and thus discover real differences and continuities in aptitudes* » (BYRNE 1993).

Toutefois, BYRNE fait preuve d'une certaine naïveté vis-à-vis des potentialités offertes par ce formalisme : « *The fact that production systems have already been used successfully to simulate human skills are a guarantee of this potential for comparing animal and human behaviour* » (*op. cit.*).

Autant il semble judicieux d'utiliser un tel formalisme pour automatiser des raisonnements, dans un contexte où des connaissances symboliques et inférentielles peuvent être extraites de tel ou tel savoir-faire (avec les difficultés

#### 4.1. LES APPROCHES INSPIRÉES DE L'IA CLASSIQUE

qu'on connaît), autant il est hasardeux d'extrapoler ce modèle à des individus dont on ignore les capacités cognitives réelles.

On peut comprendre que ce paradigme soit fructueux comme grille d'interprétation, ne serait-ce qu'en projetant de l'intentionnalité sur les phénomènes observés. En revanche, il ne peut servir au mieux qu'à *décrire* les observations sans en fournir d'explication. Les comportements de tromperie sont posés sous leur forme actuelle, le modèle ne fait que fournir un cadre dans lequel les formuler.

Quant à l'utilisation concrète de ce modèle dans une simulation informatique, elle se heurte à deux obstacles :

1. BYRNE ne donne aucune indication sur l'origine de ces règles : sont-elles innées ? sont-elles apprises ? peuvent-elles évoluer ? – autant de points qui, pour difficiles à trancher qu'ils soient du point de vue du biologiste, sont nécessaires à l'informaticien dans un but de simulation.
2. Les prémisses et les procédures proposées dans ce modèle sont trop abstraites et trop subjectives (l'un des exemples donnés par BYRNE est le suivant : « le dominé semble *coupable* »!). On peut supposer qu'un primatologue saura, sur le terrain, noter les observations pertinentes et reconnaître telle ou telle situation, mais pour cela, une *interprétation humaine* est nécessaire, alors que pour la réalisation d'une simulation on a besoin d'éléments objectifs.

##### 4.1.2 Les scripts

Partant des travaux précédents, Robert WORDEN a formulé une théorie computationnelle de l'intelligence sociale des primates, toujours dans une perspective cognitiviste. Il soutient les thèses suivantes (WORDEN 1996) :

1. La cognition sociale rencontrée chez les primates repose sur une représentation *symbolique* et *discrète* des situations d'interaction sociale.
2. Les scripts sont une structure adéquate pour représenter ces connaissances sociales. On peut construire une théorie de la cognition sociale *complète* et cohérente en utilisant des scripts et trois opérations sur ces structures.
3. Cette théorie rend compte de nombreuses observations (telles que la reconnaissance de liens de parenté et de dominance, les cris d'alerte, etc.)

◇4. TRAVAUX EXISTANTS

4. Elle conduit à une intelligence sociable adaptative, capable d'un apprentissage rapide de nouvelles régularités dans les phénomènes sociaux, en parfait accord avec le comportement des primates dans leur milieu.

Les *scripts* utilisés par WORDEN dans sa théorie sont des structures arborescentes destinées à représenter soit des faits (*scripts factuels*, cf. figure 4.1), soit des lois comportementales (*scripts de règles*, cf. figure 4.2). Un script met en relation deux *scènes*, composées d'*acteurs* accomplissant des *actions*. Certains champs peuvent être remplacés par des variables. Cette structuration fixe permet, contrairement aux règles de production de BYRNE, de construire une description manipulable des situations sociales.

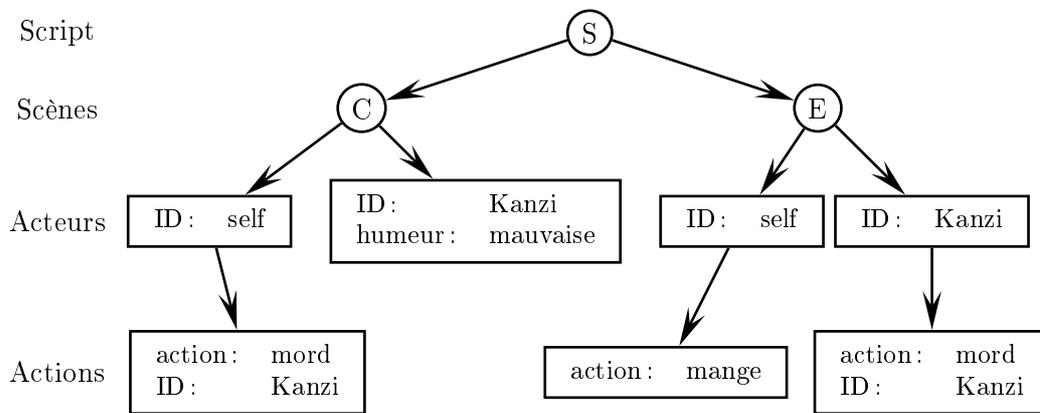


FIG. 4.1 – Un exemple de script factuel. Deux scènes (C et E) se succèdent dans le temps : dans la première, l'agent (self) mord Kanzi, qui est de mauvaise humeur. Dans la seconde, alors que l'agent (self) est occupé à manger, Kanzi le mord.

Trois opérations sont définies pour manipuler ces structures :

1. L'*unification* ( $S_1 \cup S_2$ ), assez semblable à l'opération du même nom dans le langage PROLOG (CLOCKSIN et MELLISH 1979) ; elle sert par exemple à produire un nouveau script factuel à partir d'un script de règle appliqué dans une situation décrite par un premier script factuel.
2. L'*inclusion* ( $S_1 \subseteq S_2$ ) qui dénote le fait que toutes les informations contenues dans  $S_1$  le sont aussi dans  $S_2$  : c'est donc l'inverse de la subsomption en programmation logique (par exemple, un script de règle  $R$  qui sous-tend un script factuel  $F$  est inclus dans  $F$ ).

#### 4.1. LES APPROCHES INSPIRÉES DE L'IA CLASSIQUE

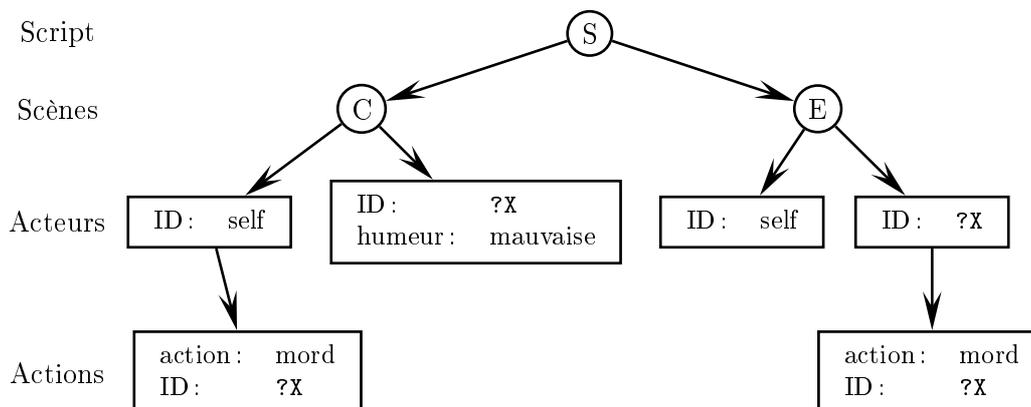


FIG. 4.2 – Un exemple de script de règle. Lorsque l'agent a été confronté à des situations proches du script de la figure 4.1, il peut procéder par induction pour généraliser en un script de règle dans lequel les agents individuels ont été remplacés par des variables (?X), et où seuls les caractéristiques pertinentes ont été retenues. Le script résultant peut alors être utilisé pour anticiper le résultat des actions de l'agent.

3. L'*intersection* ( $S_1 \cap S_2$ ), opération qui cherche à maximiser le recouvrement d'information de deux scripts, en n'en retenant que la structure commune, et en créant le cas échant des variables. On peut s'en servir pour généraliser deux faits ( $R = F_1 \cap F_2$  peut être un bon candidat pour la règle sous-jacente à  $F_1$  et  $F_2$ ).

À l'aide de ces outils, WORDEN propose de mettre en œuvre un apprentissage de règles générales en procédant à l'intersection de scripts déjà acquis avec des scripts factuels issus de l'expérience des individus.

La rapidité d'une telle méthode, comparée à un apprentissage par essais-erreurs, correspond assez aux données empiriques de la primatologie : l'acquisition des comportements de tromperie sociale, notamment, *est et doit être* suffisamment rapide pour ne pas être détectée en tant que tromperie.

Ce modèle a l'avantage d'offrir un support assez formel aux règles de production esquissées par BYRNE. Toutefois, il laisse en suspens le problème difficile qui sous-tend plus généralement la plupart des systèmes cognitifs : celui de l'ancrage dans le monde réel et dans les perceptions que les individus ont de leur environnement, y compris de leur environnement social.

#### ◇4. TRAVAUX EXISTANTS

En effet, WORDEN reconnaît explicitement supposer que les primates disposent d'un module cognitif spécialisé de traitement des informations sociales. Ce module a pour rôle non seulement de reconnaître certains comportements comme sociaux, ce qui reste une hypothèse faible, mais même de traduire un ensemble de phénomènes perçus sous la forme d'une représentation sociale extrêmement structurée (des scripts factuels et leurs constituants).

Cette hypothèse, outre les difficultés qu'elle soulève quant à sa validation ou son infirmation par les données empiriques, a le tort d'être improductive par rapport à nos objectifs à terme (l'utilisation de modèles sociaux pour des agents opérant dans un environnement ouvert, dans lequel il n'est pas possible de caractériser *a priori* tous les événements susceptibles de se produire).

## 4.2 Les approches inspirées de l'éthologie

Les travaux cités ci-dessus, fortement influencés ou explicitement inspirés par l'IA classique, peuvent éventuellement aider le primatologue à formaliser ses observations, à les rationaliser, les généraliser. Pour cela, ces modèles s'appuient sur des hypothèses cognitives fortes et sur l'existence *a priori* d'un module de construction des représentations sociales, dont ils sont impropres à expliquer le fonctionnement.

Or, par rapport à nos objectifs, il est préférable de disposer de modèles capables précisément de rendre compte de cette construction des représentations sociales. Nous allons donc examiner des travaux allant dans ce sens, inspirés plutôt par l'éthologie.

### Les travaux de HOGEWEG et HESPER

Le courant de recherche initié par P. HOGEWEG et B. HESPER à travers le paradigme MIRROR (HOGEWEG et HESPER 1983; HOGEWEG et HESPER 1985), compte parmi les premières tentatives de modélisation biologique par *comportements individuels*. Illustrée en particulier par l'étude des relations interindividuelles existant dans les colonies d'insectes (*Bombus terrestris*), la méthodologie de simulation MIRROR s'appuie sur les caractéristiques suivantes :

- une définition locale des entités (c'est-à-dire des individus simulés) ;
- une définition empirique des relations entre individus ;
- l'observabilité de l'ensemble des événements survenant dans le système.

## 4.2. LES APPROCHES INSPIRÉES DE L'ÉTHOLOGIE

Cette approche est assez voisine de celle qu'on retrouvera dans les simulations multi-agents (DROGOUL 1993). Des modèles construits de la sorte permettent en effet d'aborder des comportements macroscopiques (collectifs) à partir de comportements microscopiques (individuels), sans requérir d'hypothèses quant aux comportements collectifs. Ceux-ci ne sont produits que par les interactions entre individus et par les interactions entre ces individus et leur environnement – résultant elles-mêmes de comportements explicitement donnés aux agents.

Les simulations menées au moyen de MIRROR ont permis de montrer que des comportements individuels simples suffisent à reproduire la structure des interactions sociales observées dans la colonie. Des modèles analogues ont été utilisés pour la répartition spatiale chez les primates (TE BOEKHORST et HOGEWEG 1994).

### Les travaux de C. HEMELRIJK

Ces travaux ont été poursuivis en primatologie par Charlotte HEMELRIJK (HEMELRIJK 1998b; HEMELRIJK 2000; HEMELRIJK 1998a) à travers la notion d'« éthologie synthétique » et de « laboratoire virtuel » déjà évoqués § 2.2.3.

Les expériences qu'elle a conçues visent à suggérer des hypothèses nouvelles en primatologie, en essayant d'expliquer un certain nombre de phénomènes observés – tout particulièrement la répartition spatiale des membres d'un groupe autour des dominants – en fonction de paramètres aussi simples et peu nombreux que possible.

Les agents mis en scène dans les simulations sont des modèles réduits de primates, au sens où ils ne font que passer leur temps à des interactions agonistiques. Leurs comportements sont entièrement réactifs et régis en grande partie par la perception des autres agents :

- Lorsqu'un agent en voit un autre dans un rayon critique (*PerSpace*), il se lance dans une interaction de dominance avec celui-ci. S'il gagne, il se dirige vers son adversaire, sinon il s'enfuit.
- Lorsqu'il n'y a personne dans le rayon critique, mais que des agents sont visibles à une distance proche (*NearView*), l'agent continue de se déplacer dans sa direction actuelle.
- Si un agent perçoit d'autres agents entre *NearView* et une distance maximale *MaxView*, il se dirige vers eux.

#### ◇4. TRAVAUX EXISTANTS

- Enfin, s’il ne voit personne, il tourne aléatoirement pour chercher des congénères.

À partir de ce cadre simplifié, diverses hypothèses peuvent être testées indépendamment les unes des autres pour estimer l’influence de tel ou tel paramètre sur la forme que prend la société : règles de regroupement, valeur des distances *PerSpace*, *NearView* et *MaxView*, modalités des interactions de dominances, taille du groupe...

La société artificielle ainsi constituée peut alors être observée et analysée *avec les mêmes moyens qu’un groupe réel* : mesure de la fréquence des attaques, distances moyennes entre les agents, etc.

Ainsi, s’il ne s’agit pas à proprement parler d’une « simulation », mais plutôt d’expérimentations sur un modèle assez abstrait et sur des phénomènes sociaux bien ciblés, cette approche permet d’une part de déterminer objectivement l’influence de la forme du modèle sur la forme de la structure sociale résultante, et d’autre part, en comparant dans les mêmes termes cette structure avec des observations menées sur de vrais primates, il est possible d’estimer la validité des prédictions du modèle et la plausibilité biologique de ce que produit le système.

Par suite, lorsque les phénomènes simulés concordent avec des observations empiriques, on peut estimer que les hypothèses sous-jacentes au modèle informatique ont une légitimité à expliquer le comportement des primates réels.

Toutefois, on peut émettre une réserve quant au choix de n’introduire dans le système que des comportements sociaux, tout spécialement dans le cas d’études sur la répartition spatiale, dans la mesure où d’autres activités peuvent entraîner des biais dans la gestion de l’espace, qui disparaissent alors du modèle.

Par ailleurs, l’extrême abstraction du modèle empêche de poser la question de la *fonction* de la socialité, par rapport à l’acquisition de certaines ressources par exemple. Il n’en reste pas moins que cette démarche se conforme parfaitement aux principes de parcimonie et d’incrémentalité que nous avons choisis (cf. § 2.2.2).

### 4.3 Notre point de départ : la plate-forme MACACA

La démarche intellectuelle d’HEMELRIJK a exercé une influence certaine sur nos propres travaux ; toutefois, plutôt que de recréer *ex nihilo* un labora-

toire virtuel, nous sommes partis de travaux de simulation de primates réalisés au LAFORIA (prédécesseur du LIP6) : la plate-forme MACACA.

La plate-forme de simulation MACACA (acronyme pour *a Multi-Agent Computer simulation of Animal Communities based on Alliances*, et dénomination latine du genre des Macaques) a été développée par Steffen LALANDE en collaboration avec Bernard THIERRY (Laboratoire de Psychophysiologie de l'Université Louis Pasteur à Strasbourg).

L'objectif des simulations menées au moyen de cet outil (LALANDE et al. 1995) était de tester des hypothèses de B. THIERRY concernant la formation de structures hiérarchiques plus ou moins rigides chez les diverses espèces de macaques. On y observe en effet d'importantes variations, allant du « despotisme » où les relations hiérarchiques sont fortement marquées à l'« égalitarisme » dans lequel la contestation, violente ou non, est fréquente et conduit à un équilibre des forces (THIERRY 1985; THIERRY 1990).

### E-EMF

La plate-forme, programmée en SMALLTALK, s'appuie sur les travaux d'éthomodélisation d'Alexis DROGOUL (DROGOUL 1993), plus exactement sur EMF (*EthoModelling Framework*).

EMF permet de définir les comportements des agents du système sous la forme de *tâches* constituées par le chaînage d'un certain nombre de *primitives comportementales* correspondant aux actions élémentaires de l'agent (cf. figure 4.3). À chaque pas de temps dans la simulation, chaque agent exécute une primitive correspondant à la tâche active. Chaque primitive retourne une valeur booléenne qui peut être utilisée dans le chaînage au sein des tâches.

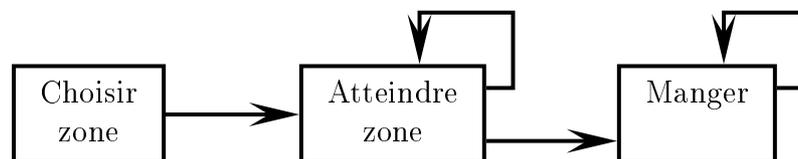


FIG. 4.3 – Un exemple de tâche EMF. Ici, le comportement consiste à choisir une zone d'alimentation, s'y rendre et manger. Les primitives sont chaînées les unes aux autres en fonction de la valeur booléenne qu'elles renvoient, avec possibilité de bouclage.

#### ◇4. TRAVAUX EXISTANTS

À chaque tâche est attaché un certain niveau de priorité, fixé par le modélisateur lors de la définition des comportements. Ces tâches sont déclenchées par des stimuli qui peuvent être exogènes (signal émis par un autre agent par exemple) ou endogènes (motivation à attaquer, besoin de s'alimenter, etc.). Dans MACACA, une version modifiée d'EMF, baptisée E-EMF (pour *Enhanced EthoModelling Framework*) permet de prendre en compte plus généralement des préconditions qui peuvent être des stimuli, des événements survenant lors de la simulation et susceptibles d'être perçus par l'agent (combat entre deux individus) ou plus généralement toute expression retournant une valeur booléenne.

Lorsque les préconditions d'une tâche sont toutes vérifiées, celle-ci devient éligible. C'est alors, parmi les tâches éligibles, celle dont la priorité est la plus élevée qui est choisie pour être exécutée. Les actions et communications des agents sont réifiées sous la forme d'événements (comportant le type d'événement et l'identité des participants).

À chaque pas de temps de la simulation, un agent exécute donc successivement les étapes suivantes :

1. perception des événements présents dans l'environnement, dans un voisinage donné ;
2. déclenchement d'un certain nombre de « règles cognitives » correspondant aux événements perçus et calcul des stimuli ;
3. sélection des tâches activables (celles dont les préconditions sont remplies ou tâches en cours) ;
4. choix d'une ou plusieurs tâches à exécuter (E-EMF permet en principe le parallélisme de certaines tâches, mais cette possibilité n'a pas été exploitée en pratique) : une tâche en cours peut être interrompue si son niveau de priorité est inférieur à celui des autres tâches éligibles.
5. exécution de la primitive suivante de la tâche en cours si celle-ci n'a pas été modifiée, ou de la première primitive de la nouvelle tâche.

### **L'environnement**

Les simulations menées par Steffen LALANDE n'ayant pas pour but d'étudier l'impact des conditions écologiques sur l'organisation sociale, l'environnement utilisé n'est pas particulièrement complexe : il s'agit d'un rectangle contenant des zones de repos et d'alimentation de diverses tailles. L'action des

agents dans ces régions (comme « manger » ou « se reposer ») ne font que produire des événements, sans modifier l'environnement : la compétition qui peut se produire entre les agents n'est donc pas directement reliée à des ressources réelles, mais médiatisée par l'espace.

### Comportements des agents

Les agents sont dotés de comportements élémentaires (sur lesquels nous reviendrons plus en détail dans le chapitre suivant) :

- 1° des comportements consommatoires : se reposer, se nourrir ;
- 2° des comportements agonistiques (liés à l'agression) : attaquer, combattre, fuir ;
- 3° des comportements affiliatifs : initiative de toilettage et toilettage.

Suivant une démarche incrémentale, les classes d'agents successivement étudiées complexifient le modèle de départ pour l'enrichir comportementalement. Les primates disposent en effet de signaux de menace et de soumission permettant d'éviter d'en venir aux combats. Des comportements de signalisation ont donc été ajoutés (menace, soumission, inhibition de l'agressivité). Quelques expériences ont également été menées sur la possibilité de coalition.

Le comportement social des agents s'inspire en grande partie des travaux de HOGEWEG et HESPER, mentionnés précédemment (HOGEWEG et HESPER 1985) et de règles élaborées avec B. THIERRY :

- L'occurrence d'une interaction affiliative entre deux individus augmente la probabilité qu'ils renouvellent cette interaction et diminue la probabilité qu'ils se combattent.
- L'occurrence d'un combat entre deux individus diminue la probabilité qu'ils entament une interaction affiliative et accroît la probabilité qu'un combat ultérieur ait la même issue (i.e. le même vainqueur)

Autrement dit, les choix des partenaires pour les comportements affiliatifs, ainsi que l'issue des combats, sont supposés suivre des dynamiques engendrées par des rétroactions positives.

Pour ce faire, les agents disposent d'une *mémoire des interactions*. Dans les premiers modèles d'agents, il ne s'agit pas à proprement parler d'une *représentation* au sens défini dans (VAUCLAIR 1992), dans la mesure où elle ne sert pas d'emblée à *anticiper* l'issue possible des interactions (cela n'est mis en œuvre que dans un second temps).

#### ◇4. TRAVAUX EXISTANTS

La mémoire des interactions d'un agent  $A_i$  peut être représentée par un repère à deux dimensions, dans lequel tout individu  $A_j$  est associé à un point. L'abscisse de ce point correspond au degré d'affiliation entre  $A_i$  et  $A_j$ , son ordonnée est une mémoire cumulée des résultats des interactions agonistiques (cf. figure 4.4).

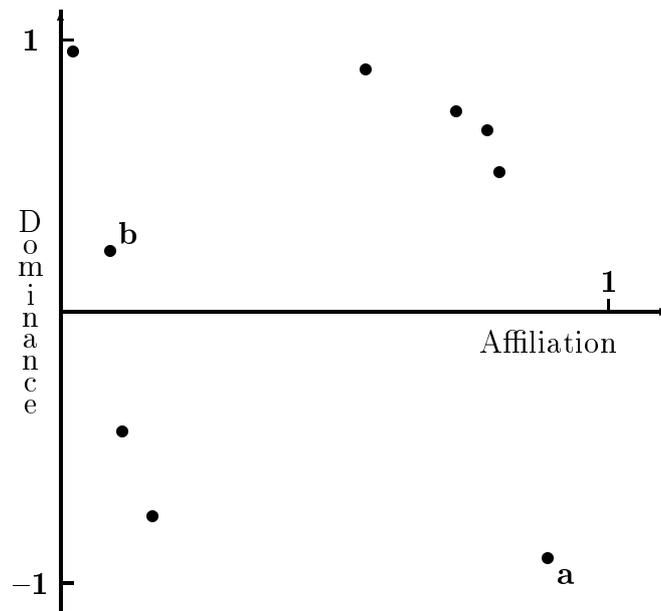


FIG. 4.4 – *Mémoire des interactions d'un agent dans MACACA. Les points représentent les autres agents : leur abscisse correspond au degré d'affiliation avec l'agent observé, leur ordonnée à la relation de dominance. Ici, l'agent étudié a une relation d'affiliation forte avec l'agent **a**, faible avec **b** ; il est dominé par **a** et domine **b**.*

Le degré d'affiliation est utilisé pour le choix du partenaire d'une interaction affiliative ; cette interaction à son tour renforce le degré d'affiliation de façon symétrique chez chacun des participants.

Inversement, l'issue d'un combat modifie asymétriquement la mémoire des interactions. Si l'ordonnée d'un agent  $A_j$  dans la mémoire de  $A_i$  est positive,  $A_i$  a plus de chances de l'emporter. Si c'est le cas, il augmente encore l'ordonnée de  $A_j$  (tandis que dans sa mémoire,  $A_j$  diminue l'ordonnée de  $A_i$ ).

Les variations suivent le principe énoncé dans (HOGEWEG et HESPER 1985), selon lequel les renforcements sont faibles lorsque l'issue d'une interaction est en accord avec la relation de dominance existante, tandis qu'un résultat contraire aux attentes produit un changement profond des relations de dominance.

Pour chaque agent, la mémoire des interactions doit être initialisée, d'une part en annulant la mémoire des conflits (ordonnées), d'autre part en plaçant à des abscisses non nulles les individus apparentés (selon un paramètre appelé *biais de parenté*).

## Résultats

Dans le modèle le plus simple (pas d'anticipation de l'issue des combats), la fréquence des rencontres agonistiques ne diminue pas (puisqu'il n'y a aucune forme d'apprentissage).

Dans les modèles utilisant la mémoire des interactions comme représentation de la relation de dominance, que ce soit ou non en conjugaison avec les signaux de menace et de soumission, le nombre d'agressions diminue au cours de la simulation. En effet, le renforcement positif établi, par brisure de symétrie, des relations de dominance entre individus (en amplifiant les divergences provoquées par les premiers combats), qui sont prises en compte dans le déclenchement ou l'inhibition des interactions sociales.

Dans ces situations se constitue une hiérarchie fortement *linéaire*, dans laquelle la mémoire des combats (ordonnées des agents dans la mémoire des interactions) atteint rapidement les extrêmes. Aucune contestation des rapports hiérarchiques ne peut plus alors avoir lieu. Selon (LALANDE et al. 1995), les modèles expérimentés n'ont ainsi pu reproduire que des formes sociales observées dans les sociétés de macaques « despotiques ».

## 4.4 Bilan

Les travaux portant sur la simulation des comportements sociaux des primates restent à l'heure actuelle peu nombreux. Une partie des primatologues s'attache d'ailleurs, plutôt qu'à la *simulation* proprement dite des comportements, à leur *formalisation* (cf. § 4.1.1).

Ces modèles, qui se prêtent facilement à une implémentation mettant en jeu des techniques issues de l'IA classique (cf. § 4.1.2), présupposent l'existence de modules cognitifs dédiés au traitement des informations sociales, capables de produire des représentations déjà sophistiquées (prédicats) à partir des perceptions.

Cela pose évidemment le problème de l'ancrage de ces représentations dans le monde réel et de leur *construction*. Or, dans un environnement ouvert,

#### ◇4. TRAVAUX EXISTANTS

il ne sera pas possible de définir *a priori* l'ensemble des prédicats pertinents.

Les approches alternatives, comme celle de C. HEMELRIJK (cf. § 4.2), sont plus prometteuses, dans la mesure où elles proposent de simuler des comportements *les plus simples possibles*, en n'imposant que très peu d'hypothèses sur les mécanismes cognitifs des primates.

Toutefois, l'objectif de C. HEMELRIJK est d'explorer les dynamiques par lesquelles une certaine forme sociale s'instaure. Aussi, toute considération extérieure à la socialité est éliminée des modèles et peut difficilement être réintroduite ultérieurement.

Nous avons donc finalement choisi de repartir de simulations existantes, celles de S. LALANDE (cf. § 4.3), en essayant de comprendre pourquoi le modèle utilisé dans MACACA ne produit qu'une forme de socialité figée, alors que nous souhaitons au contraire que celle-ci puisse se reconfigurer (selon les besoins du groupe notamment).

Nous devons de plus monter d'un degré dans le niveau d'abstraction pour ne nous intéresser qu'à des primates « virtuels » (et non telle ou telle espèce). En effet, ces simulations, en l'état, sont trop spécifiquement ciblées sur les comportements des macaques : or nous sommes plutôt intéressés par des comportements et des représentations permettant d'instaurer un ordre social (la hiérarchie par exemple), et ce de façon générale, pour pouvoir réutiliser les mêmes modèles dans d'autres applications.

Ainsi, il nous faut garder un niveau d'étude tout à la fois suffisamment général pour pouvoir être facilement transposé à d'autres applications, et relativement proche des descriptions données par les primatologues pour pouvoir comparer les résultats obtenus à des données empiriques, et juger par là de la fécondité de nos modèles.

---

# Chapitre cinquième.

## Modèles de représentations sociales

---

*Nous allons maintenant présenter les modèles proposés pour rendre compte des capacités des primates à se forger une représentation des relations sociales existant entre leurs congénères. Nous montrerons comment les difficultés évoquées précédemment ont été résolues, quels résultats ont été obtenus avec nos modèles.*

### 5.1 La représentation sociale

#### Objet de la simulation

La plupart des comportements sociaux des primates présentent un intérêt assez évident en vue d'une transposition à un environnement informatique ouvert, dans la mesure où, d'une part, ils offrent un moyen efficace d'anticiper ou de résoudre les conflits qui peuvent survenir (entre primates ou entre agents), et d'autre part, ils suscitent des interprétations fortement anthropomorphiques, ce qui peut s'avérer un inconvénient en primatologie mais constitue assurément un atout pour interagir avec des humains.

Par ailleurs, nous avons fait le choix, pour les raisons évoquées au chapitre 2, d'une *approche incrémentale* dans la construction de nos modèles. Aussi, il nous a semblé important d'amorcer le « cercle vertueux » en simulant des compétences sociales relativement simples, puis en essayant de les transposer à d'autres situations (ce qui sera l'objet du chapitre suivant), puis enfin en tentant de les mettre à l'épreuve d'applications réelles (cf. partie II et les problèmes soulevés).

Nous avons donc commencé ces travaux avec la recherche de modèles capables de rendre compte de l'une des capacités cognitives partagées par tous les primates, et qui fait leur spécificité : **la représentation des rela-**

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

**tions sociales entre congénères** (cf. § 3.2.3), ce que V. DASSER nomme des « proto-concepts » (DASSER 1988).

Il ne s'agira pas ici de tenter de *simuler, de façon réaliste, un groupe de primates d'une espèce donnée*, ce qui était le but de MACACA, mais plutôt de rechercher des modèles de représentations ou de comportements qui, s'appuyant sur des caractéristiques communes à la plupart des primates, sont capables de reproduire artificiellement des phénomènes observés empiriquement chez ces espèces.

Il est à craindre en effet qu'une simulation qui se focaliserait sur l'étude particulière d'une espèce, voire d'un groupe, soit rendue inexploitable par une accumulation de détails, de paramètres qui, nécessitant un ajustage précis, tendraient à masquer des mécanismes sous-jacents plus généraux. On serait alors dans la situation paradoxale d'un modèle hyperréaliste inutilisable en-dehors du champ d'application pour lequel il a été conçu. S'agirait-il alors toujours d'un modèle ? Ou au contraire, ne serait-il plus qu'une description algorithmique incapable de tenir lieu d'explication ?

Nous préférons donc aborder le problème par la simulation de primates « abstraits », à la façon de C. HEMELRIJK (cf. § 4.2). Ainsi, il sera ensuite plus facile de transposer ces modèles à d'autres types de SMA.

Nous allons donc, après avoir défini une base comportementale vraisemblable, proposer un modèle de représentation des relations sociales et étudier son effet sur le groupe. Nous appliquerons pour cela les principes méthodologiques définis au chapitre 2 : le principe de parcimonie doublé du principe d'incrémentalité pour la construction des modèles d'agents et l'expérimentation d'une part ; une observation mettant en vis-à-vis interprétation intentionnelle (fonctionnelle) et explication causale d'autre part.

### **Point de départ**

Des travaux évoqués précédemment, ce sont ceux qui se rapprochent le plus d'une tentative de construction incrémentale de modèles pour reproduire des phénomènes observés chez les primates (à savoir ceux de C. HEMELRIJK et de S. LALANDE) qui coïncident le plus à nos besoins (en vue d'une utilisation de mécanismes sociaux dans d'autres systèmes). Nous avons donc choisi une approche assez similaire dans son principe.

Nous sommes partis, pour nos propres expériences, de la plate-forme MACACA qui constituait un bon point de départ pour simuler les phénomènes

de reconnaissance des liens sociaux. En particulier, les comportements élémentaires implantés dans les simulations existantes ont fourni une base permettant une étude comparative avec les dynamiques observées dans les modèles de S. LALANDE.

En revanche, les choix retenus dans MACACA quant aux *représentations* des agents nous ont semblé particulièrement restrictifs.

En effet, comme chaque agent ne dispose que d'une mémoire des interactions avec chacun des autres agents (cf. § 4.3), il peut estimer la force relative d'un de ses congénères, donc estimer si un agent est dominant ou dominé par rapport à lui-même et agir en conséquence. Mais il ne peut pas se représenter les relations sociales existant entre les autres agents. Il est de plus *obligé* d'affronter un agent qu'il n'aurait pas encore rencontré pour le positionner « correctement » dans sa représentation sociale, puisqu'il ne peut tenir compte du rang que peut avoir acquis cet agent auprès des autres.

Cela pourrait expliquer d'ailleurs l'échec partiel de MACACA dans ses objectifs de simulation. La rétroaction positive mise en œuvre pour établir la relation de dominance conduit aisément, après une période de fluctuations, à l'établissement d'une hiérarchie fortement linéaire, dont il devient alors difficile de sortir. Il faut donc se focaliser tout particulièrement sur la forme de mémoire à utiliser dans nos modèles.

### Fécondité des modèles

Le problème de l'évaluation des modèles, dans notre cas, ne se posera pas vraiment en termes de « validation » par rapport à telle ou telle observation empirique de primates particuliers, puisque les entités simulées sont assez abstraites.

Au contraire, il nous faudra estimer si les résultats obtenus, qu'ils soient quantitatifs ou qualitatifs, correspondent bien aux fonctions qui ont été assignées comme but au modèle, et comment cela se traduit dans les comportements d'agents. Ceux-ci doivent évidemment être comparables qualitativement à des phénomènes observés chez les primates, liés à la représentation sociale du groupe.

On pourra également chercher à voir si le modèle permet de rendre compte de plus de choses que ce pour quoi il a été conçu. Autrement dit, il s'agit d'abord de savoir si le modèle est *fécond*, s'il est capable de produire

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

des formes de socialité intéressantes, et autant que possible proches de celles des primates.

Les éthologues ont défini un certain nombre de critères pour caractériser les sociétés animales ; nous nous en servons pour donner à des sociétés artificielles, produites par la simulation, des caractérisations en termes éthologiques, que nous pourrions comparer aux interprétations proposées par les primatologues.

## 5.2 Cadre méthodologique

### 5.2.1 CASSIOPÉE

Les modèles de représentation sociale décrits ici ont vocation à être transposés dans d'autres systèmes : cela signifie en particulier que la *fonction* réalisée par le groupe, c'est-à-dire la construction collective de représentations sociales permettant de s'organiser, devra être en grande partie indépendante de la nature exacte des comportements qui la produisent.

Cela pose un problème méthodologique d'ordre informatique, qui rejoint celui de la réutilisabilité des composants logiciels. Ces obstacles surviennent également lorsque l'on cherche à concevoir un SMA pour une application donnée, sans avoir à réinventer de toutes pièces des mécanismes d'organisation.

Aussi depuis quelques années a-t-on vu se multiplier des efforts de recherche en IAD sur les méthodologies de conception de SMA.

Nous nous appuyons ici sur la méthodologie CASSIOPÉE, qui est très fortement inspirée de l'éthologie (DROGOUL et COLLINOT 1997) en ce qu'elle permet une spécification intentionnelle des fonctions du système à concevoir, tout en laissant une liberté assez considérable quant aux architectures effectivement mises en œuvre pour l'implémentation (COLLINOT et DROGOUL 1998).

CASSIOPÉE a été utilisée avec succès comme méthodologie de conception, notamment pour l'organisation d'une équipe d'agents robotiques dans le cadre des compétitions de « robots footballeurs », RoboCup<sup>1</sup> (COLLINOT et al. 1996a; COLLINOT et al. 1996b; DROGOUL et COLLINOT 1998). Elle a donné naissance, depuis, à ANDROMÈDE, qui intègre au cadre général de CASSIOPÉE des techniques d'apprentissage (DROGOUL et ZUCKER 1998).

---

1. On pourra se reporter à (ASADA et al. 1998) et au chapitre 7 pour plus de détails sur RoboCup.

### La notion de *rôle*

Les *rôles* sont un concept central de CASSIOPÉE. Un rôle est un ensemble de comportements (décrits d'un point de vue fonctionnel) qui peut être assumé par un ou plusieurs agents ; un même agent peut également jouer plusieurs rôles (i.e. assurer plusieurs fonctions) simultanément.

L'hypothèse sous-jacente à CASSIOPÉE, qui a été reprise par d'autres méthodologies de conception (GUTKNECHT et FERBER 1999, par exemple), est la possibilité de *répartir les rôles en trois niveaux* par rapport aux degrés d'organisation du système :

1. les rôles individuels ;
2. les rôles relationnels ;
3. les rôles organisationnels.

Nous la désignerons sous le nom d'**hypothèse de séparation des rôles**.

Les rôles individuels concernent les comportements individuels nécessaires à la réalisation de la tâche collective considérée (dans le cas de la conception d'une application multi-agent). Les rôles relationnels gèrent tout ce qui touche aux interactions, aux communications entre agents. Enfin, les rôles organisationnels structurent les groupes d'agents, par exemple en mettant en place des mécanismes de coordination ou de coopération.

Ces rôles ne sont pas rigoureusement indépendants : si l'on se reporte à la figure 1.1 page 6, on doit pouvoir retrouver les influences de l'organisation sur les agents et des agents sur l'organisation. Il est donc indispensable de déterminer comment ces rôles sont liés les uns aux autres.

Il faut rajouter deux étapes pour faire le pont entre les trois niveaux de rôles :

- Les *dépendances mutuelles* qui existent entre les rôles individuels, au sens par exemple de (CASTELFRANCHI et al. 1992), servent à établir les rôles relationnels : deux rôles individuels  $R_1$  et  $R_2$  peuvent être en conflit ;  $R_1$  peut permettre la réalisation de  $R_2$  ou la faciliter, ou encore nécessiter  $R_2$ , etc. On retrouvera alors au niveau relationnel une typologie Facilitateur/Dépendant.
- Les rôles relationnels à leur tour peuvent servir de base à la constitution de *groupes*. Les rôles organisationnels vont servir à gérer la participation, l'initiation, la dissolution de ces groupes.

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

Le parcours de conception suit alors cinq étapes (cf. figure 5.1) dans lesquelles on passe alternativement d'un point de vue local, celui des rôles joués par les agents, à un point de vue global, celui de l'organisation de ces rôles les uns par rapport aux autres. Il peut être suivi de façon ascendante (en partant des rôles individuels, et en n'ayant éventuellement de rôles relationnels et organisationnels qu'émergents) ou descendante (en spécifiant les formes d'organisation voulues et en déduisant ensuite une décomposition de la fonction désirée en rôles relationnels et individuels).

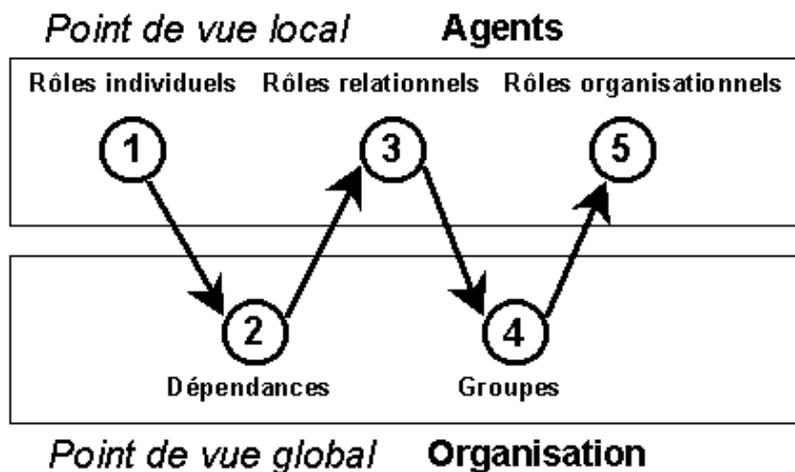


FIG. 5.1 - *Les liens entre rôles dans CASSIOPÉE (voir texte).*  
— d'après (DROGOUL 2000)

### Niveaux d'organisation et représentation sociale

La séparation entre rôles est absente de MACACA. Il nous semble pourtant qu'elle constitue une clef pour modéliser la représentation des relations sociales au sein du groupe.

On peut en effet identifier clairement des comportements relevant de fonctions purement individuelles (se déplacer, s'alimenter, se reposer), et d'autres qui font intervenir des dépendances relationnelles entre individus (les comportements agonistiques et affiliatifs, qui s'appuient sur les liens sociaux d'agent à agent). La relation de dominance tout comme la relation d'affiliation chez les primates relèvent de rôles *relationnels*, dans la mesure où elles concernent les interactions des agents pris deux à deux.

Toutefois, lors d'un combat, d'autres agents peuvent participer et former une coalition avec l'un ou l'autre protagoniste, mais on peut considérer en suivant l'hypothèse la plus faible que le déclenchement d'une coalition repose sur des liens interindividuels, et non sur l'appartenance explicite à un groupe (cf. figure 5.2). Cette interprétation de nombreuses situations d'interaction coïncide d'ailleurs avec les observations de H. KUMMER sur les relations triadiques chez les primates (KUMMER 1967).

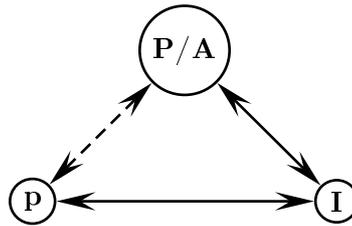


FIG. 5.2 – *Dépendances triangulaires entre trois rôles : un protecteur ou auxiliaire (P/A), un protégé (p) et un intrus (I). Entre les rôles P/A et p se développent des relations affiliatives (pointillés), tandis que les relations avec I sont agonistiques (traits pleins)*

Nous sommes donc amenés à considérer explicitement deux niveaux d'organisation au minimum : le niveau des rôles individuels et celui des rôles relationnels. Les rôles individuels sont donnés par le choix des comportements liés au domaine envisagé (ici, les comportements individuels des primates) ; les rôles relationnels vont s'appuyer quant à eux sur des représentations sociales qu'il va falloir définir (cf. § 5.3).

Nous discuterons plus loin de l'introduction du troisième niveau, celui des rôles organisationnels, pour la simulation des mécanismes de coalition.

### 5.2.2 Quelques définitions

Avant de présenter les modèles choisis, nous allons définir les grandeurs mesurées lors des expériences, ainsi que les paramètres qui servent de fondements aux représentations sociales.

### L'indice de dominance

Introduit initialement par (PARDI 1948), il est donné par la formule suivante :

$$X_i = \frac{D_i}{D_i + S_i} \quad (5.1)$$

où  $D_i$  est le nombre d'interactions dans lesquelles l'agent  $i$  a été dominant (a gagné) et  $S_i$  le nombre d'interactions où il a été dominé.

L'indice de dominance  $X_i$  représente donc la fréquence de la dominance de  $i$  sur les autres agents, avec un biais important : le nombre total d'interactions n'intervient pas. Ainsi, il suffit à un agent de gagner son premier combat pour avoir  $X_i = 1$ . Mais ce biais n'interviendra pas dans les simulations, le nombre d'interactions liées à la dominance étant suffisamment grand.

On peut étendre cette notion dans le cas de relations individualisées en posant :

$$X_{i,j} = \frac{D_{i,j}}{D_{i,j} + S_{i,j}} \quad (5.2)$$

où  $D_{i,j}$  (respectivement  $S_{i,j}$ ) représente le nombre de victoires (resp. défaites) de l'agent  $i$  sur l'agent  $j$ . On considère alors que  $i$  domine  $j$  lorsque  $X_{i,j} > X_{j,i}$ .

### Dominance réelle et dominance formelle

La relation de dominance peut être observée de deux façons :

- En mesurant l'indice de dominance  $X_i$  de chaque agent de la population, calculé à partir de *l'issue des combats*, ou mieux encore, en calculant pour tous les agents  $X_{i,j}$ . On peut alors déterminer les relations de dominance d'agent à agent, de façon *objective* (qui gagne effectivement s'il y a combat). Il s'agit de la *dominance réelle*.
- En mesurant un indice de dominance  $Y_i$  (ou mieux,  $Y_{i,j}$ ) qui ne tient compte *que des comportements ritualisés* : signaux de menace et de soumission. Ces signaux servent à réduire l'agression dans la population en rappelant et en confortant les relations déjà établies. Il s'agit de la *dominance formelle*.

### Linéarité de la hiérarchie

On peut alors mesurer deux autres grandeurs : la *linéarité* de chacune de ces relations. La hiérarchie (réelle ou formelle) instaurée par la relation de

dominance peut être plus ou moins transitive (si  $i$  domine  $j$  et  $j$  domine  $k$ , on observe également que  $i$  domine  $k$ ), ce qui peut être calculé au moyen des matrices  $(X_{i,j})$  ou  $(Y_{i,j})$ . On obtient un coefficient entre 0 et 1 (1 représentant une hiérarchie strictement linéaire, 0 l'absence de relation d'ordre).

D'ordinaire, les hiérarchies de dominance sont assez linéaires et permettent alors de définir clairement un *rang* correspondant au classement dans la hiérarchie. Mais il est possible d'observer des relations triangulaires ( $i$  domine  $j$ ,  $j$  domine  $k$ ,  $k$  domine  $i$ ) lorsque les interactions entre individus conduisent à des modifications de leur rang (DE WAAL 1982).

### 5.2.3 Outils informatiques

#### E-EMF

Les comportements que nous avons implémenté l'ont été en langage SMALLTALK, au moyen d'E-EMF (cf. § 4.3). Il s'agit donc de tâches composées d'un chaînage de primitives comportementales.

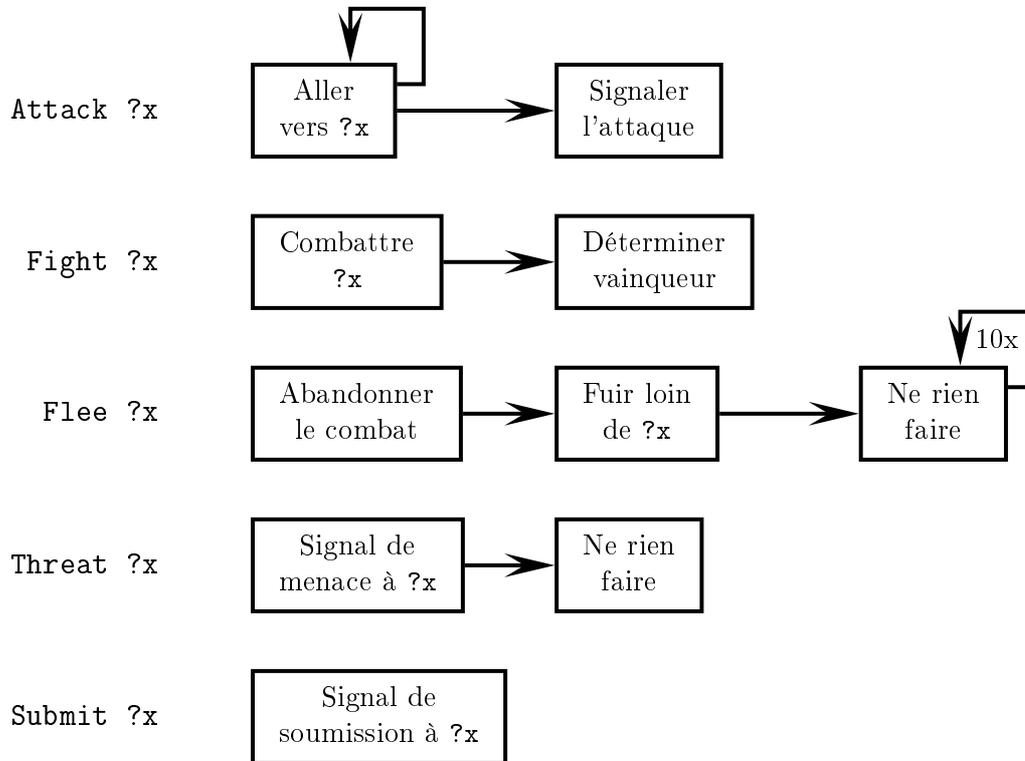
À titre d'illustration, le tableau 5.2 page 74 donne la composition des tâches intervenant pour les comportements relationnels agonistiques d'une classe d'agents utilisée pour les expériences (tâches **Attack**, **Fight**, **Flee** pour les combats et **Threat**, **Submit** pour les signaux de ritualisation). On pourra vérifier qu'elles restent particulièrement simples. Le tableau 5.1 indique les préconditions qui décident de leur éligibilité.

TAB. 5.1 – *Les préconditions d'activation des tâches. Rappelons qu'une tâche n'est activée que si elle est éligible, c'est-à-dire si ses préconditions sont vérifiées, puis en fonction de sa priorité par rapport aux autres tâches éligibles.*

Tâche	Priorité	Préconditions
<b>Attack</b>	8	$[(?x \text{ menace } self) \wedge (self \text{ s'alimente}) \wedge (self \text{ domine } ?x)] \vee [(self \text{ menace } ?x) \wedge (?x \text{ n'a pas répondu par une soumission})]$
<b>Fight ?x</b>	9	$(?x \text{ attaque } self) \vee (?x \text{ combat déjà } self)$
<b>Flee ?x</b>	10	$(\text{combat entre } self \text{ et } ?x) \wedge (X_{self,?x} < X_{?x,self})$
<b>Threat ?x</b>	5	$(\text{interaction affiliative entre } ?x \text{ et } ?y) \wedge (\text{interaction affiliative entre } self \text{ et } ?y) \wedge (self \text{ domine } ?x)$
<b>Submit ?x</b>	6	$(?x \text{ menace } self) \wedge (?x \text{ domine } self)$

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

TAB. 5.2 – Les tâches agonistiques du modèle.



### Le simulateur

La plate-forme de simulation que nous avons utilisée est dérivée de MACACA et reprend un certain nombre de ses outils.

Les comportements sont définis au moyen d'un *Task Browser*, dans lequel on fixe les caractéristiques des tâches (priorité, exclusion avec d'autres tâches, etc.) et leurs préconditions. Celles-ci sont décrites au moyen de mots-clés permettant d'invoquer la perception d'événements dans l'environnement (`MATCH:`), des conditions résultant de l'évaluation d'expressions `SMALLTALK` (`COND:`), et d'envoyer un paramètre à la tâche (`SEND:`).

Ainsi, dans l'exemple qui apparaît sur la figure 5.3 page 75, « `MATCH: $threat ?x self` » recherche si un événement nommé `$threat` est perceptible dans l'environnement. Cet événement comporte deux paramètres, on impose que le second soit `self` (l'agent lui-même). Ici, l'agent correspondant au premier paramètre (celui qui émet le signal de menace) est identifié par la variable `?x`. Si les autres conditions fixées sont vérifiées (notamment si l'agent `self` domine `?x`) la tâche `attack` devient éligible avec pour paramètre `?x` (« `SEND:`

?x »).

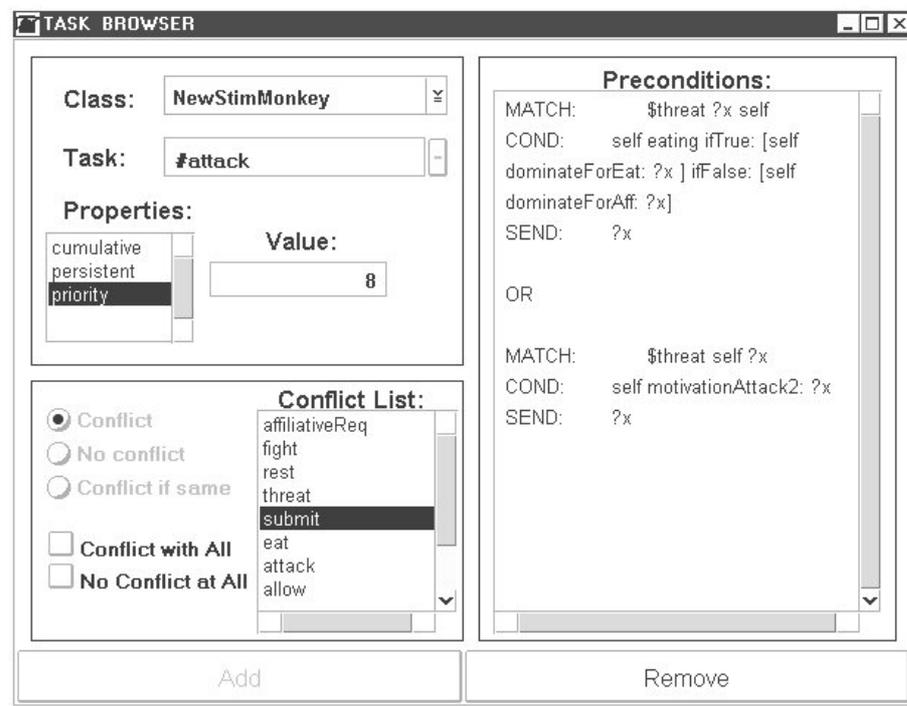


FIG. 5.3 – L’outil « *Task Browser* » qui sert à définir les propriétés et les préconditions des tâches (ici, la tâche *Attack*).

De la plate-forme MACACA, nous avons également adapté un outil d’observation statistique des résultats de la simulation (cf. figure 5.4 page 76). Nous avons notamment rajouté des fonctionnalités permettant de suivre plus finement l’évolution de la hiérarchie réelle et de la hiérarchie formelle simultanément.

◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

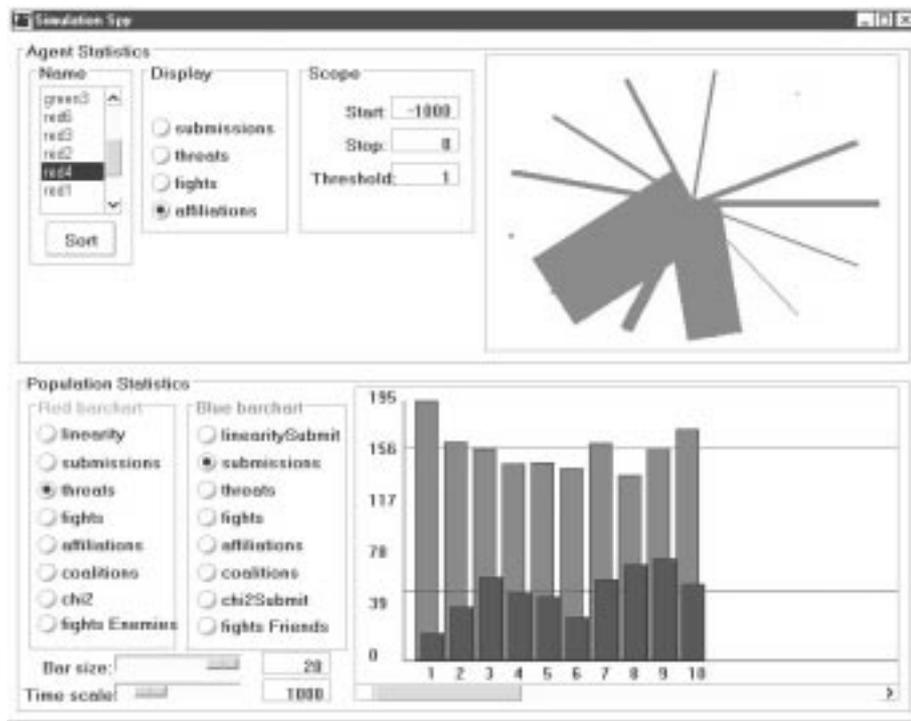


FIG. 5.4 – L’outil d’analyse statistique des simulations. La demi-fenêtre supérieure permet d’observer la fréquence des comportements relationnels entre toutes les paires d’agents (lignes plus ou moins épaisses entre un agent central observé et ses congénères représentés en cercle). En bas, on mesure pour chaque période de temps ( $T$ , à ne pas confondre avec le pas de temps de la simulation) le nombre total de comportements au sein de la population.

## 5.3 Modèle proposé

### 5.3.1 Représentation de la hiérarchie

La mémoire des interactions utilisée dans MACACA est « fermée », elle n'a *aucune conséquence visible sur l'agent* et n'est pas lisible par les autres agents.

Or, chez les primates, les changements hiérarchiques sont en grande partie liée à l'influence que les individus exercent les uns sur les autres. Les dominants semblent plus gros (non parce qu'ils sont plus corpulents, mais parce que leurs poils sont fréquemment hérissés), leur poil est mieux lustré, ils occupent une position centrale dans le groupe (cf. travaux de HEMELRIJK), etc.<sup>2</sup> Comment prendre en compte ce facteur dans le modèle de représentation de la dominance?

#### Les phéromones

Les phéromones<sup>3</sup> sont des substances chimiques qui constitueraient l'une des premières formes de communication (MCFARLAND 1990).

Elles servent dans l'ensemble du monde animal à attirer des partenaires, marquer un territoire, lancer des messages d'alerte, identifier l'appartenance d'un individu à un groupe, indiquer son statut social, etc. (phéromones de *signalisation*), ou encore contribuer au développement d'un individu (phéromones d'*amorçage*).

Les phéromones ont plus spécialement été étudiées chez les insectes pour lesquels elles constituent un mode de communication majeur. Chez les guêpes par exemple, le renforcement de la dominance se traduit par une augmentation des sécrétions phéromonales, qui à son tour influe sur l'issue des combats (THERAULAZ 1991). Chez les rongeurs (groupes de rats), ce sont encore des phéromones qui interviennent dans le marquage du statut social.

#### Phéromones et IAD réactive

Les phéromones ont par ailleurs servi de métaphore pour des modèles de comportement réactif en IAD : en simulation d'abord (DROGOUL 1993; BO-

---

2. Des études sur les réactions endocriniennes des primates montrent que les dominants sont par ailleurs *plus stressés* que les dominés.

3. Appelées aussi « phéromones », de *pherein* (transporter) et *hormân* (exciter).

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

NABEAU et THERAULAZ 1994), puis rapidement comme méthode de résolution de problèmes (BONABEAU et al. 1999a).

En tant que médium de communication, les phéromones présentent en effet des caractéristiques intéressantes pour l'IAD réactive :

- elles véhiculent des informations *simples* ;
- leur existence dans le temps est limitée (par leur évaporation) et de ce fait *les informations obsolètes sont éliminées* ;
- leur propagation dans l'espace est limitée (par des phénomènes physiques de diffusion), et par suite *l'intensité* du stimulus perçu est *porteuse d'informations* (elle permet de déterminer la distance entre le récepteur à la source et la direction de celle-ci, puisque la diffusion engendre un gradient).

### Une métaphore : les « phéromones sociales »

Bien entendu, rien jusqu'à présent n'a montré l'utilisation de phéromones comme indicateur du rang social chez les primates. Nous nous garderons bien de prétendre modéliser les *causes effectives* de la représentation sociale.

Toutefois, l'utilisation d'un modèle réactif *pour reproduire une fonction sociale donnée* s'intègre parfaitement au principe de parcimonie et à l'hypothèse de séparation des rôles.

Que doit-on entendre par « phéromones sociales » (PICAULT et COLLI-NOT 1998a; PICAULT 1998)? Par analogie avec les phéromones chimiques, il doit s'agir d'un objet porteur de signification, qui possède :

- une forme spécifique (un identifieur) donnant la nature du message ;
- une intensité ;
- une localisation dans l'espace.

De plus, pour jouer un rôle *relationnel*, nous ajouterons à cette structure *l'identité de l'émetteur*, de sorte que la réaction d'un agent à une phéromone dépendra de son émetteur.

### Réaction aux phéromones sociales

Chez les animaux, la réaction aux phéromones dépend ordinairement d'un seuil de sensibilité. Pour réagir à une phéromone relationnelle dans une population de  $N$  agents  $\{a_i\}_{i \in [1..N]}$ , chaque agent doit disposer d'un ensemble de seuils  $\{\mathcal{S}_{i,j}\}_{j \neq i}$ .

Chaque agent peut en outre percevoir des signaux émis par tous les agents  $a_i$ , d'intensité phéro( $i$ ). On décide alors qu'un agent  $a_i$  réagit au signal émis par un autre agent  $a_j$  lorsque : phéro( $j$ ) >  $\mathcal{S}_{i,j}$ .

### Application : la phéromone *rank*

Pour représenter la relation de dominance, nous allons utiliser une telle phéromone sociale, nommée *rank*.

Chaque agent dispose de seuils de réaction aux phéromones *rank* qu'il perçoit, et émet lui-même une phéromone *rank* d'intensité initiale :

$$\text{rank}(i) = \frac{1}{N-1} \sum_{j \neq i} \mathcal{S}_{i,j}$$

L'agent  $a_i$  estime qu'il domine  $a_j$  lorsque ses perceptions donnent :  $\text{rank}(j) < \mathcal{S}_{i,j}$ .

On peut en donner une illustration sur la figure 5.5. Plus l'ordonnée du point représentant le seuil de réaction à un agent  $a_j$  donné est élevée, moins il est probable que le stimulus émis par celui-ci soit suffisamment fort pour être pris en considération, donc moins il est probable que l'agent  $a_i$  soit dominé par  $a_j$ . Plus les seuils d'un agent sont élevés, plus l'intensité de sa phéromone *rank* est forte : l'agent a alors de fortes chances de se trouver au sommet de la hiérarchie.

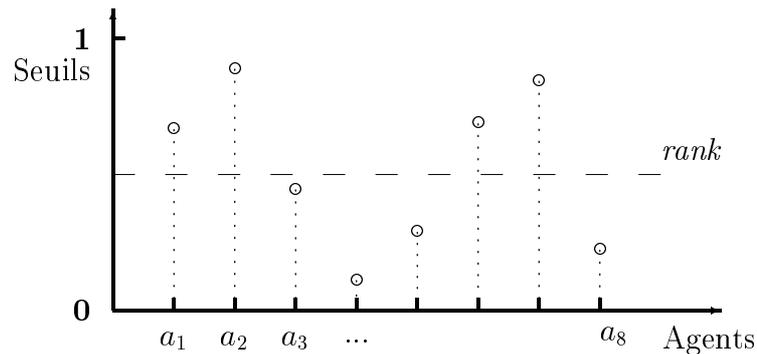


FIG. 5.5 – Représentation de la relation de dominance dans le modèle des « phéromones sociales ». Pour chaque agent  $a_j$ , on indique le seuil de sensibilité ( $\mathcal{S}_{i,j}$ ) de l'agent observé (qu'on notera  $a_i$ ) au stimulus « rank » émis par  $a_j$ . Si  $\text{rank}(a_j) < \mathcal{S}_{i,j}$ , alors  $a_i$  considère qu'il domine  $a_j$ .

### Évolution du rang social

Au fur et à mesure des interactions entre agents, l'issue des combats et des confrontations ritualisées doivent être prises en compte pour l'ajustement des seuils (et, en conséquence, de l'intensité de la phéromone *rank*).

Pour cela, on peut utiliser le même mécanisme que les modèles dérivés de MIRROR, c'est-à-dire qu'après une interaction entre  $a_i$  et  $a_j$ , les seuils sont ajustés de façon à donner plus d'importance à une interaction contraire aux représentations établies qu'à une interaction qui s'y conforme.

Pour l'agent  $a_i$ , cela donne (réciproquement  $a_j$  se livre à pareil calcul) :

$$\begin{cases} \mathcal{S}_{i,j} \leftarrow \mathcal{S}_{i,j} + \Gamma(1 - \mathcal{S}_{i,j}) & \text{si } a_i \text{ est vainqueur} \\ \mathcal{S}_{i,j} \leftarrow \mathcal{S}_{i,j} - \Gamma\mathcal{S}_{i,j} & \text{sinon} \end{cases}$$

Le facteur  $\Gamma$  est un gain paramétrable qui modifie la vitesse de différenciation sociale. Les changements opérés selon l'issue des interactions et les représentations initiales sont illustrés par les figures 5.6 et 5.7.

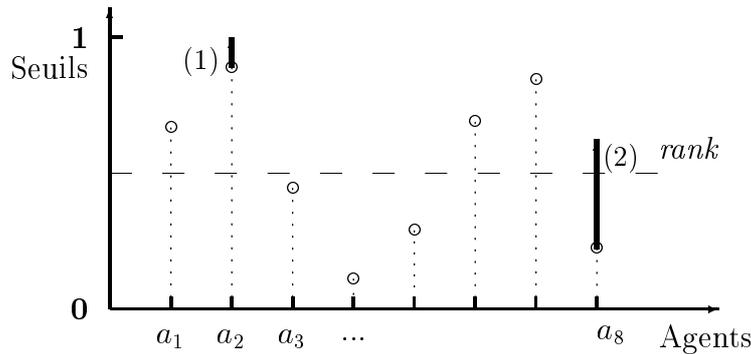


FIG. 5.6 – Évolution des seuils de réaction de  $a_i$  en cas de victoire de l'agent  $a_i$ . Si l'agent vaincu (situation 1, agent  $a_2$ ) est déjà considéré comme dominé, la variation est peu importante. Au contraire, s'il était plutôt perçu comme un dominant (situation 2, agent  $a_8$ ), l'importance de la victoire de  $a_i$  se traduit par une forte augmentation du seuil.

Ces changements dans la valeur des seuils entraînent automatiquement une modification de l'intensité de la phéromone *rank* émise par chacun des protagonistes : ainsi, une interaction sociale entre deux agents produit un résultat perceptible par les autres, et qui modifiera le comportement global de chacun des participants vis-à-vis des autres agents<sup>4</sup>.

4. Des dynamiques voisines ont été observées dans les travaux sur la polygynie chez les

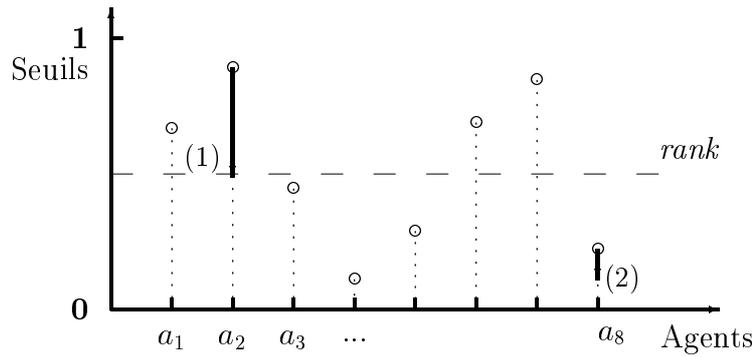


FIG. 5.7 – Évolution des seuils de réaction de  $a_i$  en cas de défaite de l'agent  $a_i$ . Si l'agent vainqueur (situation 1, agent  $a_2$ ) est considéré comme dominé, la variation est importante pour tenir compte de l'événement. Au contraire, s'il était plutôt perçu comme un dominé (situation 2, agent  $a_8$ ), le seuil est peu modifié.

### 5.3.2 Modèle de perception

Il reste à choisir une loi de propagation des « phéromones sociales » dans l'environnement. Dans MACACA, les événements ne sont perceptibles que dans un rayon de perception donné, que nous avons conservé : au-delà, les phéromones ne seront pas prises en compte.

Mais ordinairement, l'intensité d'une phéromone décroît avec la distance à la source selon les lois de la diffusion. Ceci dit, il n'est pas certain que ce point précis soit totalement pertinent dans le cadre de nos rôles relationnels. En l'occurrence, si l'on peut concevoir que l'intensité du signal puisse être modulée, la distance entre les agents (pourvu qu'elle soit inférieure au rayon de perception) ne semble pas un critère particulièrement évident.

En revanche, les phénomènes sociaux chez les primates reposent sur la *manipulation de l'attention*. Nous avons donc opté pour les conventions suivantes :

- les phéromones sont *portées par leur émetteur* ;
- leur intensité *perçue* est fonction de leur intensité à l'émission et de l'orientation des agents les uns par rapport aux autres (deux agents qui sont orientés l'un vers l'autre perçoivent un signal maximum, ceux qui se tournent le dos une intensité très faible).

Les comportements agonistiques décrits précédemment ont été bien sûr modifiés en conséquence (en rajoutant une primitive réglant l'orientation des

---

fourmis (HUTZLER 1995).

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

agents selon les interactions).

### 5.4 Expériences

#### Protocole

Les simulations réalisées avec ce modèle, pour la représentation de la relation de dominance, ont été comparées aux résultats obtenus avec les classes d'agents équivalentes dans MACACA. Il s'agit en effet d'être capable de souligner ce que ce modèle apporte par rapport à une simple mémoire des interactions.

Les populations testées comportent de 5 à 30 individus, placés dans un environnement similaire à celui de MACACA, à savoir un terrain rectangulaire possédant une zone de repos et une zone d'alimentation. Diverses valeurs du paramètre  $\Gamma$  ont été utilisées (sans changement qualitatif notable).

Les simulations ont été effectuées à chaque fois sur des périodes de l'ordre de 40 000 pas de temps (comme mentionné § 4.3, l'exécution d'une primitive comportementale appartenant à une tâche occupe un pas de temps).

#### Résultats de référence

Les simulations menées avec une mémoire des interactions comme dans MACACA donnent dans tous les cas des résultats semblables :

- La linéarité de la dominance réelle comme de la dominance formelle augmentent exponentiellement jusqu'à 1 (cf. figure 5.8).
- Parallèlement, le nombre de combats et de menaces diminue, tandis que celui des soumissions et des interactions affiliatives augmente.

#### Résultats et interprétation

Lorsqu'on observe ce qui se passe avec les « phéromones sociales », on constate au contraire des fluctuations assez importantes de la linéarité, tant réelle que formelle (cf. figures 5.9, et 5.11).

On peut montrer, en examinant les autres comportements, qu'il ne s'agit pas de comportements aberrants mais de modifications de la hiérarchie.

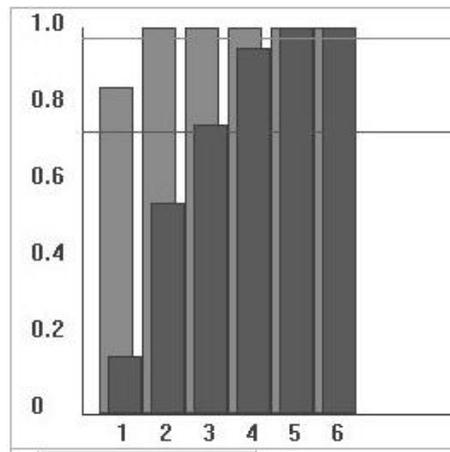


FIG. 5.8 – *Linéarité de la dominance réelle (gris foncé) et de la dominance formelle (gris clair) dans MACACA. Ces valeurs sont comptabilisées par périodes de 4000 pas de temps (en ordonnée : linéarité ; en abscisse : période).*

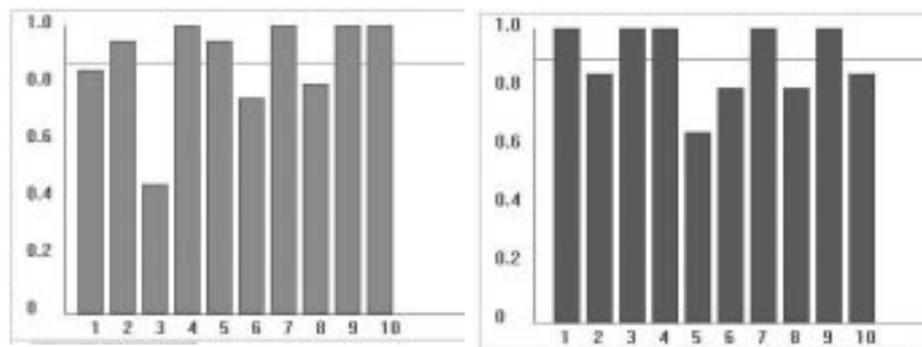


FIG. 5.9 – *Simulation A. Évolution de la linéarité de la dominance réelle (à gauche) et de la dominance formelle (à droite) dans le modèle de « phéromone sociale ».*

En voici deux exemples représentatifs des expériences réalisées (simulations A et B, figures 5.9 et 5.10 pour A et 5.11 et 5.12 pour B). Leur comparaison montre les phénomènes suivants quant aux fortes diminutions de la linéarité de la hiérarchie réelle (périodes 3, 6, 8 de la simulation A et 4, 7, 11 de la simulation B) :

- 1° Elles sont précédées d'une forte baisse de la linéarité de la hiérarchie formelle.
- 2° Elles s'accompagnent d'une augmentation notable du nombre de menaces et d'une diminution légère du nombre d'interactions affiliatives.
- 3° Elles sont suivies d'une augmentation du nombre de combats et d'une

◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

forte diminution du nombre de soumissions.

De plus, les seuils de réaction aux phéromones *rank* ne se « bloquent » pas aux extrêmes et varient de façon conséquente lors des perturbations observées.

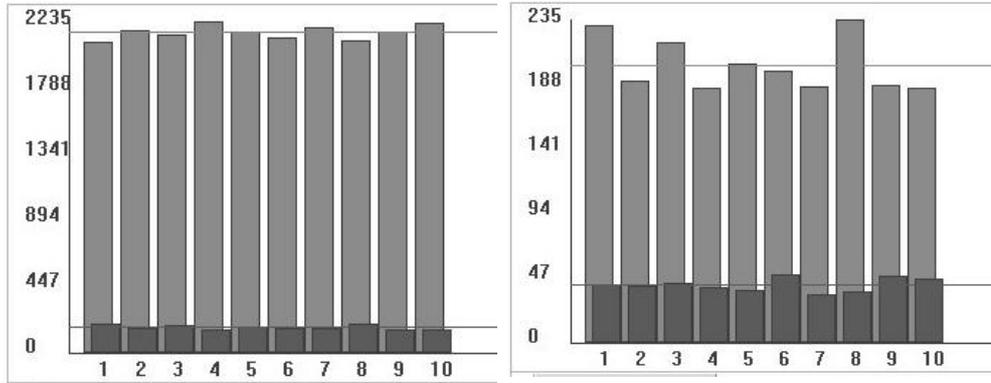


FIG. 5.10 – *Simulation A. Figure de gauche: évolution du nombre d'interactions affiliatives (gris clair) et de combats (gris foncé); figure de droite: évolution du nombre de menaces (gris clair) et de soumissions (gris foncé).*

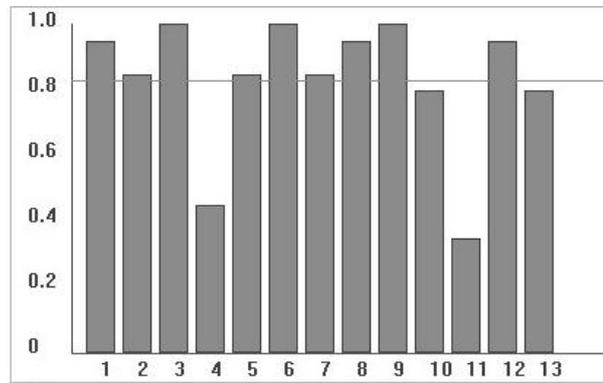


FIG. 5.11 – *Simulation B. Évolution de la linéarité réelle dans le modèle de « phéromone sociale ».*

L'analyse détaillée des comportements permet de proposer quelques explications aux phénomènes observés.

Deux agents *a* et *b* de force comparable (dominant à peu près autant de congénères l'un que l'autre) ont de fortes chances d'émettre des phéromones *rank* d'intensités très voisines. Cela ne signifie pas pour autant qu'ils ont le même *rank* au sens de la relation de dominance: l'un des deux est dominant sur l'autre (disons par exemple que *a* domine *b*). Mais la différence entre  $|\text{rank}(b) - \mathcal{S}_{a,b}|$  et  $|\text{rank}(a) - \mathcal{S}_{b,a}|$  peut être faible.

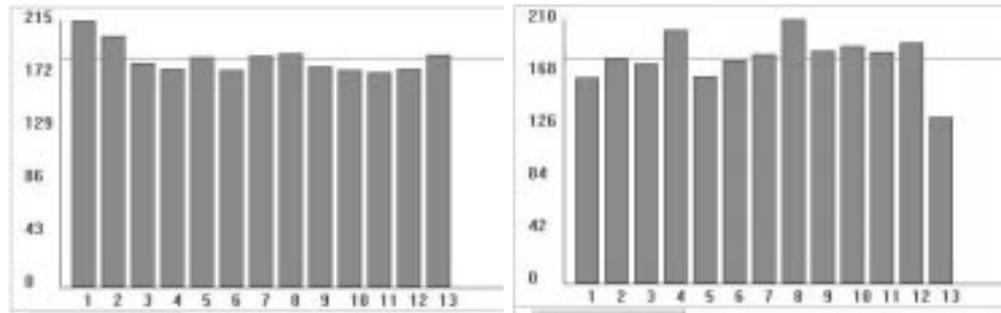


FIG. 5.12 – *Simulation B. Évolution du nombre de combats (à gauche) et de soumissions (à droite) dans le modèle de « phéromone sociale ».*

Supposons que  $a$  et  $b$  ne se rencontrent pas pendant quelque temps. Chacun, selon les agents qu'il rencontre, ajuste ses représentations sociales (c'est-à-dire ses seuils de réaction). Il est possible que le statut de  $b$  s'améliore quelque peu. Dans ce cas,  $\text{rank}(b)$  s'est élevé également. Certains de ses congénères peuvent se mettre à lui adresser des signaux de soumission : la hiérarchie formelle s'en trouve donc bousculée.

Mais pour autant, le rang de  $b$  dans la hiérarchie réelle n'a pas encore changé. Lorsque  $b$  et  $a$  se rencontrent à nouveau, ils n'ont pas eu forcément l'occasion d'ajuster leurs seuils de réponse de l'un à l'autre et on est donc confronté à un phénomène d'*hystérésis*. Tout se passe comme si les agents s'évaluaient mutuellement sur la base de *préjugés* (au sens de « représentations sociales subjectives et *a priori* » qui ne reflètent pas nécessairement une situation objective).

Si un combat entre  $a$  et  $b$  s'ensuit,  $b$  a plus de chances de l'emporter qu'auparavant (puisque  $\text{rank}(b)$  a augmenté). Il devient alors dominant sur  $a$  et  $\text{rank}(b)$  s'accroît d'autant. Il s'agit cette fois d'une modification de la hiérarchie réelle qui va se répercuter pendant quelque temps auprès de ses congénères. Le vaincu doit également trouver une nouvelle position dans la hiérarchie. La structure sociale se stabilise progressivement, jusqu'à un nouveau conflit.

Les variations aléatoires qui existent toujours dans la population se trouvent en fait amplifiées par les mécanismes de rétroaction positive et d'influence entre agents (par les « phéromones sociales ») et conduisent à leur différenciation. Ce phénomène d'« ordre par fluctuations » est fréquent dans les systèmes auto-organisés (ATLAN 1979).

### Des éléments de validation

Il reste à savoir toutefois si les comportements observés ont une quelconque vraisemblance éthologique. Or, les résultats empiriques, en particulier le lien entre les variations des relations formelle et réelle coïncident avec ceux qu'on peut décrire par exemple chez les chimpanzés :

*« Généralement, rangs formels et rangs réels se recourent. En certaines circonstances, cependant, le rang peut devenir fictif. Autrement dit, la position d'un individu dominant peut devenir intenable. [...] Si, par exemple, un subordonné est de plus en plus souvent vainqueur dans les conflits, ou si du moins il provoque régulièrement de la peur et des hésitations chez la partie adverse, ce résultat ne lui échappera pas. Si ce décalage dans la relation réelle persiste, peu à peu les « saluts » ne deviendront plus qu'une formalité creuse, puis le dominé cessera de « saluer » le dominant. De ce fait, il remettra en cause, pour ainsi dire, le statut de leur relation. Ce premier pas – arrêter de « saluer » – a été observé lors de tous les renversements de dominance survenus dans la colonie d'Arnhem » (DE WAAL 1982)*

### Autres caractéristiques du modèle

L'utilisation d'une mémoire des interactions dans MACACA présentait un autre inconvénient : l'initialisation de cette mémoire, pour tout nouvel agent introduit dans le système, imposait à celui-ci et aux autres de s'affronter tous explicitement pour établir son rang dans la population.

Le modèle de « phéromones sociales » apporte une réponse partielle à ce problème. En effet, bien qu'il soit également inévitable de devoir fixer arbitrairement la valeur initiale des seuils (0,5 en l'occurrence), ils *n'empêchent pas* de se comporter de façon adéquate face à des agents émettant une phéromone *rank* d'intensité soit plus forte, soit plus faible.

L'utilisation de ce modèle permet en outre de rendre compte de quelques phénomènes décrits dans le chapitre 3.

La phéromone *rank*, en ce qu'elle influence le comportement des autres agents et l'issue des combats, peut représenter une forme de *confiance en soi*. On peut s'en rendre compte à travers la notion de *leurre* (cf. § 3.2.5).

Si par exemple on fixe artificiellement l'intensité de la phéromone *rank* d'un agent à 1, en l'empêchant de varier durant plusieurs centaines de pas de

temps, il gagne d'abord un rang formel élevé, ce qui finit par se traduire par un rang réel élevé. Ce rang reste stable assez longtemps après la fin de son blocage forcé.

## 5.5 Conclusion sur les primates

### 5.5.1 Les résultats du modèle et son extension

#### Bilan des résultats de la modélisation

Les expériences menées sur le modèle de représentation de la relation de dominance par des « phéromones sociales » ont permis de rendre compte correctement des renversements de hiérarchie qu'on peut observer chez les primates. Il ne s'agit pas de phénomènes purement aléatoires mais de mécanismes faisant appel à une forme de « subjectivité » dans les représentations utilisées par les agents.

Il est intéressant de noter que cette seule représentation de la dominance *suffit à expliquer* des renversements de hiérarchie, alors que F. DE WAAL considère que ceux-ci sont nécessairement l'effet des coalitions (DE WAAL 1982). Les alliances entre individus jouent certainement un rôle dans la stabilisation et dans le renforcement de la hiérarchie, mais il est possible qu'existe déjà, au niveau des relations entre paires d'agents, une dérive spontanée de la hiérarchie.

Il nous faut maintenant étudier comment la fonction relationnelle de la hiérarchie peut être reliée à des fonctions individuelles ou organisationnelles (buts particuliers ou collectifs des agents).

#### Modélisation d'autres rôles relationnels

Un modèle analogue peut être utilisé pour étudier la relation d'affiliation entre agents. Toutefois, les résultats obtenus sont peu modifiés par le remplacement de la mémoire des interactions par un modèle de « phéromone affiliative », et ce quel que soit le modèle utilisé pour la dominance.

Ceci dit, l'affiliation est un comportement symétrique entre deux individus. Plutôt que de produire une structure orientée (les rangs hiérarchiques), les interactions affiliatives conduisent à des renforcements symétriques entre paires d'agents : la question n'est pas de déterminer si  $a$  domine  $b$  ou  $b$  domine  $a$  (*sens* de la relation) mais si  $a$  et  $b$  sont affiliés ou ne le sont pas (*existence* de

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

la relation). Il n'y a donc pas de raison que les dynamiques des deux relations soient comparables.

Cela signifie du reste, selon le principe de parcimonie, qu'il nous faut garder un modèle de type « mémoire affiliative » plutôt que d'utiliser des phéromones sociales dans ce cas.

### Vers des rôles organisationnels?

La notion de « phéromones sociales » transpose aux rôles relationnels des mécanismes réactifs observables pour des comportements individuels dans les sociétés d'insectes.

On peut légitimement penser à étendre ce type de modèle aux rôles organisationnels (constitution de groupes, coordination entre plusieurs agents). Cela peut présenter un intérêt comme métaphore dans diverses applications, notamment en introduisant un « environnement social », c'est-à-dire un espace dans lequel se propagent les « phéromones sociales », qu'elles soient relationnelles ou organisationnelles. Ces concepts font actuellement l'objet des recherches de la thèse de François SEMPÉ (LIP6/France Télécom R&D).

En revanche, son utilisation pour la *simulation* des mécanismes de coalition chez les primates n'est peut-être pas adéquat.

Le critère habituellement invoqué comme explication à la stabilité des coalitions est la *réciprocité* (notamment sous l'influence de la théorie des jeux). Mais des études récentes montrent que la réciprocité est un critère trop restrictif, qui ne permet pas l'instauration d'alliances stables dans la population (DESSALLES 1999). Il serait nécessaire pour cela de faire appel à la *théorie du prestige* (ZAHAVI et ZAHAVI 1997). Une étude plus fine des conditions de formation des coalitions est donc un préliminaire obligé à la simulation.

### 5.5.2 Les hypothèses formulées et la méthodologie

#### Principe de parcimonie

Le principe de parcimonie nous a conduit à proposer un modèle *réactif* opérant sur des fonctions relationnelles, de préférence à un modèle qui ferait appel à des connaissances symboliques explicites ou à des mécanismes encore plus sophistiqués.

Ce modèle réactif s'est avéré suffisant pour rendre compte de l'évolution de la relation de dominance et de ses caractéristiques (opposition entre dominance réelle et dominance formelle).

### Hypothèse de séparation des rôles

Elle a été particulièrement fructueuse ici, dans la mesure où elle a permis de reproduire une *fonction* plutôt complexe indépendamment des causes efficientes réellement à l'origine de cette fonction chez les primates.

Le modèle proposé est *empiriquement équivalent* aux mécanismes responsables de la représentation de la dominance. Or c'est précisément cette fonction qui nous intéresse en vue d'une transposition à d'autres applications multi-agents.

### Un modèle plus simple?

Récemment, le modèle classique de différenciation sociale sous l'effet d'une rétroaction positive renforçant des fluctuations aléatoires, issu du modèle MIRROR (HOGEWEG et HESPER 1983), a été remis en question. En effet, il ne peut pas être distingué, sur la base des données empiriques actuelles, d'un modèle dans lequel préexistent des différences individuelles (BONABEAU et al. 1999b) – modèle qui présente de plus l'avantage d'être *plus simple* que celui de la différenciation auto-organisée.

Toutefois, le modèle de différenciation possède un atout considérable par rapport à son concurrent : il *explique*, en donnant des mécanismes concrets, comment une structure peut se former à partir d'un groupe indifférencié. Dans ce sens il est d'un plus grand intérêt pour fournir des techniques d'organisation décentralisée à l'IAD.

À l'opposé, le modèle de différences préexistantes est plus simple. Mais si l'on reprend les schémas explicatifs de l'éthologie, ces différences doivent s'interpréter comme une *nécessité évolutive*. Le modèle n'est donc explicatif qu'au travers d'une interprétation en termes de gains adaptatifs. Or cette fonction adaptative des différences individuelles n'est pas encore clairement comprise. Pour l'instant, ce modèle n'a donc de valeur explicative que tautologique (« les individus sont différents parce qu'ils *étaient* déjà différents »).

Cela suggère de rajouter au principe de parcimonie un garde-fou méthodologique non négligeable : entre deux modèles, il ne s'agit pas de choisir

## ◇5. MODÈLES DE REPRÉSENTATIONS SOCIALES

systematiquement *le plus simple*, il faut tenir compte également de l'appartenance des modèles *au même schéma explicatif éthologique*.

Ainsi, entre deux modèles représentant une hypothèse concernant les causes effectives d'un comportement, il faudra choisir le plus simple ; de même entre deux modèles mettant en concurrence des hypothèses sur la fonction, ou sur l'histoire évolutive du comportement. Mais des modèles s'appuyant sur deux schémas différents ne doivent pas être comparés de façon simpliste.

Il nous reste maintenant, pour poursuivre le « cercle vertueux », à étudier les dynamiques des comportements hiérarchiques puis leur transposition à des SMA ouverts ; nous pourrons ainsi évaluer plus finement l'apport fonctionnel de la hiérarchie à la réalisation de tâches collectives.

---

# Chapitre sixième.

## La transposition des modèles

---

*Si l'éthologie fournit à l'IAD des métaphores de comportements sociaux intéressants et si des modèles de représentation sociale permettent de rendre compte de phénomènes observés chez des primates, se pose ensuite le problème de la transposition de ces modèles à des applications informatiques ou robotiques. Cela n'est pas sans soulever quelques problèmes méthodologiques que nous allons évoquer ici. Nous étudierons ensuite un cas concret de transposition : la construction d'un lexique commun par une population d'agents, dans un contexte social.*

### 6.1 Comprendre les dynamiques sociales

#### 6.1.1 L'étape suivante du « cercle vertueux »

Les expériences menées dans le cadre des compétences sociales des primates nous ont permis de définir des modèles qui permettent de rendre compte d'un certain nombre de capacités de représentation sociale, en particulier de la reconnaissance des rapports hiérarchiques existant au sein du groupe.

Nous avons également souligné le fait que cette hiérarchie, loin d'être figée, se déforme et se recompose régulièrement au gré des interactions entre agents.

L'étape suivante dans le paradigme méthodologique du « cercle vertueux » consiste donc à *transposer* la fonction collective ainsi produite, à savoir la *construction d'une hiérarchie adaptative*, à d'autres applications multi-agents.

Les hypothèses de la méthodologie CASSIOPÉE, et au premier chef l'hypothèse de séparation des rôles, ont rendu possible l'utilisation, pour le niveau relationnel, de modèles réactifs. Mais cette hypothèse, souvent vérifiée pour

## ◇6. LA TRANSPOSITION DES MODÈLES

la transposition de comportements individuels, est-t-elle tenable pour des modèles concernant le niveau des rôles relationnels, puisque ceux-ci reposent sur des dépendances fonctionnelles entre les rôles individuels ?

Nous pensons que cela reste parfaitement possible moyennant de trouver des *équivalences fonctionnelles* entre les rôles individuels de la situation de départ et ceux de l'application visée. Puisque les rôles relationnels sont censés se fonder sur des dépendances entre rôles individuels, une même relation de dépendance doit permettre l'utilisation d'un même rôle relationnel.

Ainsi, par exemple, si nous sommes en mesure de déterminer des comportements équivalents aux actions agonistiques des primates, il devrait être possible de transposer les modèles de représentation de la relation de dominance assez facilement.

Toutefois, si le modèle lui-même est transposable, en dépit d'un changement de la nature des rôles individuels, il n'est pas certain que les rôles relationnels conduisent à la même organisation (ils peuvent en effet entrer en interaction avec d'autres rôles relationnels, qui n'existaient pas dans la simulation initiale).

La transposition n'exclut donc pas une certaine « méfiance » et le maintien d'une méthodologie incrémentale : prétendre transposer un modèle tel quel, sous prétexte qu'il donne de bons résultats pour une situation donnée, est quelque peu naïf. L'idéal est donc de considérer la transposition comme un point de départ susceptible d'être simplifié ou complexifié selon les besoins.

Par ailleurs, il est important de comprendre les dynamiques propres au modèle qu'on cherche à transposer. En travaillant sur des applications simples (voire des problèmes-jouets), il devient souvent possible de mettre en lumière un certain nombre de mécanismes qui révèlent par exemple des conditions d'applicabilité jusqu'alors passées inaperçues.

Dans le cadre des modèles de comportements sociaux, cela revient à étudier d'une part quels rôles individuels peuvent être pertinents pour construire la structure sociale, et d'autre part comment celle-ci influe sur les activités collectives.

### 6.1.2 Approches voisines

Un certain nombre de recherches ont été menées pour comprendre comment, dans un contexte relativement abstrait, des dynamiques sociales peuvent s'instaurer et se modifier en fonction d'interactions entre individus.

### L'approche normative de la socialité

On peut notamment mentionner les travaux de C. CASTELFRANCHI et R. CONTE qui ont développé et formalisé un certain nombre de concepts fructueux issus de la psychologie et de la sociologie.

Ils ont abordé en particulier les questions de :

- la *sociabilité* : quelles sont les motivations qui poussent un agent autonome à s'engager dans une interaction avec d'autres agents ? La réponse proposée (CASTELFRANCHI et al. 1992) est qu'il existe un ensemble de *dépendances objectives entre les agents*, tant pour la réalisation de buts individuels que pour l'accès aux ressources.
- la *socialisation* : comment un agent transforme-t-il ses intérêts propres (buts individuels) en problèmes sociaux, en les faisant adopter par d'autres agents ? CASTELFRANCHI propose pour étudier ces phénomènes les concepts de *pouvoir* et d'*influence* (CASTELFRANCHI 1990).

Toutefois, ces études ont été menées en vue de concevoir des systèmes destinés à opérer dans des environnements régis par des « harmonies préétablies » (CASTELFRANCHI 1990). Dans une telle situation, les relations sociales peuvent fonctionner de façon normative à condition que *tous les agents* respectent les mêmes règles de comportement et de représentation. De plus, il s'agit d'un monde clos : il est nécessaire de disposer de connaissances *a priori* sur les compétences des agents, sur leurs interdépendances et sur les événements possibles dans l'environnement.

### Les approches constructivistes de la socialité

À l'opposé, la socialité peut être abordée d'un point de vue plus proche de la Vie Artificielle, en construisant un *monde artificiel* dans lequel on peut manipuler à volonté les comportements des agents. Il est ainsi possible, au moyen d'un tel « laboratoire virtuel », d'étudier des dynamiques sociales dans les populations d'agents en fixant arbitrairement les contraintes auxquelles elles sont soumises.

C'est notamment le cas de SUGARSCAPE, un environnement de simulation dans lequel des agents sont dotés de comportements simples au moyen desquels ils doivent survivre (EPSTEIN et AXTELL 1996). Dans SUGARSCAPE, le monde comporte deux zones circulaires riches en ressources alimentaires, lesquelles croissent à une vitesse fixe. EPSTEIN et AXTELL mènent une étude

incrémentale dans laquelle ils analysent l'influence différentielle de tel ou tel changement de comportement sur la survie des agents, sur la quantité de ressources alimentaires qu'ils amassent, sur des mouvements de groupe émergents, etc.

Ce type d'approche est tout à fait adapté à l'exploration des caractéristiques d'un modèle relationnel. Les comportements individuels peuvent être choisis arbitrairement, en vue d'étudier leur influence sur la dynamique sociale.

C'est à travers une approche voisine que nous allons explorer l'interaction entre des dynamiques sociales et la constitution d'un lexique commun par une population d'agents.

## 6.2 Langage et socialité

### 6.2.1 Point de départ

Des expériences que nous allons décrire dans ce chapitre, les premières doivent beaucoup à une collaboration avec Frédéric KAPLAN en liaison avec ses propres recherches sur la constitution de langues artificielles (KAPLAN 2001).

F. KAPLAN a abordé dans sa thèse (KAPLAN 2000b) l'étude de la construction d'un lexique commun par une population d'agents autonomes. Il a suivi pour cela une approche incrémentale, partant de l'observation de dynamiques de consensus pour aller jusqu'à une mise en situation des agents dans des corps robotiques.

#### Les « jeux de langage »

Dans ces travaux, les agents ont pour seule activité la construction d'un lexique conventionnel, c'est-à-dire un ensemble d'associations entre des signifiants (les mots) et des signifiés (des objets, des formes, etc.). Pour cela, les agents jouent à des « jeux de langage » qui se déroulent selon le protocole immuable suivant<sup>1</sup> :

- 1° Deux agents sont choisis aléatoirement dans la population ; l'un est désigné comme *locuteur*, l'autre comme *interlocuteur*.

---

1. En réalité, dans les expériences les plus complexes notamment avec des robots, il faut rajouter une phase de perception et de *contruction* des signifiés à partir de ces perceptions. Ces aspects, au demeurant extrêmement intéressants, ne seront pas abordés ici.

- 2° Le locuteur choisit un sujet (ou *topic*)  $m$  parmi  $M$  signifiés (sens ou *meanings*) possibles.
- 3° Il choisit un mot  $w$  parmi  $W$  signifiants possibles.
- 4° Il « prononce » le mot  $w$ , c'est-à-dire qu'il le transmet à l'interlocuteur (phase de communication linguistique).
- 5° L'interlocuteur recherche le sens  $m'$  associée *selon lui* au mot  $w$ .
- 6° Il l'indique au locuteur (phase de communication extra-linguistique : communication directe sans bruit ou pointage de l'objet concerné).
- 7° Le locuteur compare  $m$  et  $m'$ . Il y a succès de la communication si et seulement si  $m = m'$ . Il indique le résultat de l'interaction à l'interlocuteur.
- 8° Locuteur et interlocuteur révisent leur lexique en conséquence.

Chaque agent dispose d'une table d'associations entre signifiants et signifiés qui lui permet de choisir aussi bien le « sens » d'un mot que le mot exprimant un sujet donné.

Pareilles expériences, dans lesquelles les agents n'ont d'autre activité que celle dont on entend explorer la dynamique, rappellent la démarche fructueuse de C. HEMELRIJK (§ 4.2). Une différence majeure toutefois les en sépare : il n'existe pas de données empiriques de référence auxquelles comparer les résultats. Ceux-ci doivent être examinés pour eux-mêmes et comparativement les uns aux autres (à l'intérieur d'un même modèle, selon les variations des paramètres, et entre plusieurs modèles). En cela, la méthode d'analyse des résultats est plus proche de celle utilisée dans SUGARSCAPE (EPSTEIN et AXTELL 1996).

L'ensemble des paramètres du modèle (nombre d'agents  $N$ , nombre de sens ou *meanings*  $M$ , nombre de mots  $W$ , règles de choix du mot  $w$  selon le sujet  $m$ , du sens  $m'$  selon le mot  $w$ , et mode de révision du lexique) est complètement explicite et ajustable à volonté.

### Problèmes soulevés

La question que pose F. KAPLAN est de savoir si les agents parviennent à établir, au moyen de ce protocole, un *lexique commun*, c'est-à-dire que la table d'associations entre signifiants et signifiés de chaque agent leur permet de communiquer efficacement, autrement dit de réussir n'importe quel jeu de langage.

## ◇6. LA TRANSPOSITION DES MODÈLES

Comme il apparaît que cela est très souvent possible, un certain nombre de données expérimentales peuvent être utilisées en vue de comparer les modèles les uns aux autres :

- le temps de convergence moyen pour un jeu de paramètres ;
- le nombre de situations où la convergence n'est pas atteinte au bout d'un temps limite (cas de « non-convergence ») ;
- les lois empiriques qui permettent d'estimer le temps moyen de convergence en fonction des paramètres choisis.

D'autres aspects, comme la résistance au bruit, la capacité de discrimination des mots puis des sujets, le renouvellement des agents dans la population, ont été également étudiés (KAPLAN 2000b, pour plus de détails sur ces résultats).

### **Lien avec la socialité**

L'une des problématiques qui nous a poussé à nous pencher sur ces travaux, est la question des aspects sociaux de la communication, et réciproquement du rôle de la communication dans la socialité.

La communication, de façon générale, joue un rôle permanent et central dans les sociétés animales, et tout particulièrement dans celles qui possèdent une organisation sociale sophistiquée. Elle peut, le cas échéant, n'être qu'indirecte, c'est-à-dire ne s'appuyer que sur des modifications de l'environnement : on a alors affaire au phénomène de « stigmergie » décrit par P.-P. GRASSÉ pour expliquer la construction du nid chez les termites (GRASSÉ 1959). Plus fréquemment, la communication chimique, par phéromones, permet l'instauration du rang chez les guêpes (THERAULAZ 1991). C'est bien entendu encore la communication, visuelle ou auditive cette fois, qui donne naissance aux manifestations ritualisées de la relation de dominance (parades, mimiques faciales, cris, grognements).

Chez l'homme, le langage est lui-même intimement lié à la socialité, dans la mesure où il permet d'en exprimer explicitement les règles et de formuler les liens qui existent entre individus. Des théories récentes supposent de plus, dans la lignée des hypothèses sur l'origine sociale de l'intelligence chez les primates (cf. § 3.2.7), que le langage humain a tout d'abord joué un rôle fondamentalement social et non pas simplement « utilitaire » (c'est-à-dire informationnel). Il aurait pu ainsi remplacer les séances de toilettage servant à renforcer les relations entre individus (DUNBAR 1996), ou encore assurer un statut social

éminent aux individus capables de *pertinence* dans leur usage de la langue (DESSALLES 2000).

À l'exception de travaux sur le niveau pragmatique de la langue, dans lequel son usage social est explicitement étudié (DESSALLES 2000), il peut donc sembler paradoxal que les recherches informatiques portant sur les origines ou la construction d'un langage fassent complètement abstraction de l'organisation au sein du groupe.

En ce qui concerne la problématique plus modeste de la construction d'un lexique commun par une population d'agents, la socialité soulève deux problèmes distincts :

1° Quelle est l'influence d'une hiérarchie supposée préexistante sur la construction du lexique ? Du point de vue de la séparation entre rôles relationnels et rôles individuels, l'étude de cette question permettrait d'évaluer l'intérêt *fonctionnel* d'une hiérarchie par rapport à une tâche collective donnée.

2° Réciproquement, est-il possible de se servir de la performance des agents vis-à-vis de la tâche collective (ici, la communication) pour *construire* la hiérarchie ? À l'étude des dynamiques qui apparaissent lorsque l'on introduit une rétroaction du langage sur la socialité, s'ajoute un point méthodologique : la recherche de comportements analogues aux comportements agonistiques des primates. En effet, les modèles sociaux étudiés précédemment s'appuient sur une typologie de comportements liés à l'agression dont il faudra trouver des équivalents pour pouvoir transposer ces modèles.

### 6.2.2 Influence de la socialité sur le lexique

Dans un premier temps, nous allons considérer une population d'agents déjà structurée socialement par un ordre hiérarchique préexistant. Nous pourrions ainsi déterminer comment une relation de dominance affecte l'efficacité d'une tâche collective.

### 6.2.3 Modèle de départ

Nous repartons ici du modèle décrit dans (KAPLAN 2000b, § 3.2, modèle 2.6). Initialement implémenté en LISP, il a été réécrit en SMALLTALK pour la réalisation de l'ensemble des expériences présentées dans ce chapitre.

Les  $N$  agents doivent nommer  $M$  sens. Ils disposent pour cela de  $W$  mots possibles (on aura ici systématiquement  $W = 5M$ ).

## ◇6. LA TRANSPOSITION DES MODÈLES

### Mémoire des agents

Pour nommer les sens et comprendre les mots, chaque agent utilise une mémoire constituée d'associations (*sens*, *mot*) pourvues d'un *score*.

$$\mathcal{M} = (a_{mw})_{m \in [1..M], w \in [1..W]}$$

En l'occurrence,  $a_{mw}$  représente le score de l'association entre le sens  $m$  et le mot  $w$ .

Pour coder un sens  $m$ , l'agent utilise le mot  $w$  défini par :

$$w = \arg \max_{k \in [1..W]} \{a_{mk}\}$$

Réciproquement, le sens  $m$  auquel sera associé un mot  $w$  est donné par :

$$m = \arg \max_{l \in [1..M]} \{a_{lw}\}$$

En pratique, lorsque plusieurs associations ont le même score, un choix aléatoire les départage.

Il est alors possible de calculer la probabilité, au sein de la population, qu'un sens  $m$  soit codé par un mot  $w$  – « matrice de codage » ( $\mathcal{C}_{mw}$ ) – et réciproquement – « matrice de décodage » ( $\mathcal{D}_{wm}$ ). Par suite, on peut évaluer la cohérence en production (le fait que l'ensemble des agents n'utilisera qu'un mot pour chaque sens) et la cohérence en interprétation (le fait qu'un mot donné sera décodé de la même façon par tous les agents).

F. KAPLAN définit enfin le *succès théorique en communication*, qui mesure la probabilité de succès d'un jeu de langage quelconque au vu des associations utilisées par les agents de l'ensemble de la population. On a :

$$ST = \frac{1}{M} \sum_{m \in [1..M]} \sum_{w \in [1..W]} \mathcal{C}_{mw} \cdot \mathcal{D}_{wm}$$

### Modalités du jeu de langage

En cas de succès d'un jeu de langage, le score de l'association ( $m$ ,  $w$ ) est incrémenté d'une unité chez le locuteur comme chez l'interlocuteur. Il est au contraire décrétementé d'une unité chez tous deux en cas d'échec.

Le jeu de langage utilisé ici fait par ailleurs appel à deux mécanismes

supplémentaires (retenus pour leur efficacité dans la convergence du lexique) :

- L’inhibition latérale : en cas de succès, les associations de la même ligne et de la même colonne (dans la mémoire  $\mathcal{M}$ ) sont décrémentées d’une unité, ce qui revient à renforcer l’association  $(m, w)$  *au détriment* des associations qui sont immédiatement en compétition avec elles.
- L’indication du sujet : en cas d’échec (l’interlocuteur ayant désigné un sujet  $m' \neq m$  pour le mot  $w$ ), le locuteur désigne de façon extra-linguistique le sujet du jeu ( $m$ ) et l’interlocuteur incrémente alors d’une unité l’association  $(m, w)$ .

### Rappel des résultats préliminaires

En l’absence de socialité (dans un jeu de langage classique), on observe une relation simple entre temps de convergence et taille de la population ou nombre de sens (cf. figures 6.1 et 6.2) :

$$\begin{cases} \ln T = \mu(N) \cdot \ln M + \varepsilon & \text{à } N \text{ fixé} \\ \ln T = \nu(M) \cdot \ln N + \varepsilon & \text{à } M \text{ fixé} \end{cases}$$

ou encore :

$$\begin{cases} T \propto M^{\mu(N)} & \text{à } N \text{ fixé} \\ T \propto N^{\nu(M)} & \text{à } M \text{ fixé} \end{cases}$$

Expérimentalement, on peut mesurer les *paramètres d’échelle*,  $\mu$  et  $\nu$ , qui indiquent comment le temps de convergence croît en fonction de  $N$  ou  $M$  (cf. figures 6.12 et 6.13). En l’occurrence, pour la construction asociale du lexique,  $\mu$  varie de 2.2 à 0.9 environ<sup>2</sup> avec une décroissance rapide ( $\mu(N) \propto -N$ ) : autrement dit, plus la taille de la population augmente, plus on gagne de temps dans l’assimilation d’un grand nombre de sens (avec toutefois des temps de convergence qui restent prohibitifs). Par ailleurs,  $\nu$  varie de 3.4 à 1<sup>3</sup>, et cette fois nous verrons que la décroissance est la plus faible de tous les modèles que nous allons examiner.

---

2. Pour  $N$  compris entre 2 et 50

3. Pour  $M$  entre 2 et 20

## ◇6. LA TRANSPOSITION DES MODÈLES

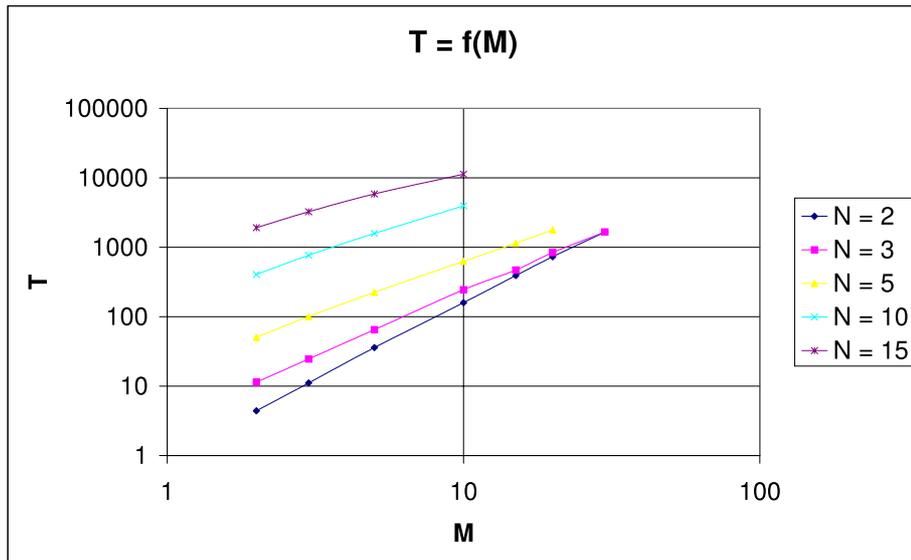


FIG. 6.1 – *Temps de convergence en fonction du nombre de signifiés (meanings) pour des populations de tailles diverses, en l'absence de socialité.*

### Rôle de la socialité dans le jeu de langage

Pour tenir compte d'une structure sociale schématique préexistante, le jeu de langage a subi une autre modification. On suppose que chaque individu  $i$  possède un « rang »  $r_i$  (à proprement parler,  $r_i$  n'est pas le rang de l'individu mais une indication du « poids social » de l'agent  $i$ ; le rang au sens usuel est obtenu en classant les individus par  $r_i$  décroissant). Ce rang est utilisé pour le *choix du locuteur*, en posant que la probabilité pour un agent  $i$  d'être choisi comme locuteur pour chaque jeu de langage est :

$$p(i) = \frac{r_i}{\sum_{j \in [1..N]} r_j}$$

Métaphoriquement, cela revient à considérer que les individus dominants accaparent une « ressource » particulière, le temps de parole, en fonction de leur importance sociale.

Pour ces expériences, nous avons étudié des structures de dominance hiérarchiques, strictement linéaires, dans lesquelles le rang des individus suit une loi géométrique :  $r_i = \gamma^{i-1}$  (avec  $\gamma > 1$ ).

Choix du locuteur mis à part, le jeu de langage reste ici parfaitement conforme à la description que nous en avons donnée (§ 6.2.1).

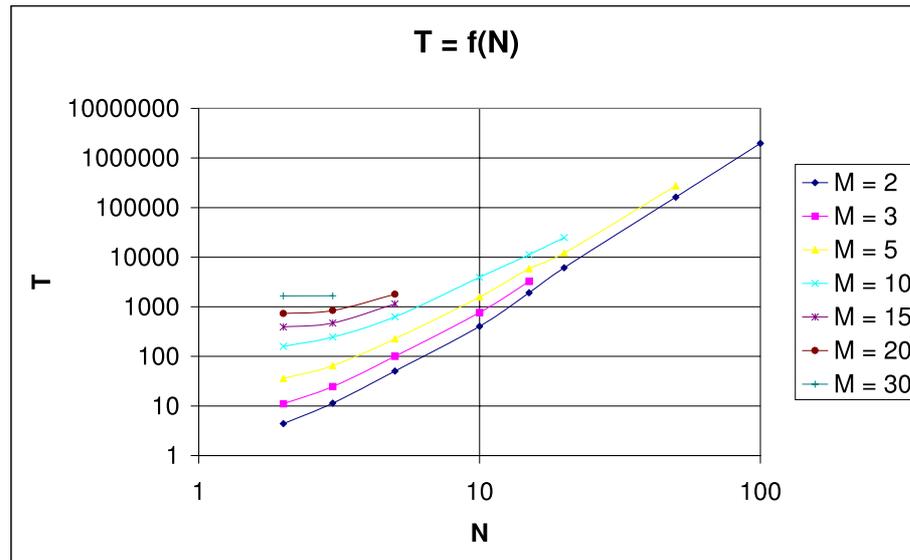


FIG. 6.2 – Temps de convergence en fonction de la taille de la population selon le nombre de signifiés, en l'absence de socialité.

### Résultats expérimentaux avec une hiérarchie fixe

Ici encore, les lois donnant  $T$  en fonction de  $M$  et  $N$  restent les mêmes, mais la convergence est bien plus rapide que dans le cas des agents asociaux, tant en valeur absolue que pour le facteur d'échelle  $\nu$  (entre 1.4 et 0.8) : pour un grand nombre de sens, la sensibilité à la taille de la population est faible.

Cette considérable accélération de la convergence est évidemment le fait d'une augmentation de la cohérence du lexique (puisque certains agents monopolisent le rôle de locuteurs, donc de producteurs de lexique).

Elle peut avoir un intérêt si l'on cherche simplement à obtenir rapidement un lexique partagé dans la population, mais au prix d'une condition : le groupe d'agents est définitivement clos, l'introduction d'un nouvel agent poserait le problème de son positionnement dans la hiérarchie.

Quelle que soit la valeur de  $\gamma$ , le temps de convergence reste du même ordre de grandeur (par rapport à celui, bien plus élevé, du modèle asocial), avec une tendance à la diminution lorsqu'on augmente  $\gamma$ . Toutefois, on observe un phénomène inattendu (cf. figure 6.5) : à  $M$  fixé, le temps de convergence peut *augmenter* progressivement pour  $\gamma > 2$ . Cet effet est d'autant plus marqué pour les populations importantes.

Cela peut s'expliquer par la mise en balance de deux tendances contraires : d'une part, plus le temps de parole se concentre en un agent, plus la cohérence

## ◇6. LA TRANSPOSITION DES MODÈLES

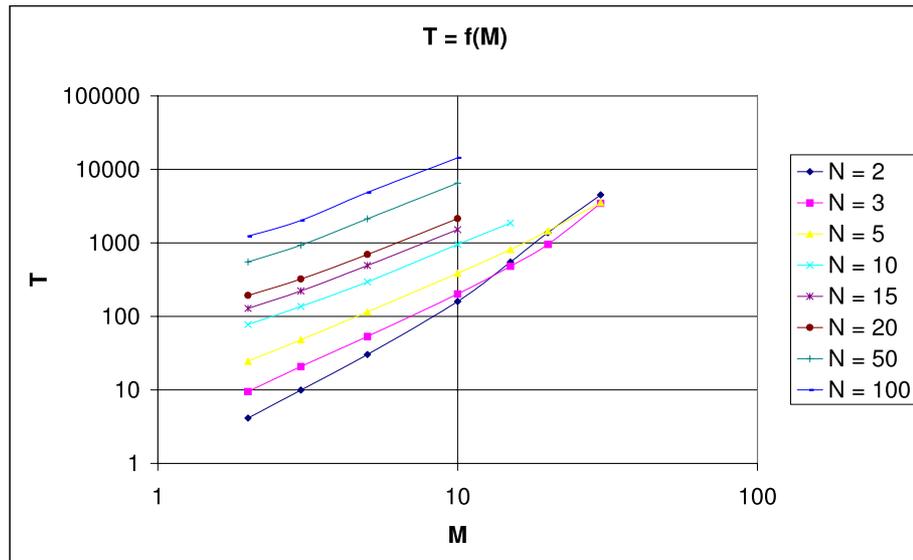


FIG. 6.3 – Temps de convergence en fonction du nombre de signifiés (meanings) pour des populations de tailles diverses. Nous utilisons ici une hiérarchie géométrique ( $\gamma = 1,1$ ).

en production augmente (donc  $T$  tend à croître avec  $\gamma$ ). Mais inversement, plus la population augmente, plus la tâche du locuteur est difficile (puisque un échec de compréhension de l'interlocuteur constitue aussi un échec de production du locuteur), or pour des valeurs de  $\gamma$  et de  $N$  élevées, on tend vers un locuteur unique. Aussi, il ne sert à rien d'avoir une structure *trop* hiérarchisée.

### 6.2.4 Rétroaction du lexique sur la socialité

Essayons maintenant de construire la relation de dominance à partir des interactions lexicales. Il nous faut d'abord définir un équivalent aux comportements agonistiques qui fondent les hiérarchies chez les animaux.

On remarquera que le jeu de langage classique est déjà, lui-même, une forme de comportement agonistique, non pas entre les individus, mais entre les associations (*sens*, *mot*) présentes dans l'ensemble de la population – le lexique construit collectivement n'est rien d'autre que le  $M$ -uplet des associations qui ont pu se maintenir au détriment des autres (KAPLAN 2000c).

Aussi, nous pouvons nous servir du résultat des interactions langagières (succès en communication ou échec) pour tenter de produire une relation de dominance.

Comme nous partirons d'agents indifférenciés, le temps de convergence

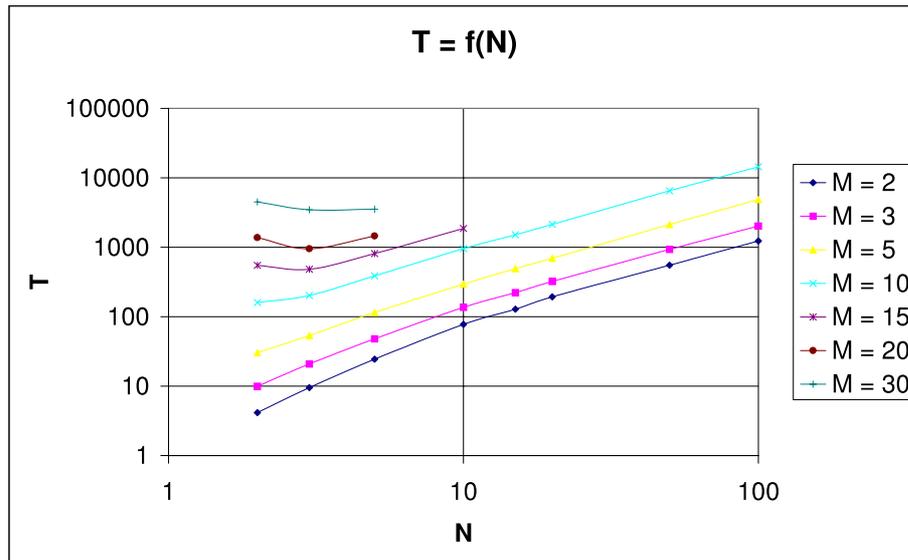


FIG. 6.4 – Temps de convergence en fonction de la taille de la population selon le nombre de signifiés. Nous utilisons ici une hiérarchie géométrique ( $\gamma = 1,1$ ).

des divers modèles que nous allons étudier vont se situer entre celui du modèle asocial et celui des hiérarchies figées.

On pourra comparer sur les figures 6.10 et 6.11 (page 109) quelques-uns des résultats expérimentaux qui sont décrits plus exhaustivement pour chaque modèle.

### Rétroaction sociale

Le premier modèle utilisant une rétroaction sociale consiste à reprendre le mécanisme de constitution de la dominance chez la plupart des espèces animales, par renforcement du rang des agents ( $0 < \Gamma \leq 1$  et  $0 \leq r_i \leq 1$ ) :

$$\begin{cases} r_i \leftarrow r_i + \Gamma(1 - r_i) & \text{en cas de succès du jeu} \\ r_i \leftarrow r_i - \Gamma r_i & \text{en cas d'échec} \end{cases}$$

La variation du temps de convergence en fonction de  $N$  et  $M$  (cf. figures 6.6 et 6.7) suit une loi semblable aux deux précédentes. De façon générale, les temps de convergence sont inférieurs à ceux obtenus pour un modèle asocial, mais restent élevés.

On constate expérimentalement que le paramètre  $\Gamma$  joue un rôle mineur : le temps de convergence, à  $N$  et  $M$  fixés, reste du même ordre de grandeur

## ◇6. LA TRANSPOSITION DES MODÈLES

(cf. figures 6.10 et 6.11).

L'étude des paramètres d'échelle est plus intéressante :

- $\mu$  varie de 2,2 à 1,4 environ, il reste donc supérieur au coefficient homologue du modèle sans lexique et sa loi de variation est proche de celle du modèle à hiérarchie fixe. La croissance du temps de convergence en fonction du nombre de sens n'est donc en rien ralentie par la taille de la population (cf. figure 6.12).
- $\nu$  au contraire varie de 2,9 à 1 pour  $\Gamma = 0,1$  et de 2,6 à 0,9 pour  $\Gamma = 0,5$ ; la décroissance de  $\nu$  en fonction de  $M$  est plus importante pour  $\Gamma = 0,1$ .

Comme ces résultats ne sont pas entièrement satisfaisants (on pouvait s'attendre à une amélioration plus marquée par rapport au modèle sans socialité), nous allons voir ce que donne un modèle social individualisé.

### Socialité individualisée

Dans ce modèle, nous reprenons la notion de « phéromone sociale » émise par chaque agent, et déclenchant ou non une réponse chez un partenaire selon un seuil de réaction (cf. § 5.3).

Dans pareille situation, le « rang » de l'individu est l'intensité de la phéromone de dominance qu'il émet, et n'a donc plus aucune raison de servir à déterminer l'espérance de temps de parole.

Nous avons donc modifié la première étape du jeu de langage :

1. Deux agents  $i$  et  $j$  sont choisis aléatoirement dans la population.
2.  $i$  est le locuteur par défaut ; mais si  $j$  domine  $i$ , c'est-à-dire si  $r_j > \mathcal{S}_{i,j}$ , alors  $j$  devient le locuteur et  $i$  prend le rôle de l'interlocuteur.

La révision des seuils s'effectue selon le mécanisme décrit § 5.3.

L'examen des résultats obtenus (cf. figures 6.8 et 6.9) donne (une fois encore) une loi faisant intervenir les mêmes facteurs d'échelle que dans les autres modèles.

En valeur absolue, le temps de convergence est légèrement supérieur à celui obtenu avec un modèle non individualisé, du moins par rapport à la taille de la population : le facteur  $\nu$  est proche de la valeur du modèle précédent.

En revanche, le facteur  $\mu$ , quant à lui, décroît sensiblement plus vite avec la taille de la population (cf. figure 6.12). Cela signifie que plus la population est importante, plus s'atténue l'effet du nombre de sens à partager sur le temps

de convergence : c'est là une amélioration très nette par rapport aux modèles précédents.

En fait, le choix du locuteur en fonction de l'identité de l'interlocuteur, dépendant en partie des interactions passées entre ces deux agents mais aussi du succès global de chacun en communication, semble permettre des « appariements » pertinents qui accélèrent la convergence, et ce d'autant plus que la population est grande.

### 6.2.5 Conclusions

À travers ces exemples, nous avons pu montrer une transposition simple de nos modèles sociaux à une tâche distribuée, la construction d'un lexique commun par les agents.

Nous avons pu étudier l'impact de la dynamique sociale sur cette tâche, et par comparaison à des modèles de représentation sociale plus simples, montrer l'avantage que présentait ici un système individualisé.

Nous allons maintenant aborder un nouveau domaine d'application, domaine dans lequel les problématiques liées à la socialité n'ont cessé de se développer depuis quelques années : la robotique (et plus généralement les systèmes informatiques ancrés dans un environnement humain).

Nous chercherons à aborder la conception de modèles de comportements sociaux dans ce nouveau contexte en essayant de conserver les principes méthodologiques que nous avons appliqués jusqu'à présent. Nous poserons également la question de la possibilité d'une transposition des modèles issus de la simulation.

◇6. LA TRANSPOSITION DES MODÈLES

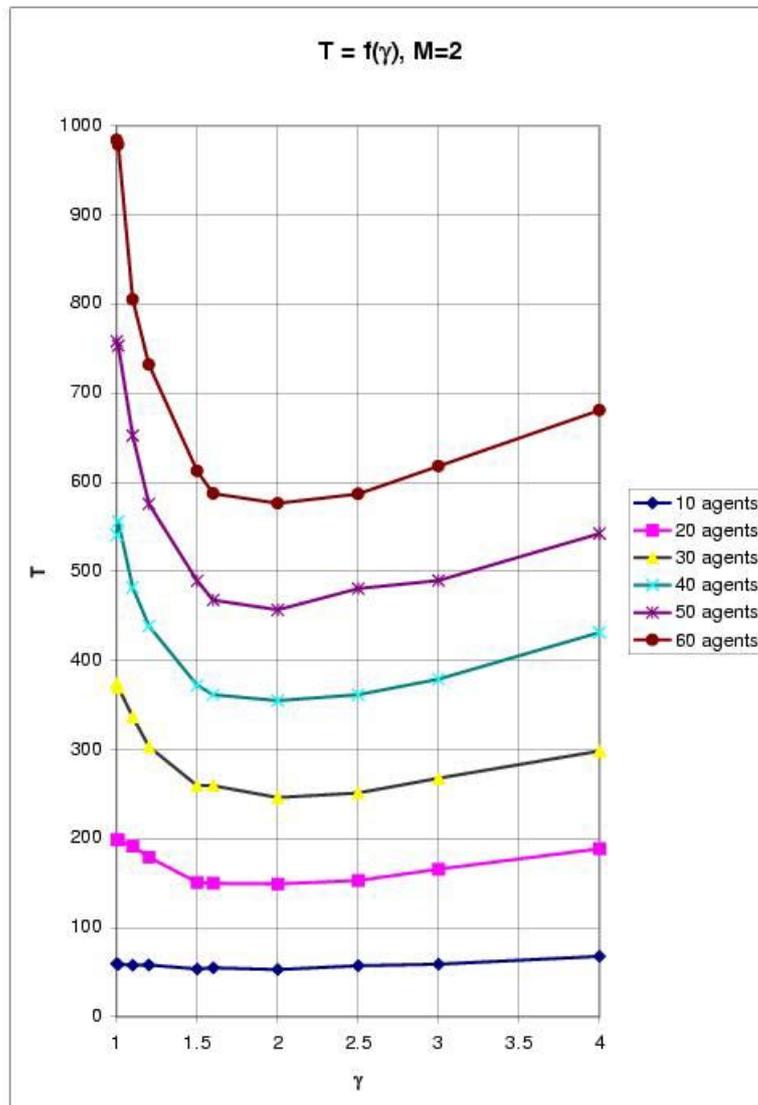


FIG. 6.5 – Temps de convergence en fonction de  $\gamma$  (raison de la suite géométrique déterminant la hiérarchie) pour diverses tailles de population ( $M = 2$ ).

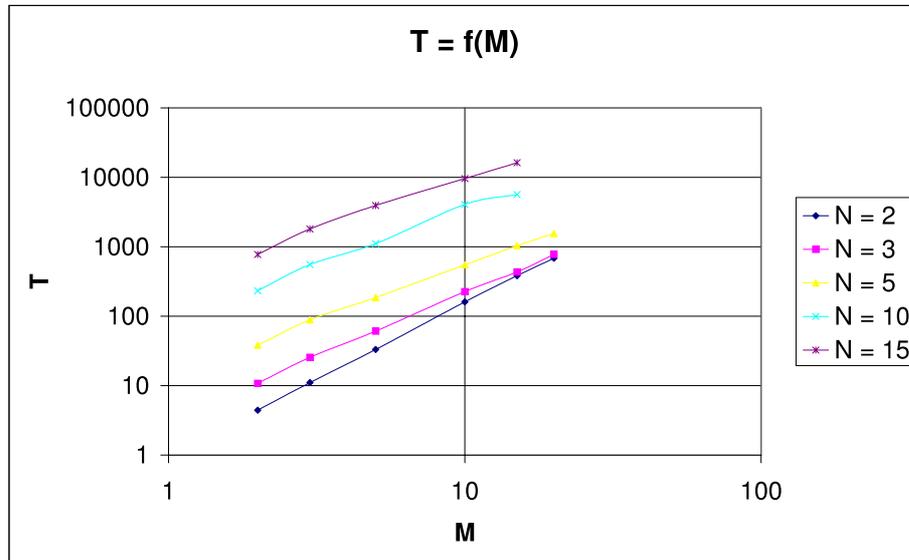


FIG. 6.6 – Temps de convergence en fonction du nombre de signifiés (meanings) pour des populations de tailles diverses. Ici, le lexique rétroagit sur le rang ( $\Gamma = 0,5$ ).

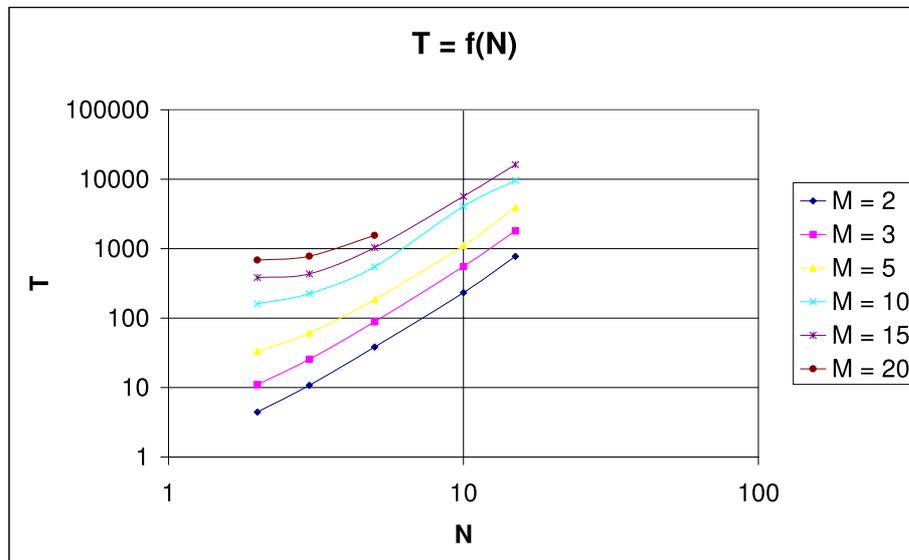


FIG. 6.7 – Temps de convergence en fonction de la taille de la population selon le nombre de signifiés. Le lexique rétroagit sur le rang ( $\Gamma = 0,5$ ).

◇6. LA TRANSPOSITION DES MODÈLES

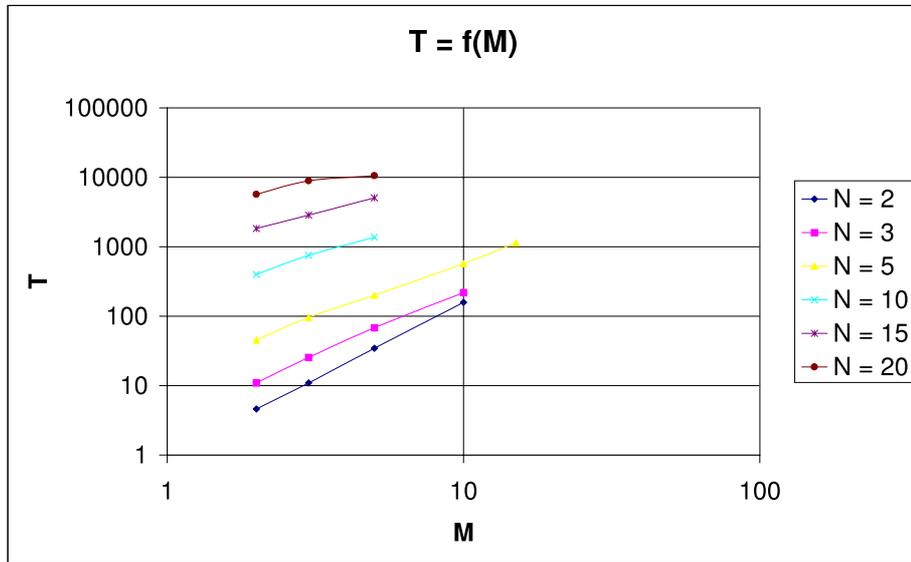


FIG. 6.8 – Temps de convergence en fonction du nombre de signifiés (meanings) pour des populations de tailles diverses. La relation de dominance est individualisée ( $\Gamma = 0,5$ ).

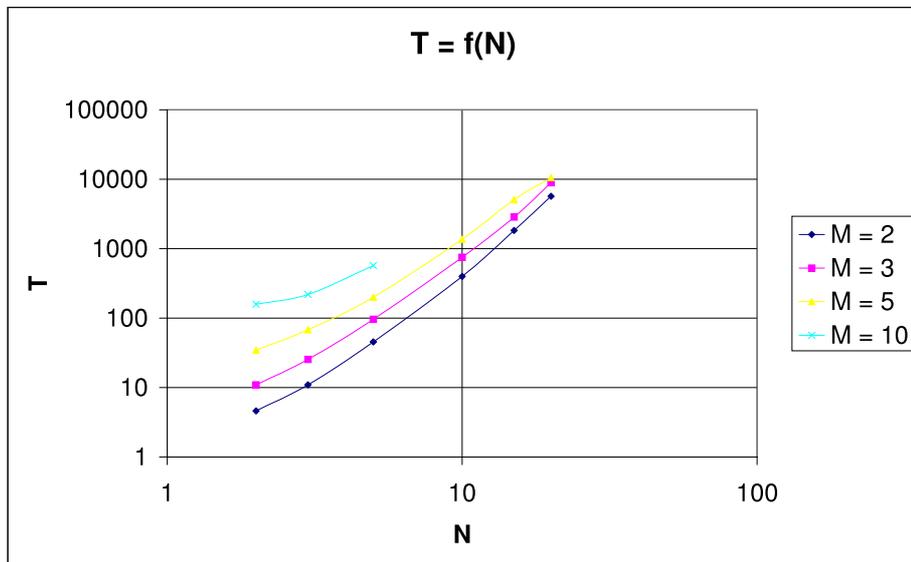


FIG. 6.9 – Temps de convergence en fonction de la taille de la population selon le nombre de signifiés. La relation de dominance est individualisée ( $\Gamma = 0,5$ ).

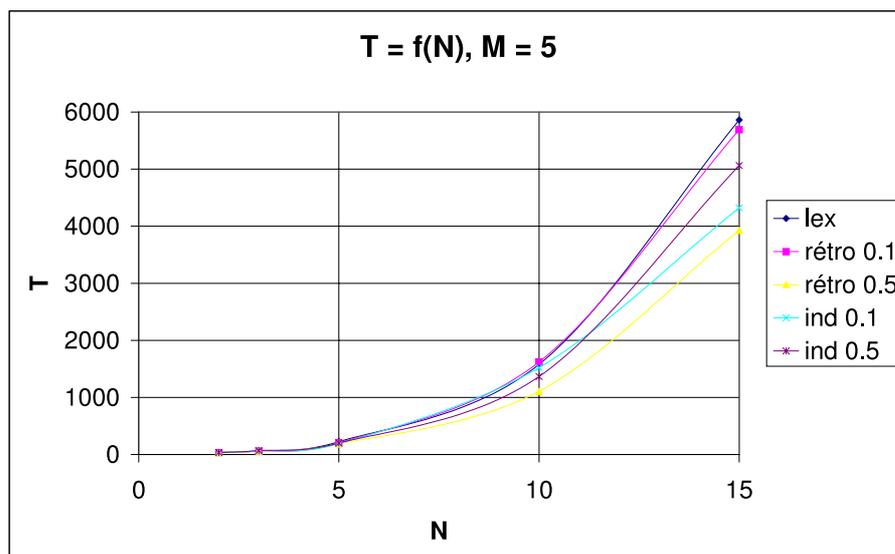


FIG. 6.10 – Comparaison du temps de convergence en fonction de la taille de la population pour les divers modèles étudiés, sauf hiérarchies ( $M = 5$ ). Légende : lex – modèle asocial; rétro – modèle avec rétroaction (et valeur de  $\Gamma$ ); ind – modèle individualisé (et valeur de  $\Gamma$ ).

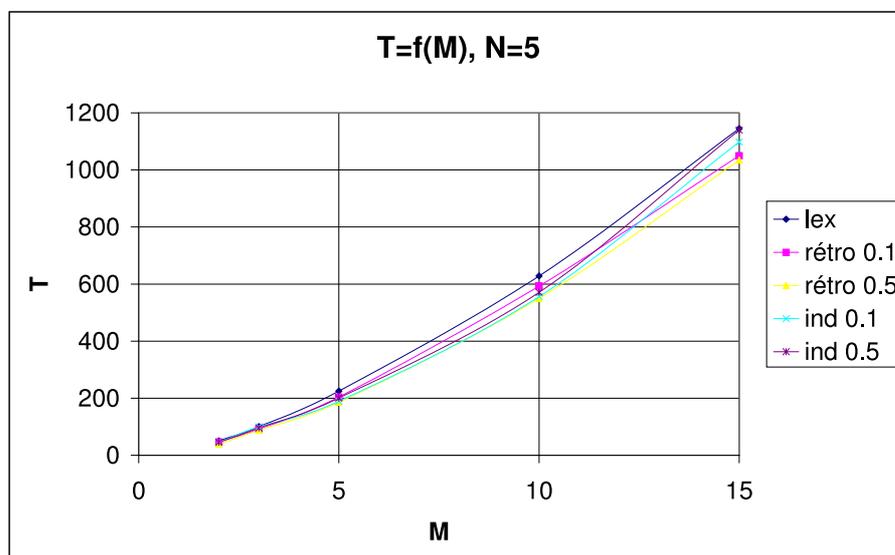


FIG. 6.11 – Comparaison du temps de convergence en fonction du nombre de sens pour les divers modèles étudiés, sauf hiérarchies ( $N = 5$ ). Légende : lex – modèle asocial; rétro – modèle avec rétroaction (et valeur de  $\Gamma$ ); ind – modèle individualisé (et valeur de  $\Gamma$ ).

◇6. LA TRANSPOSITION DES MODÈLES

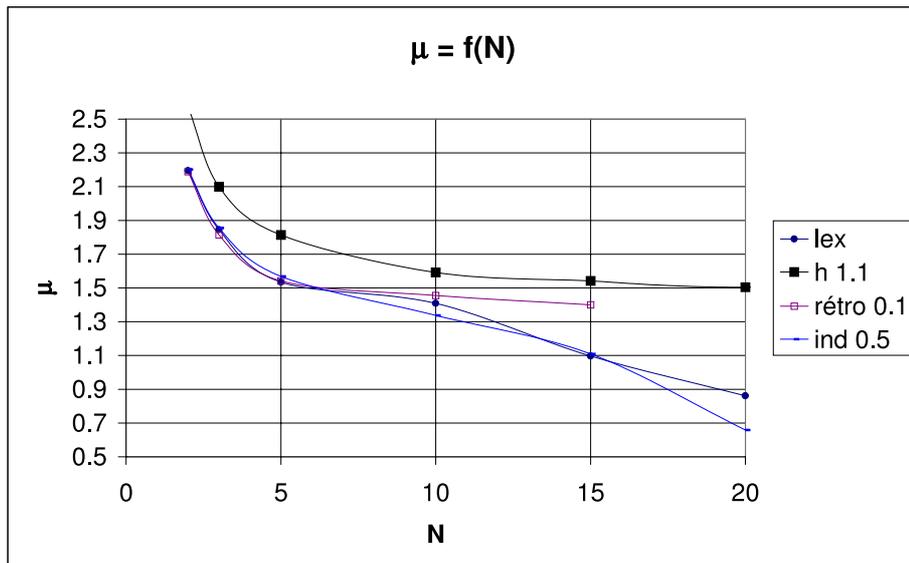


FIG. 6.12 – *Mesure expérimentale de la variation de  $\mu$  en fonction de la taille de la population pour les divers modèles étudiés. Légende : lex – modèle asocial; h – modèle hiérarchique fixe (et valeur de  $\gamma$ ); rétro – modèle avec rétroaction (et valeur de  $\Gamma$ ); ind – modèle individualisé (et valeur de  $\Gamma$ ).*

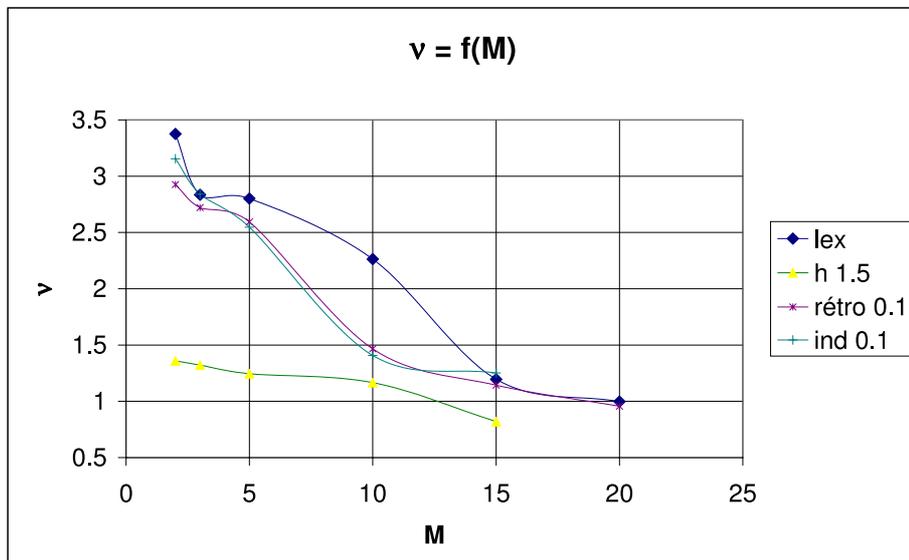


FIG. 6.13 – *Mesure expérimentale de la variation de  $\nu$  en fonction du nombre de significés pour les divers modèles étudiés. Légende : lex – modèle asocial; h – modèle hiérarchique fixe (et valeur de  $\gamma$ ); rétro – modèle avec rétroaction (et valeur de  $\Gamma$ ); ind – modèle individualisé (et valeur de  $\Gamma$ ).*

---

## Deuxième partie

# MODÈLES DE CONSTRUCTION DE LA SOCIALITÉ POUR LES SYSTÈMES SITUÉS

EDM

---

*« À un monde qui n'est pas notre représentation, les lois numériques sont tout à fait inapplicables : elles n'ont cours que dans le monde de l'homme. »*

NIETZSCHE. *Humain, trop humain*, aph. 19

## Résumé

Dans cette partie, nous allons nous intéresser à la conception de comportements sociaux dans le cadre d'un type d'application dont nous pensons qu'il constitue un excellent paradigme de système multi-agent ouvert : l'intégration de groupes de robots dans un environnement social humain.

Nous commencerons par un examen des travaux récents sur la robotique collective et sur les machines « socialement intelligentes », avant de présenter le projet MICRobES, qui vise à étudier l'immersion de groupes de robots dans des collectivités humaines.

Ce projet sera l'occasion de remettre en questions conceptuelles majeures, qui nous conduira à explorer une nouvelle piste pour l'étude des mécanismes sociaux, à travers deux métaphores complémentaires : celle de l'organisme et celle de l'écosystème. Les schémas explicatifs empruntés à l'éthologie nous pousseront à rechercher des modèles capables de produire le *développement* et l'*évolution* de SMA sociaux. Nous proposerons donc des principes méthodologiques généraux en ce sens (l'*Éthogénétique*) ainsi qu'un *framework* permettant leur application, ATNoSFERES.

---

## Chapitre septième.

# L'émergence de la socialité robotique

---

*Le maître [Hephaestos] s'appuyait sur  
deux servantes,  
Qui étaient tout en or, mais semblaient  
des vierges vivantes.  
Elles avaient de la raison, possédaient voix  
et force,  
Et tenaient des dieux mêmes leur science  
du travail.*

HOMÈRE. *Iliade*, Chant XVIII

*Depuis quelques années, on assiste à la confluence, autour de la robotique mobile, de disciplines parfois assez différentes : robotique collective, interaction homme-machine, IAD, apprentissage, etc. Les robots mobiles sont en effet considérés comme un support d'expérimentation privilégié pour un certain nombre de techniques, pour lesquelles des approches classiques s'avèrent limitantes. Nous allons voir où plongent les racines de cette évolution et les nouveaux problèmes qu'elle soulève.*

### 7.1 Le robot non fonctionnel

Qu'est-ce qu'un robot ? Le dictionnaire donne à ce mot une double acception :

1° « machine, automate à l'aspect humain, capable de se mouvoir et d'agir » (ROBERT 1993);

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

2° « mécanisme automatique à commande électronique pouvant se substituer à l'homme pour effectuer certaines opérations, et capable d'en modifier de lui-même le cycle en appréhendant son environnement. »

Cette définition du *robot* recouvre en vérité deux points de vue extrêmes : d'une part, le robot fantasmagorique, celui de la science-fiction depuis les « ouvriers artificiels » de la pièce de Karel ČAPEK (ČAPEK 1920), créé par l'homme à son image et menaçant de se rebeller, de le renverser<sup>1</sup> ; d'autre part, le robot industriel, machine-outil sophistiquée, programmé pour réaliser à la place de l'homme (et pour son plus grand confort) une tâche fonctionnelle précise avec plus ou moins d'autonomie.

Dans les deux cas, le robot est construit *en vue d'une finalité* préexistante, extérieure à lui, imposée par l'homme. Or une troisième catégorie de robots est apparue depuis quelques années : le robot *d'expérimentation*.

La robotisation de l'industrie a en effet conduit paradoxalement à des situations dans lesquelles des machines doivent opérer dans des environnements non seulement de plus en plus complexes, mais qui surtout *ne peuvent plus être simplifiés*. Dès lors, les techniques habituellement employées (programmes figés, planification, utilisation de représentations symboliques ou de fortes connaissances *a priori* sur l'environnement) deviennent inappropriées, car le monde où opère le robot n'est plus clos.

Les alternatives permettant à l'homme de compenser l'insuffisante autonomie du robot, comme la téléopération, n'ont d'intérêt que dans des domaines où le robot est un outil nouveau (la microchirurgie en est un bon exemple), mais certainement pas lorsque, déjà utilisé dans des situations facilement contrôlables, on cherche précisément à en étendre les domaines d'applications.

Aussi, il devient indispensable de trouver de nouvelles techniques permettant l'utilisation de ces robots, et cela ne peut se faire qu'à condition de tenir compte explicitement de la complexité de l'environnement dans la « programmation » du robot.

Le robot, en tant que machine directement au contact d'un environnement physique (contrairement à un programme dans un ordinateur), apparaît

---

1. Les robots de ČAPEK font écho à un être légendaire plus ancien, le *Golem* créé d'après la légende à partir d'argile par le célèbre rabbin LOEW de Prague ; destiné à protéger le ghetto et à réaliser quelques tâches domestiques, il finit par se dérégler et devenir dangereux. Ces mythes sont l'illustration des craintes liées à l'usurpation des pouvoirs démiurgiques.

donc comme un support expérimental *en lui-même* pour des techniques issues de l'Intelligence Artificielle ou de domaines proches. De simple outil, le robot devient une fin en soi, en même temps qu'un moyen d'investigation scientifique.

On a donc désormais affaire à une robotique « comportementale » (ARKIN 1998; BROOKS 1999), qui étudie comment doter un robot ou un groupe de robots d'un certain nombre de « comportements » (à peu près au même sens que pour un agent ou pour un SMA), de façon à lui (leur) donner plus de plasticité et de robustesse vis-à-vis d'environnements complexes.

## 7.2 La robotique collective : de la simulation à la mise en situation

### Les robots comme agents

La naissance des robots mobiles autonomes a conduit l'IAD à considérer les *groupes de robots* comme des applications possibles des recherches menées sur les SMA. De même que dans les simulations éthologiques, il était intuitif de modéliser l'individu par un agent, les agents ont été fréquemment traités métaphoriquement comme des robots mobiles autonomes. Les mécanismes de coopération, de coordination ou de compétition au sein d'une équipe d'agents prennent en effet tout leur sens lorsqu'ils sont appliqués à des équipes de robots, comme dans (DROGOUL et FERBER 1993; ASADA et al. 1998, entre autres).

L'assimilation du robot à l'agent est facilitée par leur *analogie structurelle* : un robot est une machine placée dans un « environnement » (en l'occurrence une région de l'espace au sens commun), qui possède des dispositifs de perception (capteurs ou senseurs) et d'action (effecteurs ou actuateurs). Cette machine est contrôlée par un certain nombre de programmes qui servent à sélectionner des actions en fonction de perceptions et des événements antérieurs.

Cela correspond très précisément à une architecture d'agent. Dès lors, on peut viser l'obtention d'un comportement de robot (ARKIN 1998) de même qu'on dote un agent d'un comportement.

Deux remarques, toutefois, s'imposent. Premièrement, ces « robots » évoqués par l'IAD sont longtemps restés, et restent encore souvent, des robots « virtuels », c'est-à-dire qu'ils sont avant tout une métaphore pour donner du sens à l'activité collective des agents opérant dans le système, mais les applications utilisant de vrais robots sont largement minoritaires. Cela n'est

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

pas sans soulever quelques questions sur lesquelles nous reviendrons : le passage de la simulation au réel est-il possible ? est-il utile ? observera-t-on les mêmes résultats ? la métaphore robotique ne sert-elle qu'à donner une forme facilement appréhendable à un problème tout en masquant une scission profonde entre agents simulés et véritables robots ?

Deuxièmement, il est généralement possible dans ces travaux de simuler les robots au moyen d'agents plus ou moins réactifs. On pourra notamment se reporter aux études sur la collecte distribuée d'échantillons, le tri collectif et autres robots fourrageurs (STEELS 1989; DENEUBOURG et al. 1991; GOSS et DENEUBOURG 1992; DROGOUL et FRESNEAU 1998, etc.). Les comportements de ces robots peuvent faire l'objet d'une conception incrémentale simple (DROGOUL 1993, 3<sup>e</sup> partie). Toutefois, si ces simulations ont su démontrer brillamment la capacité d'agents réactifs à effectuer collectivement des tâches complexes, sans en avoir de représentation globale, l'environnement dans lequel ces agents étaient placés n'en reste pas moins *entièrement contrôlable et contrôlé*. Cette hypothèse particulièrement forte est complètement irréaliste pour des robots réels : que devient- alors la méthodologie de conception ?

### Les équipes de robots

Parallèlement aux premières recherches en robotique collective (MATARIC 1993; MATARIC 1995; ARKIN et BALCH 1997; MATARIC 1997) sont apparues les compétitions RoboCup (ASADA et al. 1998; KITANO et ASADA 1997) et MiroSot (DROGOUL et DUHAUT 1996) visant à faire s'affronter des équipes de « robots footballeurs » (cf. figure 7.1). Ces compétitions ont d'abord semblé constituer une plate-forme d'expérimentation intéressante pour confronter les stratégies collectives défendues par telle ou telle équipe de recherche, non plus seulement en simulation mais également sur des robots « en chair et en os ».

Or le passage du virtuel au réel a soulevé d'énormes difficultés. Les stratégies gagnantes en simulation se sont révélées souvent inefficaces face à des adversaires aux comportements plus rudimentaires mais s'appuyant sur des caractéristiques physiques favorables (la masse des robots par exemple).

De plus, un des objectifs de ces compétitions était de pouvoir évaluer l'efficacité de divers comportements collectifs afin de les transposer par la suite à d'autres applications. Malheureusement, la recherche de robots « efficaces » *en tant que footballeurs*, dans un contexte aux règles (trop) bien définies (STONE

et VELOSO 1997), a induit systématiquement des biais importants dans la conception des comportements, des techniques de perception, de représentation, de sélection d'action, etc. (DROGOUL 2000).

Enfin, quoique déjà fort complexe, l'environnement dans lequel évoluent les robots footballeurs reste hautement contrôlé (le terrain ne dépasse pas quelques mètres carrés, il est fortement balisé), et quelques paramètres physiques imprévus (brouillage des communications radio, flashes des spectateurs, etc.) ont entraîné parfois des perturbations catastrophiques des comportements de jeu.

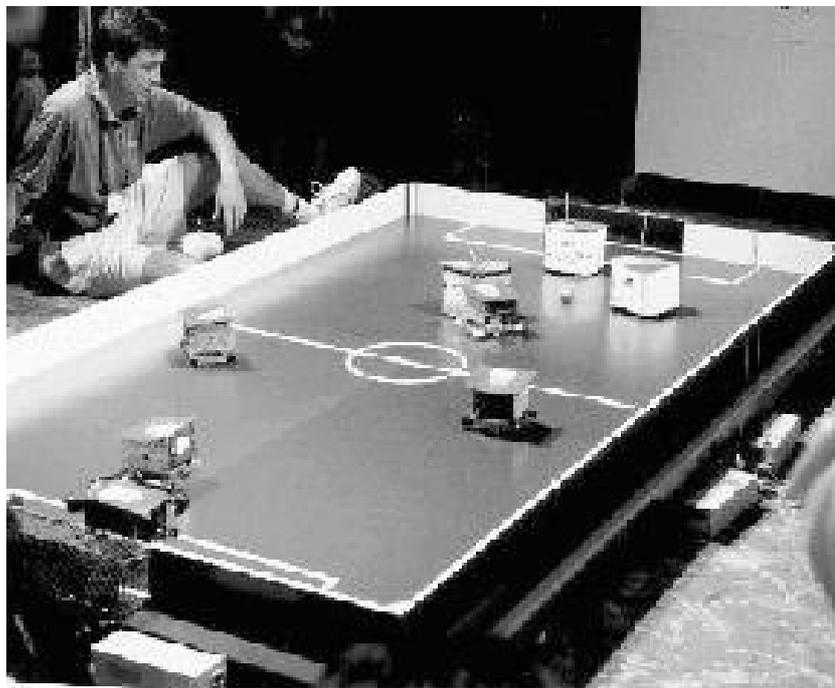


FIG. 7.1 – *Un exemple de match de RoboCup (robots de petite taille). Les équipes comptent trois robots et s'affrontent durant quelques minutes. L'environnement est fortement aménagé.*

### L'impossible passage au réel ?

Est-ce à dire que le passage aux robots réels est *trop difficile*? Que ce serait un but chimérique? Loin s'en faut. Au contraire, selon le point de vue défendu par BROOKS (BROOKS 1991), la *mise en situation*<sup>2</sup> est indispensable

2. C'est le terme que j'utiliserai pour l'anglais *situatedness*, mot à mot « le fait d'être situé », d'être plongé dans un environnement réel.

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

pour traiter la complexité du réel. Plutôt que de chercher à faire un modèle du monde le plus précis possible, et d'y appliquer des techniques de résolution de problème, de planification, etc., les architectures préconisées par BROOKS doivent s'appuyer sur les perceptions et les capacités d'action telles qu'elles se réalisent effectivement dans l'environnement physique.

Le comportement, si complexe soit-il, apparaît alors comme un phénomène émergent, rendu nécessaire par les possibilités effectives de perception et d'action (à la fois celles du robot et celles qui sont réalisables dans l'environnement considéré).

Mais cette approche se double d'une contrepartie relativement lourde : puisque le comportement *émerge* de la mise en situation du robot dans son environnement, cela signifie notamment qu'il ne saurait être question de prétendre produire un comportement adéquat à partir d'architectures conçues dans un environnement simulé. Il s'agit là d'un point de vue extrême, mais il est néanmoins de nature à jeter la suspicion sur la simulation.

Nous avons en effet mentionné dans l'Introduction (cf. § 2.1.2) qu'un modèle, qu'une simulation sont censés reproduire les éléments pertinents de la situation réelle. En l'occurrence, c'est précisément la définition exhaustive des « éléments pertinents » du monde réel, du point de vue du comportement du robot, qui pose problème. Si comme le prétend BROOKS le monde n'a pas de meilleur modèle que lui-même, tout projet de robotique basé sur la simulation préalable des comportements est voué à l'échec !

### 7.3 La robotique « sociable » : de l'ergonomie à l'implication affective

Depuis le développement de la robotique mobile et son utilisation dans des environnements peuplés d'êtres humains, une autre discipline a investi le support expérimental que constituent les robots : l'Interaction Homme-Machine (IHM).

Si la définition du robot rappelée en tête de ce chapitre est assez restrictive, il n'en reste pas moins qu'elle rend bien compte de l'usage lexical, c'est-à-dire de ce qu'*est* un robot pour la majorité des gens. En particulier, l'imaginaire de tout un chacun est marqué par la figure du robot de science-fiction, et on peut légitimement penser que cela affecte la manière dont un humain peut se comporter vis-à-vis d'un robot et ce qu'il peut en attendre.

Toutefois, l'intérêt de l'IHM pour les robots dépasse largement le cadre de la culture littéraire, pour s'enraciner en fait très profondément dans des mécanismes anthropologiques et psychologiques communs à l'ensemble de l'espèce humaine. Ses applications ne visent pas qu'à faciliter l'interaction entre l'homme et le robot, elles peuvent également tenter d'*intégrer* les robots dans un milieu humain, ou encore avoir un objectif thérapeutique.

Nous allons examiner quelques exemples d'applications dans lesquelles des robots interagissent avec des humains, dans des situations où cette interaction est recherchée et où elle constitue en fait l'essentiel de leur rôle.

### Des robots guides de musées

De tels robots sont utilisés individuellement, pour assurer une visite guidée selon un itinéraire plus ou moins prédéfini (avec en général une capacité d'adaptation).

Citons d'abord le robot SAGE, au *Carnegie Museum of Natural History*, qui suit un itinéraire préétabli à l'intérieur du *Dinosaur Hall* (NOURBAKHS 1998). Il se repère au moyen d'amers visuels placés au-dessus des portes. D'une interactivité limitée, puisqu'il ne fait que diffuser un documentaire enregistré sur CD-ROM, il est en revanche capable d'éviter des obstacles (visiteurs humains compris) et de se recharger de façon autonome.

Ce robot peut fonctionner durant plusieurs jours. L'analyse des « pannes » qui sont advenues se répartissent en trois catégories : 1° des problèmes robotiques (imprécision des sonars, blocage des roues, etc.); 2° des problèmes informatiques (bogues dans les algorithmes de vision et de navigation notamment); 3° des actions humaines inattendues (caméra dérégulée par un visiteur, robot enfermé par un vigile, etc.).

Dans tous les cas, le point à souligner est qu'une architecture « cognitive » (planification par exemple) n'aurait été d'aucun secours pour anticiper ou résoudre ce type d'accident. C'est plutôt dans une intégration de l'ensemble des perceptions qu'il faut chercher des pistes pour rendre plus robuste le comportement du robot.

Les robots RHINO (BURGARD et al. 1999) et MINERVA (THRUN et al. 1999) interagissent plus étroitement avec l'homme. Ils sont tous deux conçus pour opérer dans des environnements inconnus, par exemple pour des expositions durant quelques jours.

Dans un premier temps, ils sont téléopérés, afin de cartographier l'espace

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

dans lequel ils auront à se déplacer. Puis, en fonctionnement autonome, ils organisent leur parcours en fonction de leur localisation sur la carte, la durée de la visite, l'énergie restante, les centres d'intérêts des visiteurs, etc.

MINERVA dispose de surcroît d'un visage articulé et d'un système de synthèse vocale qui lui servent à exprimer des « humeurs » (expressions faciales et intonation). Ces états dépendent de la réaction des humains à son égard et modifient le comportement du robot. Nous allons voir comment cela peut contribuer à son intelligibilité.

### Robots à expressions faciales

D'autres expériences utilisent des robots pour étudier des compétences cognitives innées de l'espèce humaine, ou s'en servent pour permettre aux humains d'établir une relation plus complexe avec un robot qu'avec une machine ordinaire.

L'homme en effet, comme la plupart des mammifères, est extrêmement sensible aux expressions faciales, dont la très grande majorité se retrouve inchangée dans toutes les cultures, ainsi que l'ont montré nombre de travaux en éthologie humaine (EIBL-EIBESFELDT 1972; EIBL-EIBESFELDT 1976; MORRIS 1977; ZWANG 2000). Ces expressions faciales servent à exprimer et reconnaître des états affectifs (des « émotions »).

De nombreux travaux sur des robots capables de mettre en œuvre des expressions faciales simples lors de leurs interactions avec l'homme ont été entamés récemment : par exemple KISMET (cf. figure 7.2) au MIT (BREAZEAL et SCASSELLATI 1998; BREAZEAL et FITZPATRICK 2000), ou encore le robot FELIX (CAÑAMERO et FREDSLUND 2000).

Les idées qui sous-tendent ces projets sont de deux sortes. Premièrement, l'interaction avec le robot doit être facilitée. Il est question, dans ce cas, de « vraisemblance » des interactions. Le robot peut *faire illusion* lorsqu'il interagit, donner à penser qu'il « est de telle ou telle humeur », sembler « réaliste » en tant qu'*individu doué d'une intentionnalité*. En somme, il s'agit là d'une méthode proche de l'idée d'*imitation game* de TURING (TURING 1950). Pour qu'une machine puisse être considérée comme « intelligente » socialement, il suffit qu'un partenaire humain soit poussé à interagir avec elle avec autant de spontanéité et de convivialité qu'il le ferait vis-à-vis d'un autre humain.

Deuxièmement, l'intelligence du robot est d'abord considérée comme une intelligence sociale, et plus particulièrement comme une intelligence *émotion-*

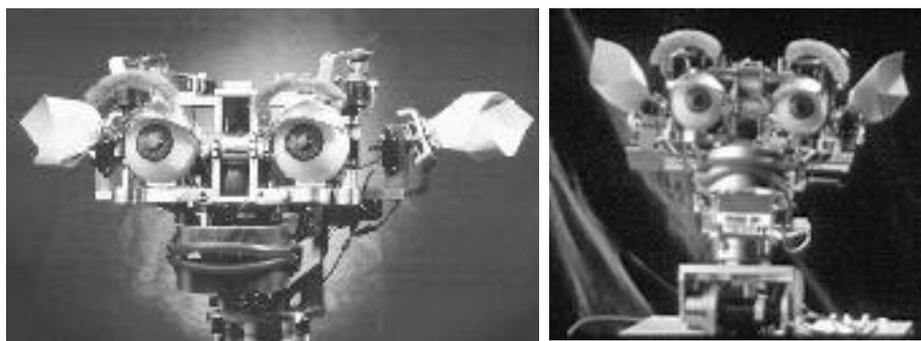


FIG. 7.2 – *Le robot KISMET, développé au MIT, a un visage articulé qui lui permet de reproduire des expressions faciales schématiques.* — d’après photos MIT et NBC.

*nelle*. Autrement dit, le comportement du robot revêt aux yeux d’un partenaire humain *plus d’intelligence* qu’il n’en a réellement dès qu’il est capable de produire de *l’empathie* : celui qui interagit avec le robot « ressent » une émotion qui est censée correspondre avec celle affichée par le robot. Ce phénomène de « compréhension mutuelle » intervient dans une part importante des relations sociales humaines et même animales (DAUTENHAHN 1997a; DAUTENHAHN 2000b).

### Robots imitateurs

La conception de robots sociables repose également sur l’hypothèse selon laquelle la socialité humaine est fortement liée à la capacité à *imiter*. L’imitation jouerait un rôle proche de celui des expressions faciales en facilitant un contact empathique entre la machine et un partenaire humain. On peut citer entre autres ROBOTTA, un robot-poupée capable d’apprendre des mouvements au cours d’une interaction et de les reproduire (BILLARD et al. 1998; BILLARD 2000).

En outre, l’imitation permet d’enseigner au robot un certain nombre d’actions, à travers une expérience physique directe de l’environnement. Cela pose le problème de la différence des capacités d’action et de perception de l’environnement par l’homme d’une part et par le robot d’autre part, ce qui revient à se demander dans quelle mesure il y a un sens pour un robot à « imiter » un humain qui ne possède pas le même corps physique que lui. Nous y reviendrons dans le chapitre 9.

### **Robotique sociale, intentionnalité et autisme**

Enfin, nous mentionnerons un dernier ensemble d'applications de la robotique sociable, à visée thérapeutique : il s'agit de l'utilisation de robots dans le cadre du traitement de l'autisme.

L'autisme est une maladie qui se caractérise par des difficultés, voire une impossibilité, à établir des relations sociales, à communiquer, et par un répertoire d'activités restreint (adoption et répétition de comportements stéréotypés, focalisation sur des environnements stables et connus, etc.). Parmi les causes invoquées pour expliquer ces troubles, on peut mentionner notamment l'hypothèse d'un désordre de la « théorie de l'esprit » (BARON-COHEN et al. 1985), celle d'une absence de mécanismes empathiques (DAUTENHAHN 1997a), ou des dysfonctionnements dans le contrôle de la pensée et de l'action (RUSSELL 1997).

Des travaux récents s'intéressent donc à l'utilisation de robots sociables pour aider les enfants autistes à améliorer la coordination de leurs interactions avec l'environnement.

C'est le cas par exemple du projet AURORA (DAUTENHAHN 1999a; DAUTENHAHN 1999b; DAUTENHAHN 2000a) qui vise à construire un robot-jouet répondant à cette fonctionnalité. Le robot social constitue alors un niveau intermédiaire entre l'environnement physique – prévisible et rassurant – et l'environnement social humain – très complexe, fortement changeant et anxiogène (DAUTENHAHN et WERRY 2000).

On peut également mentionner les travaux menés dans ce sens à l'Université de Sherbrooke, au LABORIUS (MICHAUD et al. 2000; MICHAUD 2000), en particulier le robot-boule ROBALL (MICHAUD et CARON 2000), capable de se déplacer spontanément dans n'importe quelle direction. Les études menées sur de jeunes enfants montrent qu'ils préfèrent nettement interagir avec ce type de jouet, auquel ils peuvent prêter une intentionnalité, qu'avec une balle « ordinaire ». Cette entité, identifiée comme « objet » mais manifestant pourtant une apparente intentionnalité, constitue un support thérapeutique proche des robots d'AURORA. Mais un objet comme ROBALL constitue déjà en lui-même un *jouet* d'une nature complètement nouvelle, puisque dépourvu de passivité.

### Les robots de compagnie et les robots jouets

Parallèlement aux recherches sur la sociabilité des robots et aux applications citées plus haut, on a vu se développer depuis quelques années un champ d'utilisation nouveau pour les robots mobiles : la « robotique de compagnie ».

Le robotique de compagnie s'inscrit en fait dans la lignée des « créatures digitales » amorcée en 1995 avec le TAMAGOCHI, jusqu'au FURBY (1998) en passant par des logiciels de « jeu » comme la série des CREATURES.

Le fameux robot-chien AIBO de SONY a inauguré l'ère des robots de compagnie, avec déjà quelques concurrents sur ses traces, comme le prototype R-100 de NEC ou l'humanoïde P3 de HONDA.

Ce dernier a un statut particulier au sens où il est conçu à des fins utilitaires (l'assistance aux personnes âgées ou handicapées par exemple). Mais pour la grande majorité des cas, le but affiché est une *implication affective* des humains dans leur interaction avec le robot. Cela est d'autant plus le cas avec le nouveau prototype humanoïde de SONY, le SDR-3, haut d'environ 50 cm.

De telles créatures sont totalement dénuées de fonctionnalité propre (KAPLAN 2000a) : il est certes toujours plus ou moins possible de leur en faire apprendre une, ou de la leur programmer explicitement, mais tel n'est pas leur but. Elles sont à l'opposé de la définition du robot industriel et rejoignent peu à peu celle du robot de science-fiction. Il est même assez difficile de dire *ce qu'elles sont exactement* :

*« Ces machines restent pourtant très rudimentaires, même si elles s'appuient sur de véritables prouesses techniques. Les considérer comme des « animaux » est gênant. Ce sont certes des machines, mais elles sont très particulières, et les mots pour les caractériser justement restent difficiles à trouver. »* (LESTEL 1998a)

Les interactions possibles entre un humain et une machine de ce type sont assez variables, pouvant aller d'une relative indifférence et d'un certain agacement jusqu'à une implication assez forte (on a pu s'en rendre compte lors de l'apparition des premiers TAMAGOCHIS qui ont réussi à pousser quelques-uns de leurs propriétaires à les considérer comme de « vrais » animaux).

Même si ce phénomène n'est encore qu'assez marginal, il est en train de changer notre manière de penser l'homme et son rapport à l'animal, au vivant, aux machines. Il s'agit bien d'une évolution culturelle à grande échelle, et non cantonnée aux seuls milieux scientifiques ou technologiques, puisqu'il s'agit d'objets distribués sur une échelle relativement vaste et fortement médiatisés.

## 7.4 L'émergence d'une « Robotique Collective Ouverte »

C'est dans ce contexte de confluence entre la robotique collective et la robotique sociable que nous avons proposé le concept de « Robotique Collective Ouverte » ou RCO (DROGOUL et PICAULT 1999; PICAULT et DROGOUL 2000a), nommée ainsi par analogie avec les « systèmes ouverts » (HEWITT et INMAN 1991).

### Un mouvement de fond de la robotique

À travers la notion de « système ouvert », HEWITT a su pressentir l'éclosion actuelle de collectivités mêlant agents logiciels et groupes humains (agents sur réseau, assistants personnels, etc.). Un phénomène analogue se dessine nettement dans l'évolution des travaux menés en robotique, et l'on pourra assister dans les prochaines années à la formation de véritables *collectivités mixtes hommes/robots*.

Cela suppose de faire coexister un groupe de robots et un groupe humain *dans un même environnement*, sans devoir y apporter de modifications conséquentes, étant entendu que ce sont les robots qui doivent s'adapter à un environnement peuplé d'humains, et non l'inverse.

Cet objectif est partagé par bon nombre de roboticiens (THRUN 1998), en dépit de l'absence de travaux abordant directement cette problématique. Toutes les expériences que nous avons mentionnées précédemment font appel soit à des environnements fermés, strictement contrôlés (comme dans les compétitions de la RoboCup), soit à des environnements ouverts aux humains mais fortement adaptés aux robots.

La « Robotique Collective Ouverte » visera donc à étudier les conditions nécessaires à l'intégration de groupes de robots dans des environnements habités ordinairement par des groupes humains.

Cette étude est essentiellement *expérimentale* : si les partenaires humains impliqués dans l'expérience sont peu manipulables, les robots, eux, le sont. La RCO se pose d'emblée comme une forme d'*éthologie synthétique*, dans la mesure où, comme dans un laboratoire virtuel, il est possible de modifier le comportement des agents que l'on cherche à observer (du moins, de certains d'entre eux) et de déterminer comment ces variations affectent l'ensemble de l'écosystème.

La nature intrinsèquement expérimentale de la RCO impose de procéder selon une approche *incrémentale*, au cours de laquelle des modèles d'agents, d'abord simples, seront implémentés pour contrôler les robots, observés *en situation*, analysés et seulement alors complexifiés peu à peu pour permettre l'organisation des robots et leur insertion au sein de la communauté humaine qui peuple leur environnement. Il serait parfaitement illusoire de procéder de façon monolithique, en concevant *a priori* des modèles complexes dont on prétendrait qu'ils vont produire des comportements adéquats pour un environnement quelconque.

La RCO devra également être *pluridisciplinaire*. Autant l'observation de l'organisation des robots entre eux peut mobiliser les méthodes habituelles de l'analyse d'un SMA, autant au contraire l'étude des réactions des acteurs humains à la présence des robots demande une méthode spécifique. Or c'est là le champ de la sociologie, de l'anthropologie ou de l'éthologie humaine. Ces disciplines pourront d'ailleurs également apporter leur regard propre sur les mécanismes sociaux régissant le groupe de robots.

### **RCO et incrémentalité**

Les expériences actuelles sur des robots, qu'il s'agisse de robots sociaux (organisation d'un groupe de robots) ou sociables (un robot, rarement plus, en interaction avec un ou plusieurs partenaires humains), sont considérées en général comme des préliminaires inévitables à une intégration de groupes de robots à un milieu humain.

Elles sont vues comme un premier pas dans une démarche qui consisterait à augmenter graduellement la complexité des environnements où ces robots sont capables d'opérer. Autrement dit, il faudrait commencer par savoir concevoir des comportements adéquats pour des robots dans un environnement simple, puis peu à peu paramétrer ces comportements, les modifier pour obtenir des systèmes autonomes dans des situations de plus en plus compliquées (MATARIC 1995). Cela semble faire écho au principe d'incrémentalité que nous nous sommes fixé. Mais est-ce bien le cas ?

Le principe d'incrémentalité vise à construire des modèles informatiques de complexité croissante *pour un problème donné*, en conséquence du principe de parcimonie qui veut qu'on évite un coût cognitif injustifié.

Or ici, c'est au contraire un *même modèle* – le cas échéant déjà assez sophistiqué – que l'on cherche à adapter à un *problème de complexité croissante*,

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

ou plus exactement d'ailleurs, à un problème de moins en moins contraint. Il ne s'agit donc pas d'une démarche incrémentale au sens où nous l'entendons. Celle-ci consiste au contraire à prendre en considération le problème tel qu'il se pose (à savoir l'intégration d'un groupe de robots dans un groupe humain, dans un environnement qui ne soit pas spécifiquement modifié pour les robots) et à rechercher dans ce cadre des modèles les plus simples possibles.

La nature fondamentalement expérimentale de la RCO permet de maintenir justement une incrémentalité dans la construction des modèles de comportements. On pourra être amené, dans le cadre général de la RCO, à étudier séparément diverses problématiques qu'elle soulève, étant bien entendu que cette décomposition reste orientée par la réalisation de l'objectif global. Nous retrouvons ici un parallèle avec l'éthologie : l'éthologue, s'intéressant *in fine* au comportement naturel d'un animal, peut mener par ailleurs des expériences ciblées sur telle ou telle compétence cognitive afin d'améliorer sa *compréhension* des comportements naturels. Là où l'éthologie mène à la compréhension des comportements, la RCO conduit à raffiner les modèles de *production* de comportements.

### Mise en situation sociale

Puisque la RCO implique d'aborder de front le problème visé, dans toute sa complexité, il est bien évident que les agents qui devront contrôler les robots seront *physiquement situés*.

Ne nous arrêtons pas en si bon chemin : pour la robotique collective, la présence d'humains dans l'environnement est souvent considérée comme une contrainte supplémentaire (c'est-à-dire concrètement comme un ensemble d'obstacles mobiles imprévisibles), plutôt que comme une donnée inhérente au problème (WILKES et al. 1997). De même, la coordination et l'organisation de groupes de robots est perçue comme une difficulté gratuite pour l'interaction homme-robot.

À l'inverse, la RCO devra donc s'orienter vers une *mise en situation sociale* des robots (DAUTENHAHN 1998). Un robot *socialement situé* doit, en plus de ses capacités physiques individuelles (éviter d'obstacles, navigation, etc.) être en mesure d'interagir avec des partenaires, qu'ils soient humains ou robotiques, sur lesquels il dispose de peu d'informations *a priori*.

Un robot *socialement situé* sera confronté à une vision partielle des relations sociales existant entre les individus présents dans son environnement.

Au lieu de disposer, comme dans nombre de SMA, d'informations fiables et semi-globales (par exemple une liste d'acointances), il doit traiter des données locales et peu fiables.

La seule évocation de la mise en situation sociale ne suffit pas, bien sûr, à résoudre le problème, mais il s'agit d'abord de prendre en compte en tant qu'éléments constitutifs de l'expérience, et non en tant qu'obstacle, les contraintes qui s'exercent sur des agents immergés dans un environnement social : cela demande alors de concevoir les modèles comportementaux d'une manière très différente (ce qui est considéré comme un handicap avec lequel il faut composer devient au contraire une *donnée* de l'environnement).

### Endo- et exo-socialité

Si les systèmes multi-robots disposent parfois d'aptitudes sociales en plus d'aptitudes physiques (ARKIN 1998, par exemple), celles-ci ne concernent que les relations sociales entre robots.

C'est ce que nous appellerons l'*endo-socialité* (DROGOUL et PICAULT 1999; MOSS et DAUTENHAHN 1998), par opposition à une *exo-socialité* que constituera l'ensemble des aptitudes sociales que les robots possèdent à l'égard des humains<sup>3</sup>.

Un robot social dispose d'une forme d'endo-socialité, un robot sociable d'exo-socialité. Il restera à déterminer dans quelle mesure ces deux types de compétences diffèrent dans leurs mécanismes et si des modèles semblables peuvent être utilisés pour les produire.

### Méthodologie

L'homme n'est guère habitué à partager son environnement quotidien avec des groupes d'autres animaux sociaux<sup>4</sup> : pour permettre à un ensemble de robots de s'intégrer *en tant que groupe* à une collectivité humaine, il faut sans doute donner à leurs comportements (sociaux, sociables ou autres) un certain nombre de propriétés qui les rendent acceptables.

---

3. Cette dichotomie généralise l'opposition biologique entre comportements *intraspécifiques*, tournés vers les membres de la même espèce, et comportements *interspécifiques*, entre espèces différentes, en la transposant à des « machines », étrangères à la notion d'espèce.

4. Peu de gens vivent avec des *meutes* de chiens ; quant à avoir une fourmilière dans son jardin, ce n'est pas à proprement parler *partager* l'environnement, dans la mesure où l'interaction homme-fourmi est particulièrement pauvre...

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

On pourrait penser dans un premier temps aux travaux menés sur les expressions faciales. L'être humain projette sans cesse de l'intentionnalité sur les entités auxquelles il est confronté (DENNETT 1990), et cette projection semble facilitée par l'utilisation d'expressions faciales. Mais en même temps, les interactions entre humains sont soumises à de très fortes contraintes de *pertinence* (DESSALLES 1996b; DESSALLES 2000). Les enfreindre, c'est mettre le partenaire mal à l'aise.

Il semble assez évident qu'un acteur humain n'aura pas la même réaction avec un robot « humanoïde » et un robot « classique » (à l'aspect bien plus rudimentaire). Il peut être tenté d'interagir plus volontiers avec le premier, *parce qu'il a l'air plus proche de lui*, mais c'est peut-être avec lui que l'interaction sera la plus décevante, précisément parce qu'il lui *attribuera* plus d'intelligence que la pauvre machine ne sera capable d'en montrer.

D'AIBO, ses maîtres s'attendent plus ou moins à ce qu'il se comporte comme un chien ou un chiot, ce qui est en partie le cas grâce à un travail remarquable au niveau de la mécanique de ce robot, mais il reste comportementalement pauvre. Il est à craindre que ce soit encore pire avec les humanoïdes (pour lesquels s'ajoute une projection d'origine culturelle).

En fait, les travaux sur les expressions faciales inversent la relation de causalité : on croit que l'intérêt que portent les humains aux robots humanoïdes réside dans la ressemblance morphologique. En réalité, ce qui attire les humains dans ces robots, c'est la promesse implicite d'un *comportement* humanoïde. Aussi, les carences comportementales de robots humanoïdes semblent-elles d'autant plus insupportables qu'elles contrastent avec les attentes légitimes de l'utilisateur.

Au contraire, vis-à-vis d'un robot « informe », qui ne sera pas perçu comme autre chose qu'une *machine* (dont on n'attend, en tant que machine, que fort peu d'intelligence), un partenaire humain sera plus « tolérant » et pourra attribuer plus facilement de la signification à des comportements sans que le robot ne lui oppose un démenti flagrant.

Cette conception est d'ailleurs soutenue par les expériences menées avec des robots « classiques » (DAUTENHAHN 1997a; DAUTENHAHN 1997b; DAUTENHAHN 1999b), dans lesquelles la « vraisemblance » des comportements (imitation par exemple) est mise en avant comme critère essentiel de l'interaction.

Un tel robot est en fait plus propre à *suggérer*, à *susciter des interpré-*

#### 7.4. LA « ROBOTIQUE COLLECTIVE OUVERTE »

*tations*. Autrement dit, l'intelligence sociale réside en grande partie dans l'œil de l'observateur et acteur humain.

C'est cette voie « comportementale » de la RCO que le projet MICRobES, que nous allons décrire au chapitre suivant, se propose d'explorer.

## ◇7. L'ÉMERGENCE DE LA SOCIALITÉ ROBOTIQUE

---

# Chapitre huitième.

## Le projet MICRobES

---

*Ce chapitre sera consacré à une description du projet MICRobES, mené au LIP6 depuis le courant de l'année 1999, dans le cadre de la Robotique Collective Ouverte définie précédemment. Nous y expliquerons ses objectifs scientifiques, ses participants et les sujets de recherche qu'il suscite. Nous présenterons les moyens techniques mis en jeu, tant matériels que logiciels, et les méthodes d'investigation utilisées. Enfin, nous poserons le problème de la construction de comportements sociaux dans un tel projet.*

### 8.1 Un projet de Robotique Collective Ouverte

Le projet MICRobES<sup>1</sup>, initié en 1999 au LIP6 par Alexis DROGOUL, vise à étudier les modèles et techniques permettant, à terme, l'immersion d'un groupe de robots mobiles autonomes dans un environnement social humain (DROGOUL et PICAULT 1999). Il s'agit donc très exactement d'un premier pas vers la Robotique Collective Ouverte définie précédemment (PICAULT et DROGOUL 2000a).

#### 8.1.1 Objectifs scientifiques

L'objectif à long terme d'un projet comme MICRobES est de « *construire des artefacts capables de survivre et d'opérer intelligemment, de façon autonome et collective, dans le monde réel où nous vivons [...], tout simplement capables, avant de faire quoi que ce soit pour leurs utilisateurs, de vivre en symbiose, de façon harmonieuse, avec eux (et entre eux)*<sup>2</sup> » (DROGOUL 2000).

---

1. L'interprétation canonique de cet acronyme est « Maîtrise de l'Immersion de Collectivités de ROBots dans un Environnement Social ».

2. A. DROGOUL appelle ce type d'artefacts des « Systèmes Physiques Symbiotiques » (SPS).

## ◇8. LE PROJET MICROBES



FIG. 8.1 – Deux robots *MICRobES* dans un couloir du laboratoire : ils sont placés dans un environnement prévu pour des êtres humains et non spécifiquement pour des robots.

Ce but ultime, peut-être lointain mais pas pour autant utopique, a une conséquence importante : il n'est pas question d'assigner une quelconque fonctionnalité *a priori* aux artefacts en question. Cela risquerait d'induire un biais important dans la conception des modèles de comportement. À l'inverse, ce n'est *que* lorsque l'on sera capable d'intégrer un groupe quelconque de machines dans un environnement social humain, qu'il deviendra possible de leur donner des tâches utilitaires *en plus*. Par ailleurs, rien n'interdit de doter les robots de mécanismes d'apprentissage leur permettant d'acquérir auprès des humains un savoir-faire utile : au contraire, pareille disposition au *dressage* sera sans doute une composante importante de l'interaction sociale homme-robot.

À plus court terme, ce projet fédère un ensemble de travaux de recherche liés à la robotique sociale située, autour d'une base matérielle et logicielle partagée, constituée du groupe de robots PIONEER (fabriqués par ActivMedia) et d'un certain nombre d'outils communs et de comportements « élémentaires » (tels que la navigation, la recharge autonome, la communication et la reconnaissance de balises visuelles).

Chacun des sujets de recherche intégré dans le projet MICRobES est susceptible de concourir à l'intégration, à terme, de cette communauté de robots dans l'environnement humain où ils sont placés – en l'occurrence, il s'agit d'un étage entier du LIP6 (cf. figure 8.2) pour le site d'expérimentation principal

## 8.1. UN PROJET DE ROBOTIQUE COLLECTIVE OUVERTE

(un autre se trouve dans les locaux de France Télécom R&D, partenaire du projet).

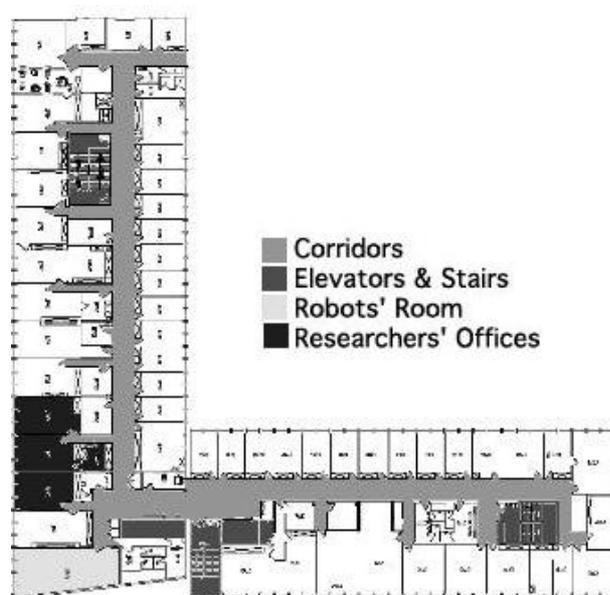


FIG. 8.2 – *La partie du laboratoire dans laquelle se déroule le projet MICRobES est un étage entier, constitué de deux longs couloirs perpendiculaires. Une salle est dédiée plus spécialement aux robots (pour réparation, recharge, etc.) mais ils circulent dans l'ensemble du site. Certaines parties de l'environnement sont d'ailleurs particulièrement dangereuses, comme les ascenseurs et les escaliers.*

Les robots disposent plus particulièrement d'une salle dans laquelle ils peuvent être réparés, se recharger, etc. mais ils peuvent circuler n'importe où (sous réserve que les portes soient ouvertes, ce qui dépend du bon vouloir des habitants du laboratoire). Certaines zones représentent un danger physique pour les robots, comme les escaliers ou les ascenseurs (cf. figure 8.5). Les seuls aménagements consentis sont : 1° l'installation de stations de recharge, indispensables à l'autonomie énergétique des robots ; 2° l'implantation de balises pour la navigation (nous reviendrons plus en détail sur ce point lorsque nous présenterons les sujets liés à la navigation).

### 8.1.2 Sujets de recherche

Les sujets de recherche qui s'articulent autour du projet MICRobES portent sur un certain nombre de compétences comportementales ou cognitives qui, pour la plupart, peuvent se décliner en termes :

- de capacités individuelles ;

## ◇8. LE PROJET MICROBES

- de capacités interactionnelles, que ce soit entre plusieurs robots ou entre un robot et un humain ;
- de capacités organisationnelles (organisation du groupe de robots ou adaptation à l'organisation humaine).

Nous allons présenter les grandes lignes des problématiques étudiées dans ce projet. Elles constituent le squelette de plusieurs thèses et de stages de DEA, sans qu'il soit cependant facile de tracer une frontière nette entre les travaux propres à chacun des membres du projet.

### Navigation et cartographie

En tant qu'individu mobile situé, la première activité d'un robot consiste à pouvoir se déplacer en évitant les obstacles. Cette étape a été rapidement franchie.

Chaque robot doit ensuite être capable de *se localiser* dans son environnement, ne serait-ce que pour pouvoir aller se recharger lorsque la borne de recharge n'est pas en vue.

Pour cela, deux approches sont possibles : *faire construire* une carte de l'environnement par le robot, ou *la lui donner* comme connaissance *a priori*. Idéalement, on se doute que la première solution est la meilleure, dans la mesure où un robot ayant des capacités de cartographie autonomes peut opérer dans un environnement quelconque.

Toutefois, la cartographie est un problème particulièrement difficile, qui fait l'objet de thèses spécifiques dans de nombreux laboratoires. Or il n'est pas possible de construire de comportement, ni individuel ni collectif, dans le cadre de la RCO, sans disposer d'un module de navigation offrant une représentation correcte de l'espace. Aussi, pour ces raisons pragmatiques, la solution finalement retenue a été de fournir aux robots une représentation *a priori* de l'environnement, doublée d'un balisage de celui-ci au moyen d'amers (en l'occurrence, des codes-barres).

La représentation spatiale retenue fait appel à la fois à un niveau global (carte topologique reliant des régions, cf. figure 8.3) permettant de naviguer d'une partie du laboratoire à une autre, et à un niveau local (« grilles d'occupation », cf. figure 8.4) en construisant à partir des perceptions une représentation dans un rayon de quelques mètres. Ces perceptions locales sont en partie mémorisées comme caractéristiques des régions traversées, ce qui permet de faire le lien entre les deux niveaux de représentation.

## 8.1. UN PROJET DE ROBOTIQUE COLLECTIVE OUVERTE

Par ailleurs, pour renforcer la fiabilité de la navigation, des balises (sous la forme de codes-barres) sont disposées dans quelques régions de l'environnement. Leur repérage et leur décodage permet aux robots de corriger les erreurs d'estimation de leur position.

Cette simplification est avant tout un choix pragmatique destiné à gagner du temps dans l'avancement du projet. Dans la mesure où d'autres travaux, menés notamment à l'AnimatLab, se penchent justement sur la question du repérage d'amers dans l'environnement (FILLIAT 1999; FILLIAT et MEYER 2000; GOURICHON 1999, au LIP6), il sera possible en temps utile de remplacer le module de cartographie actuel par un autre plus adaptatif.

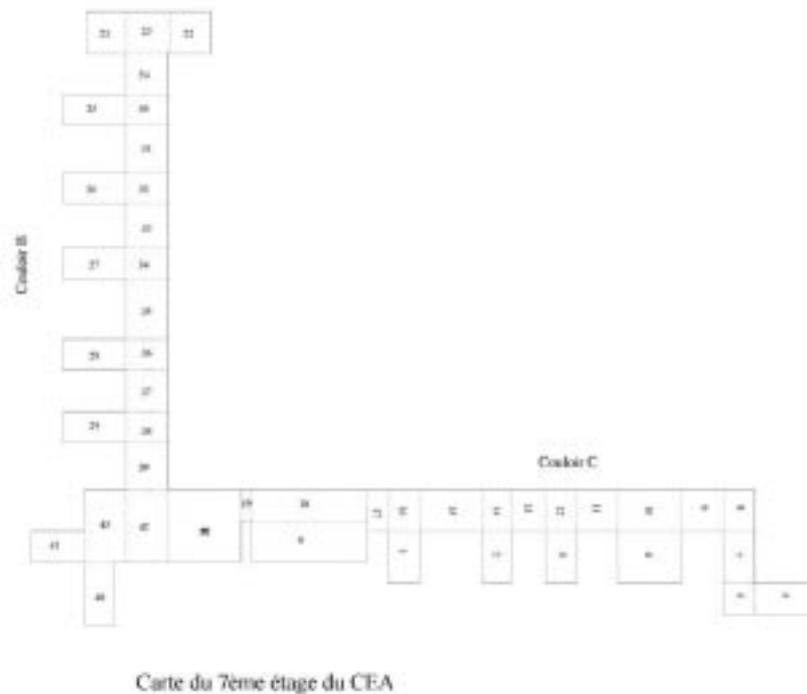


FIG. 8.3 – Carte topologique fournie au robot. Seuls les couloirs sont répertoriés ; ils sont découpés en zones. Des codes-barres placés dans certaines régions permettent une relocalisation des robots de temps à autre.

L'exploration *collective* de l'environnement constitue une question intéressante : des robots disséminés dans des lieux différents, ayant des perceptions différentes, à des moments différents, peuvent ensuite, en fusionnant leurs informations, classifier automatiquement les régions de leur environnement et faire le lien avec une représentation symbolique que pourrait leur donner un

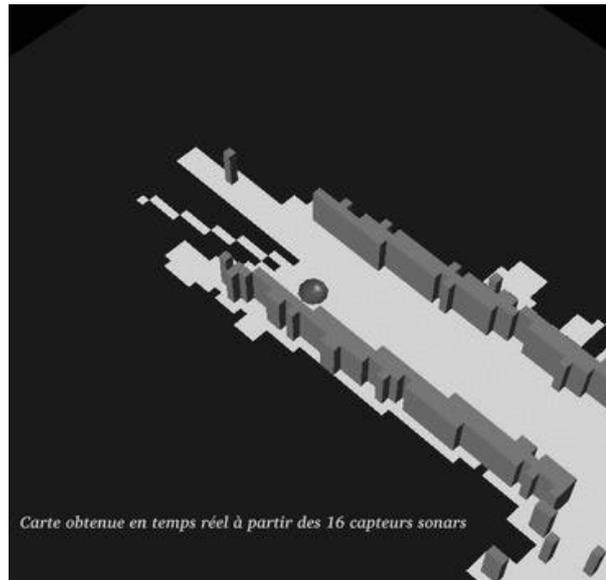


FIG. 8.4 – *Représentation construite par le robot par intégration temporelle des informations fournies par les sonars (« grille d'occupation »). Les obstacles mobiles sont ainsi éliminés au profit des obstacles fixes.* — d'après (HUGUES 2000a)

utilisateur humain (HUGUES 2000a; HUGUES 2000b; HUGUES et DROGOUL 2001a). En opérant collectivement, ils démultiplient leur expérience, réduisent les biais individuels liés au bruit des capteurs, et traitent plus rapidement l'ensemble de l'environnement.

Enfin, l'exploration de l'environnement, l'acquisition d'informations sur la localisation de telle ou telle ressource, peut servir de base à une socialité territoriale. Nous reviendrons sur ce point ultérieurement.

### Recharge autonome

Pour pouvoir être considérés comme véritablement autonomes, les robots doivent être en mesure de se recharger par eux-mêmes. Une fois la position de stations de recharge repérée (ce qui est du ressort de la navigation), il faut mettre au point des comportements d'accostage, de branchement, de surveillance du niveau de batterie et de sortie de la station.

Les bornes de recharge commercialisées par ActivMedia ne sont pas particulièrement adaptées à ce type de fonctionnement, aussi les aspects techniques de la recharge font-ils l'objet d'un partenariat avec France Télécom R&D.

## 8.1. UN PROJET DE ROBOTIQUE COLLECTIVE OUVERTE



FIG. 8.5 – *Image de gauche : les robots PIONEER sortant de leur salle, en file indienne. Image de droite : deux robots dangereusement proches d'un ascenseur.*

La recharge pose, sur le plan collectif, la question de l'attribution des ressources : il ne sera pas possible à deux robots de se recharger au même moment. Il n'est sans doute pas judicieux d'utiliser un protocole figé pour décider du robot à recharger, dans la mesure où un procédé aussi mécanique ne peut tenir compte de la répartition des activités des robots (en pratique, on peut vouloir favoriser ceux qui interagissent fréquemment avec les humains, qui explorent leur environnement, etc.).

De plus, pour un observateur humain, un protocole figé présente le désavantage de paraître trop « mécanique », trop peu « vraisemblable » en termes d'attribution d'intentionnalité. On peut donc plutôt envisager de se servir d'une organisation sociale pour réguler l'accès aux bornes de recharge, et plus généralement aux ressources dont peuvent avoir besoin les robots (ces ressources pouvant être un assistant humain).

Enfin, d'un point de vue organisationnel, des mécanismes de coopération et de coordination pour le « sauvetage » de robots en panne ou déchargés

## ◇8. LE PROJET MICROBES

sont également à l'étude (MUÑOZ-MELÉNDEZ et DROGOUL 2000; MUÑOZ-MELÉNDEZ 2001).

### **Interaction avec les humains**

Un axe important de sous-projets a trait aux interactions entre hommes et robots.

Certains travaux portent sur l'apprentissage d'informations symboliques concernant l'environnement (BREDÈCHE et ZUCKER 2000), se plaçant d'emblée à un niveau de communication langagier, au moins par mots-clefs. Les robots doivent être capables d'associer un mot prononcé par un humain à un contexte particulier, et par induction ne retenir que les associations pertinentes.

Un autre sujet prépare l'utilisation « fonctionnelle » des robots par les humains (comme entités autonomes et non simplement comme outils), et s'intéresse à la répartition de tâches demandées par plusieurs utilisateurs au groupe de robots (en l'état actuel, ces tâches restent virtuelles, le but étant de concevoir des mécanismes généraux de coopération distribuée dans un contexte ouvert). Il incombe au groupe de s'organiser pour répondre aux demandes, en mettant en place des délégations mutuelles de tâches ou d'actions concourant à l'exécution de ces tâches. Ces aspects sont étudiés en partenariat avec France Télécom R&D, et font appel à la notion d'« environnement social » que nous avons évoqué dans § 5.5.1.

Enfin, une dernière composante de l'interaction homme-robot consiste à doter les machines de capacités d'apprentissage par démonstration (HUGUES et DROGOUL 2001c; HUGUES et DROGOUL 2001b). Un humain qui souhaite « enseigner » un comportement au robot doit pouvoir le faire d'une façon relativement intuitive et immédiate, en « montrant » la séquence d'actions deux ou trois fois, à charge pour le robot de discriminer les éléments pertinents des scènes qu'il a observées.

Pour cela, la voie explorée actuellement est celle d'un « pilotage » du robot par un humain, constituant ainsi des exemples de comportements pendant lesquels le robot recherche des perceptions saillantes qu'il s'essayera ensuite à reproduire. Ce mécanisme permettra à terme d'ancrer la communication homme-robot dans la morphologie de ce dernier, et de faire le lien avec l'apprentissage symbolique.

### **Proprioception**

La difficulté pour le concepteur à choisir des comportements pertinents pour lier de façon réactive capteurs et effecteurs, d'une part, et le besoin de pouvoir traiter les pannes ou défections de certaines parties non vitales des robots (capteurs par exemple), d'autre part, ont suscité un axe de recherche autour des capacités de proprioception.

Il s'agit de recourir à des méthodes d'apprentissage automatique permettant de développer une architecture de contrôle pour les comportements réactifs du robot et de la réorganiser en cas de défaillance d'un capteur ou d'un effecteur (pour ne plus tenir compte d'informations erronées).

Ces travaux demandent une approche incrémentale et modulaire, de façon à pouvoir ajouter, couche après couche, diverses fonctionnalités à une architecture initiale assez simple. L'architecture ainsi construite doit pouvoir se reconfigurer « en ligne » (c'est-à-dire durant le fonctionnement du robot dans son environnement) pour s'adapter à des événements imprévus (dysfonctionnements).

Toutefois, une grosse part de la sélection des architectures adéquates doit être réalisée en simulation, dans la mesure où les échelles de temps manipulées sont considérables par rapport aux durées d'expérimentations et à la puissance de calcul des ordinateurs embarqués à bord des robots.

La proprioception peut s'entendre également au niveau collectif, celui du groupe, pour lequel elle consiste à maintenir une structure spatiale ou fonctionnelle, y compris en présence d'obstacles ou de dysfonctionnements (de tout ou partie d'un robot).

L'évolution darwinienne, utilisée fréquemment pour le développement d'architectures de contrôle (KODJABACHIAN et MEYER 1998, par exemple), fournit un cadre prometteur pour ce type de recherches.

Les travaux commencés dans ce sens par Samuel LANDAU (LANDAU et al. 2001a) et nos propres réflexions sur les spécificités d'un projet comme MLCRobES (cf. notamment le chapitre 9) nous ont conduit avec lui à définir une approche nouvelle pour l'évolution et le développement dans les SMA situés, que nous présenterons dans les chapitres 10 et 11.

### Compétences sociales

Les recherches sur les aptitudes sociales font écho aux problématiques d'endo- et d'exo-socialité évoquées à propos de la RCO. Dans les deux cas, l'environnement social (robotique ou humain) est dynamique et des structures sociales figées ne sauraient être adaptées.

On peut donc dans un premier temps envisager de transposer les modèles de représentation de la hiérarchie sociale, étudiés précédemment, aux robots. Cette hiérarchie pourra leur servir à anticiper ou à résoudre des conflits.

Une perspective plus lointaine consiste à rechercher des modèles permettant aux agents robotiques de se représenter une partie du fonctionnement social des humains. On peut songer notamment à tout ce qui se traduit par des changements environnementaux (horaires de travail par exemple).

D'un point de vue méthodologique, la difficulté de la transposition consiste essentiellement à trouver des comportements « équivalents » aux comportements agonistiques. Nous avons vu au chapitre 6 que cela reste tout à fait envisageable lorsque l'on peut attribuer au groupe une tâche collective (que celle-ci soit explicite ou non).

En revanche, cela est sans doute plus difficile pour des robots dépourvus de fonctionnalité *a priori*. On peut s'appuyer par exemple sur les « performances » des robots (notamment en termes d'exploration, d'interaction, etc.) ou sur une construction de « territoires » associés à des ressources. Mais cela n'est pas très satisfaisant, dans la mesure où ces choix sont assez arbitraires. Nous reviendrons sur ces difficultés au § 8.4.2.

## 8.2 Acteurs et méthodes

### Participants

Le projet MICRobES est le fruit d'une étroite collaboration entre :

- des chercheurs permanents de trois équipes différentes du LIP6 (l'équipe multi-agent, MIRIAD ; l'AnimatLab ; l'équipe apprentissage, ACASA) : Alexis DROGOUL, Jean-Arcady MEYER, Jean-Daniel ZUCKER ;
- des doctorants de ces équipes : Nicolas BREDÈCHE, Louis HUGUES, Samuel LANDAU, Angélica MUÑOZ, Sébastien PICAULT, François SEMPÉ ;
- des stagiaires de DEA : Thomas MICONI, Pierre-Emmanuel VIEL ;

- des chercheurs de France Télécom R&D : Jean-Marc RAIBAUD, Laurent PONTTHOU ;
- des anthropologues et des sociologues du CSI (École des Mines de Paris) et de l'ÉHESS : Bruno LATOUR, Albert PIETTE, Laurent THÉVENOT, Fabrice ROUSSEAU.

### **Approche incrémentale et modulaire**

Quels que soient les sujets de recherche, la même approche méthodologique prévaut, basée sur des principes d'incrémentalité et de modularité. Les modèles et les techniques étudiés doivent pouvoir être implantés, testés et évalués indépendamment les uns des autres, pour juger de leur validité intrinsèque et de leur apport spécifique au comportement observable des robots pris individuellement, du groupe de robots, mais aussi des humains.

Cela est une nécessité absolue. Non seulement il est inenvisageable d'aborder un problème tel que la mise en situation physique et sociale de robots au moyen de modèles trop complexes d'emblée (car alors on n'est plus en mesure d'expliquer comment le système fonctionne), mais en outre la multiplicité des sujets abordés demande une coordination rigoureuse entre les participants pour pouvoir, à terme, intégrer plusieurs modules.

### **Méthode expérimentale et observation**

Comme dans les travaux précédents, c'est l'éthologie qui sert de cadre de référence pour l'observation, l'interprétation, l'analyse des résultats.

Les robots peuvent en effet être vus comme une *espèce sociale* partageant le même écosystème qu'un groupe d'*Homo sapiens* (PICAULT et DROGOU 2000b). Les comportements des robots ne doivent pas seulement être analysés pour eux-mêmes mais aussi en relation avec les réactions des humains qui les côtoient.

C'est en ce sens que des anthropologues et des sociologues du CSI et l'ÉHESS participent au projet. Ils visent à étudier deux points plus particulièrement :

- comment les humains occupant le laboratoire réagissent, au fil du temps, à la présence des robots. Cela permet à l'équipe informatique de corriger des comportements en tenant compte d'une analyse rigoureuse des problèmes.

## ◇8. LE PROJET MICROBES

- sur quelles bases les membres du projet élaborent et choisissent les comportements dont ils dotent les robots.

Cette analyse complètera nos propres observations et constituera un point de vue essentiel pour prétendre, ultérieurement, étendre la RCO à des publics plus larges que des chercheurs en informatique.

### 8.3 Quelques détails techniques

#### 8.3.1 Les robots

Les robots utilisés à l'heure actuelle dans le projet MICRobES sont des PIONEER 2DX, de la société américaine ActivMedia (importés par RoboSoft). Six de ces robots sont utilisés au LIP6, deux autres servent à réaliser des expériences plus spécifiques au sein de France Télécom R&D.

Ces robots, dans leur version de base (cf. figure 8.6), sont constitués d'un ordinateur PC classique (sur une carte PC104) et d'un micro-contrôleur (Siemens 88C166) qui gère les capteurs et les effecteurs. Il peut contenir jusqu'à trois batteries.

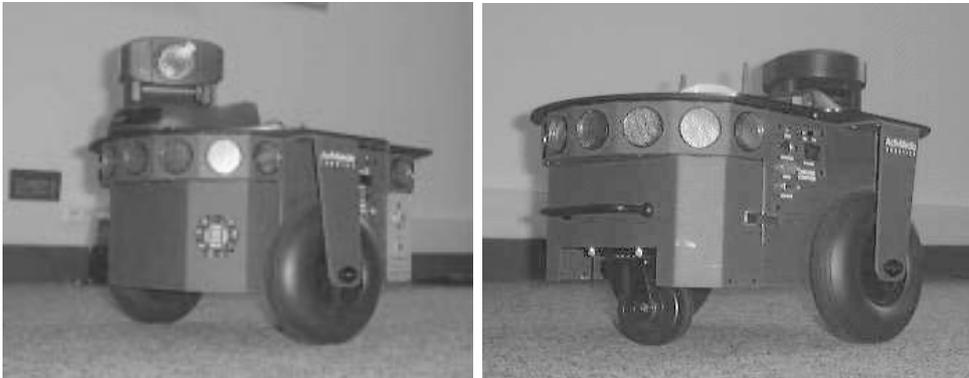


FIG. 8.6 – Vues rapprochées d'un robot PIONEER.

Chaque robot est équipé des capteurs suivants :

- 8 sonars à l'avant, 8 à l'arrière, d'une portée de quelques mètres, renvoyant une information de distance aux obstacles potentiels ;
- une ceinture de 6 « bumpers », à l'avant comme à l'arrière, qui réagissent au contact ;
- un capteur odométrique, qui mesure la distance parcourue en fonction du nombre de tours des deux roues avant ;

### 8.3. QUELQUES DÉTAILS TECHNIQUES

- un compas magnétique donnant l'orientation absolue du robot ;
- une caméra vidéo couleur ;
- un modem radio.

Quant aux effecteurs, ils se limitent aux deux roues motrices placées à l'avant et aux moteurs de la caméra (ainsi qu'au modem radio).

Une extension appelée *PeopleBot*, installée sur deux robots, consiste en une tourelle « d'interaction » placée sur la base du PIONEER, surélevant ainsi la caméra, et ajoutant 8 sonars en hauteur, un microphone et des hauts-parleurs.



FIG. 8.7 – *Un couloir du laboratoire, vu à partir de la caméra du robot, résolution maximale.*

#### 8.3.2 Système embarqué

Le micro-contrôleur possède son propre système d'exploitation (P2OS) et met à jour les paquets en provenance des capteurs ou à destination des effecteurs toutes les 100 ms.

Quant au PC, il fonctionne sous Linux et s'interface avec le micro-contrôleur au moyen d'une bibliothèque écrite en C par ActivMedia : SAPHIRA. SAPHIRA a un fonctionnement client/serveur qui permet de faire fonctionner les mêmes comportements soit sur un robot, soit sur un simulateur. Louis HUGUES s'est livré, sur cette base, à un énorme travail d'encapsulation des comportements en C++ qui permet de disposer, pour la programmation des robots, d'une interface orientée objets dans laquelle les principaux outils de SAPHIRA sont plus facilement manipulables.

## ◇8. LE PROJET MICROBES

Ainsi, le système de contrôle du robot fonctionne actuellement au moyen d'une boucle de 100 ms gérée par SAPHIRA, pendant laquelle peuvent être exécutés un certain nombre de « micro-tâches ». Le rôle donné à ces micro-tâches est parfaitement libre, toutefois il faut autant que possible se limiter à des calculs légers pour rester dans la limite totale des 100 ms.

Ces micro-tâches servent en général à contrôler l'activité de primitives comportementales (appelées *Behaviors*) et de *threads* lancés en parallèle.

### 8.3.3 Outils communs

Le moteur de sélection d'action EMF (DROGOUL 1993) a été adapté aux contraintes logicielles des robots. Les *tâches* s'écrivent toujours comme chaînage de primitives, qui en l'occurrence correspondent désormais à des *behaviors*. Ces tâches peuvent être définies au moyen d'un fichier de configuration écrit en XML. Il fait appel à des micro-tâches qui ont pour rôle respectif de :

1. mettre à jour les senseurs ;
2. calculer des stimuli ou des préconditions ;
3. sélectionner une tâche
4. exécuter une primitive ;
5. mettre à jour les actuateurs.

Actuellement, tous les robots ont été dotés d'une « base comportementale » minimale leur permettant de fonctionner de façon autonome : parcours non dirigé avec évitement d'obstacles, navigation de région en région, utilisation des balises pour la relocalisation, traçage des comportements et des perceptions, reconnaissance mutuelle des robots au moyen de balises, communication radio entre robots, accostage de la station de recharge.

Ces comportements et ces modules plus ou moins cognitifs peuvent servir de point de départ pour l'ensemble des travaux gravitant autour du projet.

## 8.4 Difficultés rencontrées

### 8.4.1 Problèmes d'origine techniques et implications

L'utilisation de vrais robots, pourtant vendus « clef en main », n'est pas sans poser quelques problèmes, d'origine technique, mais dont certains ne sont pas pour autant dépourvus de conséquences plus abstraites.

### Capteurs et effecteurs

En premier lieu, on peut mentionner la faible fiabilité des capteurs. Le compas, par exemple, affiche des variations pouvant atteindre  $30^\circ$  à  $40^\circ$  pour une même direction. La mesure de l'odométrie dérive au cours du déplacement des robots, si bien qu'il leur est nécessaire de se relocaliser au moyens d'amers visuels tels que des balises placées dans l'environnement.

Ces problèmes sont inhérents à la nature même des robots et impliquent justement de travailler sur des informations locales, partiellement fiables seulement. Toutefois, il existe divers degrés de fiabilité et de précision dans les dispositifs techniques utilisables comme capteurs et effecteurs. Par exemple, un télémètre laser sera infiniment plus précis que des sonars, et un GPS résoudra immédiatement la question de la localisation.

Cela amène donc à un problème qui n'est pas anodin et qui reste difficile à trancher : dans quelle mesure les comportements construits *sous certaines contraintes techniques* ne risquent-ils pas de perdre tout intérêt avec la diffusion d'un nouveau dispositif électronique ? Autrement dit, le point crucial consiste à identifier des problématiques générales, relativement indépendantes des moyens techniques mis en œuvre, mais cela n'a rien d'évident avec des robots.

### Robustesse

Une autre source de difficulté réside simplement dans le manque de robustesse des robots (de manière très générale, et quels que soient les constructeurs), qui conduit à une succession de pannes et dysfonctionnements en tous genres. De plus certains accessoires (station de recharge par exemple) ne sont pas adaptés à un usage autonome des robots.

Ces obstacles sont actuellement surmontés (depuis peu de temps) mais ils ont induit un retard considérable dans le lancement des expériences. Dans la mesure où la base commune à l'ensemble des sujets de recherche est maintenant implémentée et testée, des expériences plus ciblées vont pouvoir être menées dans de meilleures conditions.

### Interdépendance des comportements

Toutefois, au cours de la mise en place des comportements « de base » des robots, ainsi que lors d'expériences préliminaires, nous avons constaté une

forte interdépendance entre les comportements des robots, ainsi qu'entre ces comportements et des particularités physiques des machines par exemple, pour une activité aussi « élémentaire » que la navigation réactive, il a parfois été nécessaire de se livrer à de patients jeux d'essais pour vérifier que tel ou tel mouvement ne risquait pas de bloquer le robot sous un meuble, ou pour équilibrer le poids de deux comportements (évitement d'obstacles et exploration).

Ces problèmes sont peut-être de nature à remettre en question la possibilité réelle d'une conception modulaire des comportements. Ils conduisent à paramétrer très finement des comportements dont on attendrait au contraire une certaine généralité. En tout cas, ils posent une question à laquelle nous allons tenter de répondre au chapitre suivant : dans quelle mesure et par quels moyens le substrat physique du robot doit-il être pris en compte dans la conception de ses comportements ?

## 8.4.2 Transposition, simulation et réel

### Transposition

Nous avons évoqué, à propos des sujets de recherche sur les mécanismes sociaux pour les robots, les difficultés soulevées par la transposition de modèles utilisés pour la simulation.

La transposition d'un modèle s'appliquant aux rôles relationnels pose le problème de l'*équivalence fonctionnelle* des comportements, comme nous l'avons vu au chapitre 6. Dans le cas des robots, on peut envisager d'utiliser des mécanismes sociaux pour réguler l'accès aux stations de recharge, l'exploration et la navigation dans l'environnement (avec une notion de territorialité par exemple), etc. On constate d'emblée la grande variété des situations où la socialité peut être *utilisée*. Le problème qui se pose alors est, une fois choisi l'usage de la socialité, de déterminer *comment produire* des représentations sociales en rapport avec les activités des agents : autrement dit, établir une équivalence entre les dépendances existant entre les rôles individuels des agents.

Ce problème ne se pose pas en simulation, et ce pour deux raisons : 1<sup>o</sup> les comportements jugés pertinents au regard des objectifs de la simulation sont peu nombreux, il est donc possible d'explicitier ces dépendances et de les implémenter ; 2<sup>o</sup> la *manière* dont les comportements relationnels s'ancrent dans les comportements individuels (chez les entités réelles qu'on cherche à simuler) n'a pas forcément de pertinence et l'on peut donc en *faire abstraction* lors de la démarche de simulation.

En revanche, cela n'est plus vrai lorsqu'on cherche à *produire* des comportements de robots. La réalisation d'une *fonction* passe par un corps physique, par l'utilisation de comportements qui peuvent être construits arbitrairement (à la fois en qualité et en quantité). Il devient impossible de se cantonner dans le niveau d'abstraction qui était celui de la simulation, il est indispensable de décider des moindres détails du comportement.

### Du « cercle vertueux » au cercle vicieux

Une solution qui peut sembler assez naturelle consisterait à procéder par expériences successives des divers modèles, et à juger de leur pertinence vis-à-vis du fonctionnement collectif des robots.

Cela, bien entendu, est rendu quasiment impossible par les durées d'expérimentation nécessaires. Pour tester les variantes des modèles envisagés, il faut passer par la simulation. Mais alors, on perd tout l'intérêt d'une mise en situation et l'on se retrouve dans un cercle vicieux. En effet, il faudrait disposer d'un simulateur dont on soit certain qu'il reproduise tous les éléments pertinents de l'environnement réel, et nous avons vu ce qu'il en était.

Cet obstacle que nous avons rencontré lors de tentatives pour transposer nos modèles fait apparaître deux points importants :

- 1° un élément constitutif indispensable des modèles à utiliser pour conserver une approche située : la *corporéité* du robot, i.e. sa nature physique ;
- 2° un problème méthodologique lié à la simulation : puisqu'il n'y a aucun sens à passer par la simulation pour valider un *modèle de comportement* situé, il faut définir un modèle qui soit ne nécessite pas de simulation (adaptation « en ligne »), soit recourt à la simulation à un niveau « méta », par exemple pour valider les capacités d'*adaptation* du modèle et non ses caractéristiques en elles-mêmes.

Nous allons approfondir ces questions dans les chapitres qui suivent.

◇8. LE PROJET MICROBES

---

## Chapitre neuvième.

# Construire une socialité située

---

*« En essayant d'imiter l'esprit humain adulte, il va nous falloir beaucoup réfléchir au processus qui l'a amené à l'état où il se trouve. [...] Au lieu de produire un programme qui simule l'esprit de l'adulte, pourquoi ne pas essayer plutôt d'en produire un qui simule celui de l'enfant? [...] Il y a un lien évident entre ce processus et l'évolution, à travers les identités suivantes :*

*structure de la machine-enfant = matériel héréditaire ;  
changement dans la machine-enfant = mutations ;  
sélection naturelle = jugement de l'expérimentateur »*

ALAN TURING. *Computing Machinery and Intelligence*, traduit dans (TURING et GIRARD 1995)

La notion de mise en situation, indispensable pour concevoir des comportements de robots autonomes dans un environnement complexe, conduit progressivement au concept de « corporéité », que nous allons définir, et qui s'avère, du moins à moyen ou long terme, indispensable à la construction de robots sociables et sociaux. Par ailleurs, il nous apparaît également nécessaire d'étudier la socialité d'un groupe de robots dans une perspective évolutionniste, de façon à mettre en correspondance la fonction de la socialité et les mécanismes qui l'instaurent. Ces deux approches impliquent des révisions conceptuelles et méthodologiques importantes que nous expliquerons.

## 9.1 Corporéité et socialité

Nous avons souligné à propos du projet MICRobES le fait que les comportements effectifs des robots dans leur environnement ne sont pas complètement indépendants de leurs particularités physiques. Cette problématique n'est pas propre à MICRobES, elle affecte plus généralement toutes les recherches portant sur la conception de comportements adaptatifs (PFEIFER 2000).

Aussi, la première piste alternative que nous examinerons pour la construction de robots sociaux fait appel à la notion de *corps*. Ce terme apparemment sans ambiguïté lorsqu'il s'applique aux êtres vivants, se révèle problématique pour des machines constituées d'un assemblage de composants.

Nous allons tenter de définir ce que peut être le « corps » d'un robot, et déterminer le rôle clef qu'il peut jouer dans le développement de la socialité (qu'il s'agisse en fait d'endo-socialité ou d'exo-socialité).

### 9.1.1 De la mise en situation à la corporéité

#### Le corps d'un robot ?

Le terme « corps », qui semble si familier, regroupe pourtant déjà trois définitions :

- 1° la structure matérielle et la substance d'un *organisme* – le corps d'un être vivant tout particulièrement ;
- 2° un regroupement fermé de matière – c'est la définition utilisée en physique ;
- 3° un ensemble d'individus considéré comme une entité – par exemple le corps législatif ou un corps d'armée.

Un robot possède bien évidemment un corps au sens 2°, puisqu'il est constitué d'un assemblage de pièces matérielles. Selon cette définition, le robot est *clos* au sens mathématique (CARDON 2001a).

Jusqu'à présent, nous nous sommes intéressé à la construction de modèles d'agents permettant de produire un comportement *spécifié* par le concepteur, et plus particulièrement un comportement à visée sociale.

Pour cela, nous nous sommes appuyé métaphoriquement sur des comportements animaux. Or, un animal possède un corps au sens 1°. Tant que l'on se restreint à la conception d'agents logiciels, pour lesquels l'existence même du

« corps » selon l'un des sens cités fait débat<sup>1</sup>, la métaphore des comportements animaux ne pose pas vraiment problème : on est en fait contraint dans ce cas à *abstraire* les comportements de leurs racines corporelles et à transposer une *fonction*.

En revanche, dès que l'on travaille sur des robots, l'existence indéniable du corps au sens physique met en péril la pertinence même de la métaphore : est-il judicieux de prétendre *transposer* des comportements « semblables » d'un individu *observé*, doté d'un corps *biologique*, à un individu *construit*, doté « seulement » d'un corps *physique* ? Quand l'agent auquel on cherche à transposer un modèle n'a pas de corps, les comportements s'appuient sur des capacités de perception et d'action « idéales ». À l'opposé, lorsqu'on transpose un modèle à un agent possédant un corps (physique), les capacités de perception et d'action réelles passent par un médium dont les propriétés doivent être prises en considération, et leur immersion dans un environnement naturel les rend infiniment plus complexes que leur formulation idéale.

Si cette question peut sembler dans un premier temps accessoire en IA, où l'hypothèse historique affirme que l'intelligence est d'abord affaire d'algorithmes et non de substrat, elle prend plus de force si l'on songe que l'éthologie, quant à elle, s'appuie en dernier ressort sur les ressemblances ou les différences corporelles pour comparer les comportements – puisque celles-là font partie des causes efficientes de ceux-ci. Un autre exemple « inquiétant » nous vient des « tortues » construits dans les années 1940 par le neurophysiologiste Grey WALTER : il s'agissait de robots reposant sur un corps physique sophistiqué contrôlé par *deux transistors* seulement, et néanmoins capables de navigation réactive, d'évitement d'obstacles, de recharge autonome, autant de comportements qui ne sont pas si faciles à obtenir sur des robots *informatisés*. Bien malin, alors, qui pourrait « extraire » d'une telle machine la « fonction » pour la reproduire sur un robot moderne : elle ne réside pas tant dans le câblage des transistors que dans la forme et les propriétés mécaniques de la machine.

Ainsi, il paraît finalement nécessaire de poser au moins le problème et d'étudier l'importance que peut revêtir la part du corps dans la production de comportements.

---

1. La notion de corporéité « au sens large » ne repose pas spécifiquement sur un substrat matériel robotique mais plutôt sur un couplage fort entre des dispositifs de perception et d'action dans un environnement complexe et des structures de contrôle du comportement (DAUTENHAHN 1999a; CARDON 2001a).

### Corporéité et situation physique

La notion de *mise en situation*, centrale dans le projet MICRobES, suppose déjà que l'environnement façonne les comportements possibles des agents.

Si l'on poursuit cette idée jusqu'à son terme, cela implique que les architectures logicielles qui contrôlent les actions et les perceptions d'un robot doivent s'inscrire *dans le substrat physique du robot*, dans son corps, pour produire des comportements adéquats.

Cette *inscription dans le corps physique* n'est rien d'autre ici que la mise en situation dans un environnement « élargi », qui en plus d'un milieu physique extérieur au robot, comprend également la machinerie mécanique et électronique qui le constitue. C'est en ce sens que se sont développés quelques projets autour de la construction située de comportements de robots humanoïdes, comme COG (BROOKS 1996).

Ce lien fondamental entre la corporéité et la mise en situation a été défendu d'ailleurs par certains courants récents de la théorie de l'action sociale :

« *La perception humaine, en tant qu'elle est fondée dans la motricité intentionnelle du corps, se distingue [...] de l'image traditionnelle de la perception, telle qu'elle a été entretenue depuis Descartes jusqu'à l'« intelligence artificielle ».* [...] *La supériorité de la perception humaine sur les techniques par lesquelles on tente de la simuler ne résulte pas d'une meilleure « qualité » de l'esprit, qui ne pourrait être égalée par de tels programmes, mais de l'ancrage de la perception dans la corporéité de l'homme.* » (JOAS 1999)

De même que la prise en compte de la situation physique d'un agent ne suffit pas à résoudre ses problèmes comportementaux, mais permet au moins de guider la conception de comportements adéquats en pointant les aspects fondamentalement locaux, imprécis et dynamiques des perceptions, de même la prise en compte de la constitution corporelle d'un robot doit rajouter des contraintes sur le type de comportements que l'on peut raisonnablement espérer mettre en œuvre dans un robot.

Cela n'est pas sans dangers : concevoir un modèle d'agent capable de tenir compte d'un corps physique donné risque de conduire à un modèle *ad hoc*, inutilisable sur d'autres robots ou dans d'autres conditions d'utilisation. On pressent déjà qu'il faudra se doter d'outils méthodologiques capables d'éviter pareil problème.

### Corporéité et situation sociale

Le corps joue également un rôle central dans la constitution du lien social : puisque le corps, chez les animaux, est le médium obligé des comportements, y compris des comportements sociaux, il est compréhensible que la transposition directe de modèles d'interaction sociale utilisés pour simuler des groupes de primates conduise à de sérieux obstacles (cf. § 8.4.2). La simulation permet de faire abstraction de la corporéité en se focalisant sur la fonction du comportement, implémentée immédiatement ; mais l'implantation de ces comportements sur des robots doit en revanche passer nécessairement par un corps.

Nous avons déjà mentionné les travaux portant sur la relation entre la socialité et des mécanismes liés au corps : expressions faciales, imitation, empathie, etc. (cf. § 7.3). Ces approches mettent en avant le rôle central de la corporéité dans la constitution du lien social, tant avec un partenaire humain qu'entre robots (DAUTENHAHN 1999a).

La possibilité d'une transposition reste alors ouverte, moyennant toutefois le maintien de l'hypothèse de *séparation des rôles*. En effet, tant que celle-ci est posée, le problème difficile est celui de la conception de comportements individuels équivalents entre les modèles utilisés en simulation et ceux utilisés dans d'autres applications (comme nous l'avons vu dans la partie I). Dans ce cas, trouver l'équivalent d'un comportement d'agression permet de retrouver les dépendances entre rôles individuels, sur lesquelles s'appuient les rôles relationnels.

En revanche, si cette hypothèse doit être abandonnée, c'en est fini de la dynamique du « cercle vertueux » classique ! Il reste certes possible d'espérer transposer des comportements individuels<sup>2</sup>, mais ce n'est plus le cas pour les comportements relationnels et organisationnels, dans la mesure où ceux-ci deviennent complètement dépendants de la corporéité de l'agent.

Or, le problème qui se pose lorsqu'on passe d'agents logiciels à des agents robotiques est principalement la double impossibilité de contrôle par le concepteur d'une part, des capacités de perception et d'action du robot, et d'autre part de la dynamique de l'environnement. Concevoir des agents logiciels laisse une marge de manœuvre considérable dans le choix du modèle de perception et d'action des agents, avant même de se pencher sur leurs comportements.

---

2. Cela est même envisageable d'un corps robotique à un autre si l'on est en mesure de déterminer une fonction de transfert entre les deux espaces sensori-moteurs impliqués : cf. par exemple (KAPLAN et OUDEYER 2000; BRAZIER et al. 2001).

## ◇9. CONSTRUIRE UNE SOCIALITÉ SITUÉE

Plus exactement, il est en partie possible de décider de certaines capacités de perception et d'action *en vue de comportements spécifiés*, voire *en vue d'un environnement donné* (même complexe).

À l'opposé, cela n'est plus possible sur un robot, dont les perceptions et les actions sont imposées à la fois par la nature, le nombre, les caractéristiques physiques des capteurs et des effecteurs, et par la nature de l'environnement.

Dès lors, rien ne garantit qu'un modèle simplement transposé d'un agent logiciel dans un agent robotique produise un comportement qui ait un sens. L'idéal serait que l'agent robotique apprenne à « manipuler » son corps physique et à *donner un sens* aux événements qu'il peut percevoir<sup>3</sup>.

En somme, il s'agit là d'un problème d'*ontogenèse* : pour agir, et notamment pour agir socialement, il est nécessaire de *se* construire en tant qu'individu situé dans un environnement physique et social.

### 9.1.2 Comment construire un corps ?

Il existe plusieurs façons d'aborder le problème de l'ontogenèse, c'est-à-dire en l'occurrence d'une représentation du corps manipulable pour percevoir et agir dans l'environnement. Nous distinguerons principalement les approches mono-agents des approches multi-agents.

#### **Le développement mono-agent**

Dans cette première approche, chaque robot (l'individu qui se développe) est contrôlé par un agent. L'ontogenèse de ses comportements, c'est-à-dire leur développement au cours de la « vie » du robot, peut se formuler en termes d'*apprentissage*.

L'agent va devoir construire une représentation du corps robotique qu'il habite, basée sur un couplage fort entre perceptions/actions et comportements. Ce couplage rend l'agent « autobiographique » (DAUTENHAHN 1997b; DAUTENHAHN 1999a) c'est-à-dire qu'il dispose d'une trace sensori-motrice de son expérience dans le monde où il est immergé et peut la « revivre » virtuellement.

---

3. Cette capacité à interpréter un flot d'informations est déjà difficile à réaliser lorsqu'on a affaire à un monde symbolique : cf. les travaux sur la reconnaissance et la constitution de chroniques (CARLE et al. 1998; SANSONNET 1999) ou l'analyse des interactions en situation de crise (LESAGE 2000).

Sous diverses variantes, ces mécanismes permettent une certaine forme d'empathie, d'adaptation à l'environnement, mais aussi d'apprentissage de comportements à partir de démonstrations réalisées par un expérimentateur humain (HUGUES et DROGOUL 2001c; HUGUES et DROGOUL 2001b).

Les robots ainsi constitués peuvent être indifféremment sociaux ou sociaux : en fait, selon ce point de vue, ils seront tels que l'environnement dans lequel ils ont été immergés les aura façonnés. La socialité en tant que telle cesse alors d'être une propriété du modèle de comportement pour devenir une caractéristique de l'environnement.

### Le développement multi-agent

La seconde approche fait appel, pour contrôler le robot et lui donner une corporéité, à un *système multi-agent* et non plus à un unique agent.

Il s'agit donc de construire un SMA fortement adaptatif, puisque c'est lui qui va servir à la fois de représentation du corps et d'« individu ».

On peut particulièrement mentionner le modèle proposé par Alain CARDON (CARDON 1999; CARDON 2001a), dans lequel le SMA se compose de deux sous-systèmes étroitement couplés, l'un affecté à la manipulation des données provenant des capteurs et envoyées aux effecteurs, et l'autre construisant une représentation de l'activité du premier dans un espace morphologique. Cette capacité d'auto-observation et d'auto-adaptation du système lui donnerait alors une forme de « conscience » de son propre corps.

Le problème qui reste posé est celui de la *construction* d'un tel système. On peut de nouveau chercher des métaphores en biologie, par exemple dans la construction du corps des êtres pluricellulaires.

Tout être vivant est construit à partir d'une cellule unique contenant, entre autres, une molécule d'ADN, considérée comme le support de l'hérédité. L'embryologie classique est fortement finaliste : elle conçoit l'ADN comme un « programme » conduisant *nécessairement* (sauf « accident ») à la formation d'un individu, par des mécanismes effroyablement compliqués d'envoi de messages entre cellules pour induire des différenciations aux lieux et moments adéquats.

Outre qu'elle ne résout en rien notre problème (puisque'il faudrait être capable d'écrire un programme de développement d'un système distribué, alors que c'est précisément la difficulté), cette conception de l'ADN comme programme a été récemment beaucoup critiquée (ATLAN 1999; KUPIEC et SONIGO

2000). La métaphore du programme a été empruntée abusivement à l'informatique en dépit du fait que l'ADN est une molécule inerte : poser l'ADN comme structure de données aurait plus de sens. En pratique, les théories les plus prometteuses formulent plutôt ces modèles en termes de systèmes complexes (KUPIEC et SONIGO 2000). Nous reviendrons dans les chapitres suivants sur les modèles informatiques qui peuvent s'inspirer de cette approche.

### La robotique évolutionniste

Une autre voie, qui ne fait pas nécessairement appel au concept d'agent, permet de prendre en compte la notion de corporéité : il s'agit de la robotique « évolutionniste » (NOLFI et FLOREANO 2000), qui fait appel à des mécanismes darwiniens pour produire des comportements adaptés ou adaptatifs<sup>4</sup>.

Sous sa forme extrême, la robotique évolutionniste consisterait à faire évoluer les corps physiques des robots de façon à ce que leur morphologie elle-même s'adapte aux contraintes de l'environnement. Cette voie a donné des résultats intéressants en simulation, sur des créatures virtuelles (SIMS 1995, par exemple). Pour des robots réels, elle n'en est qu'à ses débuts (HARA et PFEIFER 2000; EGGENBERGER 1997; EGGENBERGER 1999) et un passage par la simulation reste généralement nécessaire.

En revanche, il est tout à fait possible de faire évoluer une architecture de contrôle en parfaite adéquation avec les dispositifs de perception et d'action. Cette architecture peut être matérielle, comme un circuit reprogrammable (HIGUCHI et al. 1993), ou logicielle, par exemple un réseau de neurones<sup>5</sup> (ACKLEY et LITTMAN 1991; NOLFI et al. 1994; KODJABACHIAN et MEYER 1998; FLOREANO et URZELAI 2000). Le problème de l'ontogenèse, c'est-à-dire du développement de comportements adaptatifs vis-à-vis de l'environnement, consiste alors à trouver une méthode de sélection pertinente.

Nous allons étudier plus précisément le rôle que peut jouer l'évolution darwinienne dans la construction de comportements sociaux.

---

4. Nous donnerons plus de détails sur les mécanismes de l'évolution darwinienne dans la section 9.2.

5. Les réseaux de neurones sont fréquemment utilisés, dans la mesure où ils permettent au robot de continuer à apprendre après que le réseau a été construit.

## 9.2 Évolution et socialité

L'utilisation de méthodes évolutionnistes constitue l'autre principale approche du problème de la socialité. Nous allons rappeler brièvement les principes de la théorie de l'Évolution par la sélection naturelle, formulée par Charles DARWIN en 1859 (DARWIN 1859) ainsi que le rôle qu'elle joue dans la constitution des sociétés animales ; puis nous verrons comment les mêmes principes ont été réutilisés en informatique.

### 9.2.1 L'Évolution par sélection darwinienne

#### Principes de la sélection darwinienne

La théorie darwinienne de l'Évolution des espèces s'appuie sur le fait qu'un environnement naturel donné fournit une quantité de ressources *limitée*, donc ne peut assurer la survie d'un nombre indéfini d'individus. Par suite, ne peuvent survivre et se reproduire que ceux qui sont les plus adaptés à leur environnement (à un moment donné) : c'est le principe de la *sélection*.

La seconde observation est la ressemblance entre enfants et parents (l'hérédité) : des traits communs sont transmis d'une génération à l'autre, parmi lesquels ceux qui ont permis aux parents d'être adaptés à leur environnement. Mais les enfants ne sont pas des copies conformes des parents : il y a *variation*. Du fait de ces variations, certains individus seront plus adaptés que d'autres et par suite favorisés par la sélection.

L'hypothèse sous-jacente à la théorie darwinienne est l'*indépendance* entre les causes des variations et celles de la sélection : les variations sont en ce sens « aléatoires », sans rapport avec les contraintes qui vont déterminer le degré d'adaptation d'un individu. DARWIN avait en effet noté que les éleveurs, qui pratiquent une sélection artificielle en croisant les animaux pour obtenir certains caractères, ne parviennent pas à susciter des variations correspondant aux critères désirés : ils ne peuvent que choisir des traits intéressants portés par certain individus parmi des variations aléatoires. L'extrapolation à une sélection « naturelle » dans laquelle les contraintes de l'environnement favorisent la propagation de certains caractères aux dépens d'autres constitue l'idée centrale de la théorie darwinienne (DARWIN 1859).

Le processus ainsi produit peut être vu alors comme un algorithme très général capable de produire de l'ordre, de la forme, des comportements de

## ◇9. CONSTRUIRE UNE SOCIALITÉ SITUÉE

façon graduelle à partir de désordre, d'aléatoire (DENNETT 1996).

### Évolution, socialité et éthologie

La sélection darwinienne fournit un cadre de pensée extrêmement intéressant, notamment parce qu'elle a pour effet, sous certaines conditions (DAWKINS 1982, pour une discussion précise sur l'adaptationnisme), de produire des entités *adaptées à leur environnement*, et ce par un procédé absolument mécanique.

C'est un moyen de donner une interprétation intentionnelle d'un phénomène dépourvu de toute téléologie. Nous avons vu (§ 2.2.1) que l'éthologie s'appuie précisément sur ce type de démarche et qu'elle fait appel à la sélection naturelle pour passer du mécanisme à la fonction : la fonction d'un comportement correspond précisément à sa valeur, son rôle vis-à-vis de la sélection.

On comprend assez aisément comment la téléonomie apparaît : bien que les animaux n'aient pas pour *but* explicite d'être adaptés à leur environnement, celui-ci agit comme un filtre qui favorise les individus qui, de fait, sont adaptés. Le même raisonnement peut se poursuivre au niveau « méta » avec les capacités à s'adapter durant la vie de l'animal (DENNETT 1996).

Qu'en est-il de la socialité ? L'adaptation produite par la sélection naturelle vaut pour l'environnement aussi bien physique que social. Autrement dit, plutôt que de concevoir des modèles de comportements sociaux et de les transposer laborieusement sur des robots, il serait plus intéressant de *produire des comportements sociaux par évolution*, en sélectionnant des *générations* successives de robots.

L'intérêt majeur d'une telle approche est de réduire considérablement la part d'arbitraire qui réside dans le choix de telle ou telle métaphore, de tel ou tel modèle, dans la transposition d'un modèle à une application, etc. En effet, la dynamique sélectionniste a pour effet de produire des adaptations *graduelles* à l'environnement. Elle joue donc exactement le rôle d'un *principe de parcimonie*, mais fait le travail en lieu et place d'un concepteur, et avec plus de généralité, puisqu'il s'agit d'un même processus à l'œuvre dans des situations différentes.

En revanche, si le résultat produit est *nécessaire* en ce qu'il reflète les contraintes qui l'ont façonné, il est également *accidentel* dans la mesure où l'on n'est jamais certain d'obtenir deux fois le même résultat. Toutefois, cette tendance à l'indéterminisme des processus évolutionnistes, qui rappelle celle

des SMA (cf. § 2.1.2), n'a rien de systématique : on observe de nombreux phénomènes d'évolution convergente dans la nature (c'est-à-dire le fait deux « solutions » morphologiques ou comportementales semblables soient trouvées indépendamment par des espèces différentes), ce qui laisse à penser qu'il existe en fait des attracteurs et des « passages obligés » (DENNETT 1996).

### 9.2.2 Évolution et SMA

Puisque le processus de sélection naturelle décrit par DARWIN est un algorithme (DENNETT 1996), il a été utilisé assez tôt en informatique (HOLLAND 1975) sous diverses variantes.

#### Principes d'un algorithme évolutionniste

Tout « algorithme évolutionniste » repose sur les principes suivants<sup>6</sup> :

- 1° Une population d'« individus » est placée dans un « environnement ». Ces individus disposent d'une structure de données héréditaire appelée « génotype » par analogie avec les chromosomes des êtres vivants. Ils ne sont pas nécessairement (voire rarement) des agents et représentent souvent des solutions potentielles à un problème, celui-ci pouvant assez fréquemment être formulé en termes de recherche d'extréma d'une fonction. Quant à leur environnement, il peut s'agir de la formulation fonctionnelle d'un problème à résoudre, d'un milieu simulé ou artificiel ou du monde physique (par exemple pour une population de robots).
- 2° Le « comportement » de ces individus, appelé « phénotype », est construit à partir du génotype. Cela peut être l'ensemble de paramètres candidat à la résolution d'un problème, ou un comportement au sens propre (pour des agents ou des animats).
- 3° L'adaptation des individus à leur milieu est évaluée soit *explicitement* (au moyen d'une fonction de notation ou fonction de *fitness*), soit *implicitement* (par exemple selon la capacité des individus à « survivre » ou à se reproduire avec d'autres agents).
- 4° Les individus se « reproduisent » en mêlant leurs génotypes pour construire de nouveaux individus, tandis que d'autres sont retirés de

---

6. Il ne s'agit là que d'une description générale de n'importe quel algorithme évolutionniste. De plus amples informations peuvent être trouvées dans (HEITKÖTTER et BEASLEY 1998; BEASLEY et al. 1993; MITCHELL et TAYLOR 1999; NOLFI et FLOREANO 2000; WHITLEY 1994; BÄCK et SCHWEFEL 1993, ...).

## ◇9. CONSTRUIRE UNE SOCIALITÉ SITUÉE

la population (le succès reproductif comme l'élimination peuvent être explicitement calculés à partir d'une évaluation explicite des individus, ou au contraire résulter d'une évaluation implicite de leur adaptation).

5° On répète ce processus avec la nouvelle population.

On comprendra que les recherches dans ce domaine portent principalement sur :

1° la *construction du comportement* ou phénotype à partir du génotype ;

2° les mécanismes d'*évaluation* explicite ou implicite de la fonction adaptative du phénotype des individus ;

3° le choix d'« opérateurs génétiques » qui vont permettre le *croisement* entre les génotypes et leur variation par *mutation*.

Sous réserve de choisir judicieusement les techniques utilisées pour chacun de ces points, les algorithmes évolutionnistes constituent des méthodes de résolution de problème particulièrement efficaces et permettent d'explorer artificiellement des problématiques de biologie. Nous examinerons dans le § 10.2.2 quelques exemples des choix retenus par les approches évolutionnistes classiques.

### Une phylogénèse de l'endo-socialité?

Nous pouvons alors envisager d'utiliser un algorithme évolutionniste pour produire des agents destinés à contrôler les robots, en vue de faire *émerger* des comportements sociaux qui soient rendus *nécessaires* par les contraintes pesant sur le groupe de robots.

La forme de la société reflèterait dans ce cas des dépendances objectives, *situées*, entre les robots, qui ne sont pas nécessairement les mêmes que celles qu'on importe en transposant un modèle existant, ni les mêmes que celles qu'un concepteur peut supposer *a priori* en concevant un modèle *ex nihilo*.

On aurait affaire alors à une production *phylogénétique* des comportements sociaux, c'est-à-dire une construction graduelle de ceux-ci par des « lignées » d'agents animant successivement des corps robotiques.

Toutefois, la construction sélectionniste d'un tel SMA semble devoir se heurter à un handicap technique. Pour qu'un algorithme évolutionniste donne des résultats intéressants, il est en effet nécessaire que la population d'individus manipulée soit relativement importante. Ce ne sera manifestement pas le cas dans un projet comme MICRobES, fonctionnant avec un faible nombre

de robots. Il faut donc travailler sur des populations d'agents beaucoup plus nombreuses, en mettant en application l'une des deux solutions suivantes :

1. Soit travailler d'abord en simulation seulement, puis implanter dans les robots des comportements dont on estime qu'ils ont atteint un degré d'adaptation suffisant. Dans ce cas, il est nécessaire de disposer d'un simulateur réaliste, sans quoi les agents obtenus peuvent être très bien adaptés à des caractéristiques propres au simulateur et absentes du réel.
2. Soit utiliser d'emblée les robots comme source de la pression sélective, en « testant » les agents par sous-groupes implantés sur les robots. Le biais précédent est alors éliminé, mais on risque de devoir faire face à des temps de calcul prohibitifs (selon la durée nécessaire pour évaluer chaque agent et la taille de la population).

L'idéal serait donc que les agents produits par le processus évolutionniste disposent de comportements non seulement adaptés, mais *adaptatifs*, de façon à passer sans trop de difficultés d'un environnement simulé à un environnement réel. On peut également, comme nous le verrons, faire agir la sélection sur une architecture multi-agent de contrôle de robot.

## 9.3 Conséquences méthodologiques

Nous avons proposé deux pistes possibles pour la construction de mécanismes d'organisation sociale en robotique ouverte :

- la corporéité, conséquence ultime de la mise en situation, qui demande un processus de *développement* (une ontogenèse) ;
- la sélection naturelle qui, en forçant les agents à s'adapter les uns aux autres de génération en génération (phylogenèse), entraîne la constitution d'une structure sociale en rapport avec les contraintes de l'environnement.

Dans ces deux approches, l'hypothèse de *séparation des rôles* est menacée : la corporéité suppose un couplage fort des rôles relationnels et organisationnels aux rôles individuels et au substrat sensori-moteur de l'agent ; quant à l'Évolution, elle opère sur l'ensemble des comportements de chaque agent, pris comme un tout, et sur les interactions entre ces comportements.

Cette hypothèse levée, *la transposition de modèles n'est plus possible*. Mais en même temps elle n'est plus *nécessaire*, puisque le modèle fait l'objet d'une *construction automatique* par les agents ou par des lignées d'agents.

## ◇9. CONSTRUIRE UNE SOCIALITÉ SITUÉE

Si l'on se reporte au tableau des schémas explicatifs de l'éthologie (tab. 2.1 page 19), on voit que nous sommes passé de l'étude des comportements tels qu'ils sont (ou, d'un point de vue de concepteur, tels que nous les souhaitons), qui oppose des causes efficientes (mécanismes) à leur interprétation téléonomique (fonction), à l'étude de l'« histoire » des comportements, c'est-à-dire la manière dont un individu peut acquérir, construire, modifier, adapter ses comportements durant sa vie (ontogenèse) d'une part, et une interprétation finale (« pourquoi » le comportement s'est-il construit, quelle est sa fonction adaptative?) à travers la phylogenèse.

Ainsi, nous gardons encore le cadre épistémologique de l'éthologie, mais nous sommes passé à un niveau différent, dans lequel le comportement est dans tout les cas construit par le système et non posé par le concepteur.

Dans le chapitre suivant, nous allons tenter de concilier les approches ontogénétique et phylogénétique au moyen d'un même concept et nous proposerons un modèle informatique spécifiquement conçu pour l'évolution et le développement multi-agents.

---

# Chapitre dixième.

## Modèles d'agents évolutionnistes

---

*« Si les atomes ont, par sort, formé tant de sortes de figures, pourquoi ne se sont-ils jamais rencontrés à faire une maison, un soulier ? Pourquoi, de même, ne croit-on pas qu'un nombre infini de lettres grecques versées emmi la place, seraient pour arriver à la contexture de l'Iliade ? »*

MONTAIGNE. *Essais*, II.XII

Nous avons vu que la conception de comportements sociaux situés peut faire appel à deux mécanismes : le développement ou ontogenèse (pour tenir compte de la corporéité des robots) et l'évolution ou phylogenèse (qui donne une raison d'être aux comportements). Nous allons rapprocher ces deux notions et rechercher des modèles d'évolution de comportements d'agents et de SMA. Les prérequis à une approche de ce genre n'étant pas assurés par les techniques évolutionnistes classiques, nous proposerons alors notre propre paradigme, l'Éthogénétique, et en donnerons une implémentation à travers le framework ATNoSFERES.

### 10.1 La dualité ontogenèse–phylogenèse

Nous proposons d'établir un pont entre les deux concepts d'ontogenèse (liée au développement) et de phylogenèse (liée à l'évolution). En effet, nous nous trouvons face à deux métaphores qui peuvent être mise en parallèle assez

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

facilement :

- Celle de **l'organisme**, c'est-à-dire d'une entité possédant un *comportement global* qui se *développe*, placée dans un *environnement* auquel elle s'adapte. La *corporéité* d'un organisme est le résultat d'un processus de couplage entre un corps physique (dispositifs situés de perception et d'action) et une structure de contrôle.
- Celle de **l'écosystème**, c'est-à-dire d'un ensemble d'entités en interaction dans un *environnement*, qui *évoluent*, doivent s'adapter les unes aux autres sous l'effet d'un ensemble de *pressions sélectives*, et par conséquent construisent une *organisation sociale*.

Nous avons déjà fait allusion à la possibilité de traiter la corporéité d'un individu au moyen de modèles multi-agents. Ce n'est pas alors les agents qui possèdent un corps, mais le système dans son ensemble.

En comparant ces deux points de vue, on voit qu'ils conduisent en fait à des représentations *duales* du même phénomène : l'immersion d'un système multi-agent adaptatif dans un environnement contraignant. Selon qu'on s'intéresse à un schéma téléonomique ou efficient, on le considérera plutôt comme un organisme qui se développe qu'un écosystème dont les entités évoluent.

De même, un schéma final s'attachera plutôt au comportement de l'ensemble du système, alors qu'une interprétation purement causale préférera décrire le même phénomène en termes d'organisation.

Ce rapprochement a été proposé également en biologie par (KUPIEC et SONIGO 2000), à travers le concept d'*ontophylogenèse* : en abandonnant la conception finaliste de l'embryologie, selon laquelle l'organisme est le *but* d'un « programme génétique » contenu dans chacune des cellules, il devient possible de considérer un individu pluri-cellulaire comme une « colonie » fortement organisée de cellules :

« *L'ontogenèse et la phylogenèse sont produites en même temps, par le même processus, que nous appelons l'ontophylogenèse. [...] Les conditions de l'environnement [structurent] la colonie (ontogenèse) et [provoquent] son évolution (phylogenèse)* » (KUPIEC et SONIGO 2000)

L'unification des concepts d'ontogenèse et de phylogenèse à travers un modèle multi-agent permet de faire d'une pierre deux coups : le SMA va constituer un *individu* capable de se développer et de se construire une corporéité, en s'appuyant sur un processus évolutif au niveau des agents qui le composent.

Pour aborder cette problématique, il nous faut donc utiliser une approche permettant *l'évolution de comportements d'agents dans un système situé*.

## 10.2 Vers une « Éthogénétique »

### 10.2.1 Principes

Les travaux que nous allons présenter ici et dans le chapitre suivant sont nés de la conjonction dans le projet MICRobES de nos propres recherches sur les comportements sociaux et de celles menées par Samuel LANDAU sur la proprioception des robots, assignant à la question de la corporéité un rôle central.

Ce que nous baptisons ici l'« Éthogénétique » (PICAULT et LANDAU 2001c) n'est rien d'autre que la définition d'une approche évolutionniste

- s'appliquant à des agents *situés* et à des *systèmes multi-agents situés* ;
- permettant à des *comportements*<sup>1</sup> d'agents d'évoluer.

Prendre en compte dans le choix des techniques évolutionnistes des spécificités du paradigme agent amène à imposer un certain nombre de prérequis que nous allons détailler.

#### 1° complexité arbitraire des comportements

Le comportement d'un agent peut être produit de multiples façons : tâches composées de primitives et déclenchées par des stimuli comme dans EMF, réseaux de neurones, systèmes de classeurs, démonstrateur de théorèmes, réseaux de Petri, etc. (FERBER 1995).

Dans tous les cas, il peut atteindre un niveau de complexité arbitrairement choisi (il est question ici non plus de complexité cognitive, mais de complexité informatique). Il doit permettre, *même pour des agents réactifs*, d'exprimer une téléonomie, c'est-à-dire de tendre vers un but (explicite ou non). Il doit également avoir une extension temporelle aussi longue que nécessaire.

Par conséquent, faire évoluer des agents consistera d'abord et surtout à faire évoluer leurs comportements, *dans toute leur complexité*. On doit pouvoir passer, de génération en génération, de comportements assez élémentaires

---

1. D'où son nom.

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

à d'autres plus sophistiqués, ou inversement des comportements inutiles compliqués doivent pouvoir se simplifier au fur et à mesure du processus évolutif, ce qui suppose en fait une certaine forme de *modularité* dans la structure produisant le comportement.

Il n'est donc pas question ici de se limiter par exemple à l'évolution de *paramètres* modifiant l'effet d'un *comportement prédéfini* : ce serait passer à côté de ce qui fait la spécificité d'un agent.

### 2° intelligibilité des comportements

Nous imposerons en outre que les comportements produits par le processus évolutif restent *immédiatement compréhensibles*. En effet, comme nous nous plaçons toujours dans une démarche expérimentale et incrémentale, il peut être nécessaire, au cours de l'évolution du système, d'analyser finement ce qui se passe, de pouvoir expliciter facilement l'ensemble des comportements des agents, de comprendre *quels comportements* sont sélectionnés et *pourquoi* (autrement dit, déterminer leur fonction adaptative).

Il sera peut-être également nécessaire d'*amorcer* le système avec des comportements donnés *a priori*, ou encore de *réutiliser* des modules.

Si jusqu'à présent nous avons toujours utilisé dans nos agents des modèles de comportement explicites et manipulables sous forme d'informations symboliques (c'est le cas d'EMF), cette contrainte est loin d'être vérifiée par toutes les architectures de sélection d'action : les réseaux de neurones, par exemple, ne s'y conforment pas.

### 3° possibilité d'une évolution « aveugle »

Les processus à l'œuvre dans la sélection darwinienne sont « aveugles » (DAWKINS 1986) au sens où les variations qui sont produites dans le génotype<sup>2</sup> sont « aléatoires » par rapport aux critères qui interviennent dans la sélection (LESTIENNE 1993) ; autrement dit, les variations ne sont *dirigées par aucune finalité*.

Au niveau moléculaire, cela se traduit par des erreurs de copie, des insertions ou délétions de séquences de bases dans l'ADN, etc., autant de méca-

---

2. On rappelle que le génotype est le support de l'information héréditaire, sujette à variations. Le phénotype est le caractère résultant, observable, par exemple un comportement. Dans certaines approches, génotype et phénotype reposent sur une même structure de données.

nismes chimiques opérant indépendamment du « sens » que peuvent prendre ces modifications lorsqu'elles affecteront la synthèse d'une protéine.

Par comparaison avec les machines, qui tombent si facilement en panne, les êtres vivants sont étonnamment résistants à ces modifications. En fait, il apparaît que la plupart des mutations survenant au niveau moléculaire sont *sélectivement neutres*, c'est-à-dire qu'elles n'ont pas d'incidence sur l'adaptation de l'être vivant à son milieu (KIMURA 1990).

Quelques mutations sont délétères, d'autres enfin ont des effets bénéfiques. Dans un système artificiel, implanté sur un ordinateur, on peut légitimement penser que le risque est grand de rencontrer des mutations délétères, si l'on songe à la fragilité des programmes informatiques – nous reviendrons plus précisément sur ce point lorsque nous examinerons les principaux paradigmes évolutionnistes existants.

Pourtant, l'évolution darwinienne apparaît comme une forme d'apprentissage non supervisé et d'exploration d'un espace de comportements (DENNETT 1996). Elle fonctionne en tant qu'algorithme de création d'ordre et c'est ce qui fait son avantage sur des approches plus proches de l'ingénierie.

Pour conserver cette possibilité d'une évolution aveugle, il faudra donc veiller à ce que le modèle soit *résistant aux mutations*, c'est-à-dire que des variations dépourvues de sémantique affectant le génotype n'aient pas des effets catastrophiques sur le phénotype (le comportement des agents).

#### 4<sup>o</sup> possibilité d'une adaptation cumulative

Il ne faut pas non plus oublier que l'évolution est avant tout basée sur l'*hérédité* : les enfants ressemblent à leurs parents, ils possèdent certains de leurs caractères.

En termes informatiques, cela signifie que les variations affectant un point donné du génotype ne doivent (statistiquement) pas avoir d'effet majeur sur l'ensemble de la construction du phénotype. Sans quoi, le croisement de deux individus peut produire des descendants *très différents* des parents, et l'on tomberait alors dans une exploration aléatoire de l'espace des comportements. On souhaite au contraire que des individus correctement adaptés à un environnement produisent une progéniture possédant un degré d'adaptation du même ordre (quelques-uns peuvent être mieux adaptés, d'autres moins : mais ces variations doivent rester faibles).

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

Cela impose notamment qu'existe une *quasi-continuité* entre le génotype et le phénotype (DESSALLES 1992; DESSALLES 1996a) : de faibles variations dans le génotype ne doivent entraîner, dans la grande majorité des cas, que de faibles modifications du phénotype (en l'occurrence, le comportement).

### 5° possibilité d'un contrôle autonome de l'évolution

Dans les modèles biologiques classiques (VOGEL et ANGERMANN 1994) un seul type de molécule, les protéines, assure des « fonctions » dans la cellule, relevant à la fois :

- de la morphologie (protéines « de structure ») ;
- de la survie de la cellule dans son environnement, par exemple le métabolisme (transformation de substances chimiques pour produire de l'énergie) ;
- de l'auto-reproduction de la cellule (sa capacité à se dupliquer) ;
- de l'expression des gènes, c'est-à-dire de la manière dont est utilisé le matériel héréditaire pour produire un phénotype ;
- du rôle d'« opérateurs génétiques » : la recopie du matériel héréditaire (avec erreurs), l'appariement et l'échange de portions entre deux brins d'ADN de chromosomes homologues, etc.

Ces phénomènes contribuent à faire des êtres vivants les plus simples des systèmes d'une grande complexité (ATLAN 1999), mais en même temps leur confèrent des capacités d'adaptation énormes.

Aussi, pour construire des SMA fortement adaptatifs en vue d'un fonctionnement situé, nous pouvons exiger que le modèle *permette d'emblée la gestion par les mêmes entités* de la plupart des rôles mentionnés ci-dessus.

Ne nous y trompons pas : il est important de garder une *démarche incrémentale* en construisant le système pas à pas, module après module si nécessaire. Toutefois, il faut aussi n'avoir pas à tout reconstruire s'il s'avère que le système doit avoir par exemple suffisamment d'autonomie pour contrôler sa propre évolution.

La proposition que nous faisons consiste donc essentiellement à imposer que toutes les entités participant au fonctionnement évolutif du système (opérateurs génétiques par exemple) soient des agents construits de la même façon que ceux assurant collectivement la *fonction* du SMA – c'est-à-dire que leur comportement soit *susceptible d'évoluer*.

En résumé, ces contraintes imposent à la structure produisant les comportements les caractéristiques suivantes :

- 1° riche *pouvoir expressif* et *modularité* ;
- 2° *intelligibilité immédiate* ;
- 3° *robustesse vis-à-vis des mutations* survenant dans le génotype ;
- 4° *quasi-continuité* entre le génotype et le phénotype ;
- 5° *homogénéité* du système (permettant le cas échéant *autonomie* et *réflexivité*).

### 10.2.2 Approches évolutionnistes existantes

Nous allons donc maintenant examiner les principaux paradigmes d'algorithmes évolutionnistes (essentiellement les Algorithmes Génétiques et la Programmation Génétique<sup>3</sup>) en essayant de voir lesquels satisfont aux prérequis de l'Éthogénétique. Nous mentionnerons également quelques modèles moins connus présentant des caractéristiques intéressantes.

#### Algorithmes génétiques et stratégies évolutionnistes

Dans le paradigme des « Algorithmes Génétiques » (HOLLAND 1975; DE JONG 1975; GOLDBERG 1989, ...) – respectivement le paradigme « Stratégies Évolutionnistes » (RECHENBERG 1973; SCHWEFEL 1975; BÄCK et SCHWEFEL 1993) – le génotype est une chaîne (en général binaire) – resp. un tableau de nombres décimaux.

Cette chaîne code un ensemble de valeurs correspondant par exemple :

- aux variables d'une fonction à optimiser ;
- les poids d'un réseau de neurones ;
- des paramètres pour un comportement *préexistant*.

La chaîne en elle-même est dépourvue de sémantique, seuls les paramètres issus de son décodage ont un sens qui leur est donné par le cadre où ils sont utilisés.

La chaîne – resp. le tableau – est manipulée par des opérateurs génétiques fixés (mutation inversant un bit, croisement de deux chaînes pour produire un

---

3. Nous nous livrons dans (LANDAU et PICAULT 2001b) à une étude comparative plus poussée avec d'autres approches.

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

nouvel individu). Ces opérateurs sont « aveugles », c'est-à-dire qu'ils opèrent sur une structure homogène et sans signification.

Le principe des algorithmes génétiques permet, si l'on cherche à le transposer à l'évolution de comportements, de suivre le schéma suivant (la première ligne représente la façon dont le phénotype est produit à partir du génotype; la seconde, comment sont produites les variations du génotype) :

$$\left\{ \begin{array}{l} \text{Chaîne} \xrightarrow{\text{Traduction}} \text{Paramètres} \xrightarrow{\text{Utilisation}} \text{Comportement} \\ \text{Opérateurs prédéfinis} \rightsquigarrow \text{Chaînes} \end{array} \right.$$

Comme la chaîne constitue un génotype de granularité fine, moyennant un codage judicieux on peut garantir une *quasi-continuité* entre le génotype et le phénotype. En revanche, leur *pouvoir expressif* est très faible, puisqu'il se limite à explorer un espace de *paramètres* qui modifient des comportements préétablis. Le comportement en tant que tel n'évolue pas.

### La programmation génétique

Avec le paradigme « Programmation Génétique » (KOZA 1992) ou les approches qui en dérivent (GRUAU 1994), le comportement d'un individu est *directement produit par l'exécution d'un programme*.

Ce programme est représenté sous la forme d'un arbre contenant des instructions et des paramètres. Il peut servir à calculer une fonction, choisir des actions en fonction de perceptions, etc. et possède donc un *pouvoir expressif* considérable : la complexité des comportements obtenus ne dépend en effet que de la grammaire appliquée aux arbres.

Mais cet arbre qui permet de représenter et d'exécuter le programme (donc le phénotype) constitue en outre le support de l'hérédité (le génotype).

$$\left\{ \begin{array}{l} \text{Arbre de lexèmes} \xrightarrow{\text{Parcours}} \text{Programme} \xrightarrow{\text{Exécution}} \text{Comportement} \\ \text{Opérateurs prédéfinis} \rightsquigarrow \text{Arbres de lexèmes} \end{array} \right.$$

Cela conduit à des problèmes difficiles :

- les opérateurs génétiques doivent opérer en tenant compte des contraintes syntaxiques qui pèsent sur l'arbre (nature et arité des nœuds, typage des valeurs, ...);

- les variations génétiques s'exercent sur des entités de forte granularité (des instructions ou des paramètres, chargés d'une forte sémantique) ;
- les variations affectant les nœuds proches de la racine ont des conséquences très lourdes par rapport à celles affectant les feuilles.

On perd donc en partie l'hypothèse de quasi-continuité entre le génotype et le phénotype, et la construction d'un individu à partir de ses parents pose de sérieux problèmes de sémantique.

### Auto-réplication

Nous quittons ici le domaine des paradigmes classiques pour mentionner quelques approches visant à reproduire des caractéristiques propres aux êtres vivants naturels, telles qu'une capacité d'auto-réplication du génotype, ou d'auto-reproduction de la cellule.

**La « Typogénétique »** inventée par HOFSTADTER (HOFSTADTER 1985) n'est pas à proprement parler une méthode évolutionniste, bien qu'inspirée des mécanismes de la biologie moléculaire. C'est en réalité un système de production, dans lequel des chaînes sont produites par l'action de « typoenzymes » sur des chaînes existantes.

Ces typoenzymes sont des programmes (séquences d'instructions) issus d'une traduction déterministe de ces chaînes. Seules les typoenzymes (produites par des chaînes) peuvent affecter les chaînes : il n'y a pas ici d'opérateurs génétiques à proprement parler.

$$\left\{ \begin{array}{l} \text{Chaîne} \xrightarrow{\text{Traduction}} \text{Typoenzymes (séquence d'instructions)} \\ \text{Typoenzymes (en tant que règle de production)} \rightsquigarrow \text{Chaînes} \end{array} \right.$$

La typogénétique ne vise pas particulièrement à produire des comportements *adaptés à un environnement* : HOFSTADTER s'intéresse plutôt à la question de l'existence de chaînes *auto-répliquantes*. Une chaîne est auto-répliquante si l'action de la typoenzyme qu'elle code, appliquée à elle-même, produit une copie de la chaîne initiale.

Ce système a comme intérêt majeur d'être *réflexif*, au sens où ce qui correspondrait au « phénotype » (l'action des typoenzymes) s'applique au « génotype » (les chaînes), alors que celui-ci est d'ordinaire manipulé par des mécanismes de sélection et de reproduction distincts.

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

La typogénétique apparaît donc essentiellement comme un jeu :

- La traduction des chaînes en typoenzymes, comme l'action des typoenzymes sur les chaînes, est *déterministe*. La chaîne de départ une fois choisie, celles qui seront produites sont déjà fixées. Autrement dit, ce système ne peut pas *s'adapter*, il n'y a aucune *sélection*
- Le seul rôle des typoenzymes est d'agir sur les chaînes, alors que dans une cellule vivante, l'action des enzymes concourt en outre aux diverses « fonctions » nécessaires à sa survie et à sa reproduction (métabolisme entre autres).

**TIERRA** de Tom RAY (RAY 1991), s'intéresse à l'évolution de programmes peuplant la mémoire d'une machine virtuelle. Cette approche s'inscrit résolument dans une perspective de « Vie Artificielle » et permet d'étudier des problématiques qu'on rencontre dans des écosystèmes naturels (compétition pour des ressources, lutte entre proies et prédateurs, parasitisme, etc.).

La mémoire (qui constitue le milieu dans lequel vont évoluer les individus) estensemencée avec un programme auto-réplicateur. Le « processeur » de la machine virtuelle produit des erreurs aléatoires lors de l'exécution des instructions, assurant ainsi, entre autres, une variation du code lorsque les programmes cherchent à se répliquer.

$$\left\{ \begin{array}{l} \text{Code} \xrightarrow{\text{Traduction}} \text{Programme (langage-machine)} \xrightarrow{\text{Exécution}} \text{Comportement} \\ \text{Comportements (des programmes)} \rightsquigarrow \text{Code} \end{array} \right.$$

Ces mécanismes introduisent ainsi une perspective d'Évolution darwinienne dans un système proche par ailleurs de la typogénétique. Cette fois encore, le système est réflexif : tout programme est capable de traiter ses propres instructions et celles des autres programmes comme des données (à la façon d'un langage-machine). C'est donc en fait la même structure qui joue le rôle de génotype et de phénotype.

Toutefois on ne saurait « utiliser » TIERRA pour quelque fonctionnalité que ce soit, autre que l'observation d'un écosystème artificiel, puisque l'« environnement » dans lequel agissent les programmes se limite à l'ensemble des programmes présents dans la mémoire de la machine virtuelle. Ces programmes ne peuvent rien faire d'autre que se reproduire et agir sur d'autres programmes.

**EDEN.** Pour dépasser les limitations comportementales des algorithmes génétiques et de la programmation génétique, on peut songer à leur adjoindre des caractéristiques présentes dans la typogénétique d'HOFSTADTER ou dans TIERRA, comme la possibilité d'une action du système sur lui-même.

On peut par exemple envisager un algorithme « métagénétique » qui fait des opérateurs génétiques *le* problème à optimiser ; le système peut également être « endosémantique », c'est-à-dire laisser le génotype contrôler sa propre signification en influant sur le processus de traduction (DESSALLES 1996a).

Le modèle EDEN (PICAULT et al. 1997) est une première tentative dans ce sens, visant à renforcer l'autonomie de l'expression génétique tout en permettant une « utilisation » du système dans un cadre fonctionnel.

Dans EDEN, une chaîne est traduite en une séquence d'instructions composant un programme appelé « enzyme ». Ces enzymes peuvent avoir diverses fonctions (selon la définition du langage de programmation) ; en particulier, elles peuvent effectuer la traduction de la chaîne en d'autres enzymes et réaliser des opérations génétiques (mutations, croisements, etc.).

La métaphore biologique correspondante consiste à voir l'ADN comme une chaîne codant des séquences d'acides aminés (les instructions) composant des enzymes (programmes) qui seules ont une action dans la cellule – vision éminemment réductrice des phénomènes biologiques, mais en somme nettement moins que celle des algorithmes génétiques. Rappelons que notre but n'est nullement de *simuler* une cellule mais de s'en inspirer de façon métaphorique.

$$\left\{ \begin{array}{l} \text{Chaîne} \xrightarrow[\text{= Enzyme}]{\text{Traducteur}} \text{Enzymes (séqu. d'instructions)} \xrightarrow{\text{Interprétation}} \text{Comportements} \\ \text{Comportements (des enzymes)} \rightsquigarrow \text{Chaîne} \end{array} \right.$$

Nous ne décrivons pas en détail les caractéristiques d'un tel système. L'expression du matériel génétique et l'évolution des cellules sont auto-régulées selon des critères implicites (stabilité des opérateurs de mutation et des enzymes de traductions, « survie » des individus, etc.). L'obstacle majeur rencontré dans EDEN est la faible robustesse des enzymes vis-à-vis des mutations.

En effet, comme dans les arbres de la Programmation Génétique, une mutation a des conséquences d'autant plus graves qu'elle opère sur le début de la séquence d'instructions composant le programme. De plus se pose un problème de construction de programmes « bien formés » (ayant un sens) : une mutation perturbe l'intégralité de la fonction réalisée par un programme (voire lui fait perdre toute fonction).

## 10.3 ATNoSFERES, un framework pour l'Éthogénétique

Force est de constater qu'aucune des approches évolutionnistes classiques ne présente simultanément toutes les caractéristiques requises pour l'évolution de comportements d'agents, ni *a fortiori* pour l'ontophylogénèse d'un SMA :

- le système doit pouvoir être construit par un processus évolutionniste aveugle, donc en particulier maintenir l'hypothèse de quasi-continuité entre génotype et phénotype : cette quasi-continuité fait défaut aux approches utilisant un génotype fortement structuré (par exemple la Programmation Génétique) ;
- le comportement doit pouvoir être de complexité quelconque : ce n'est pas le cas avec les Algorithmes Génétiques ;
- la construction des comportements doit être incrémentale et modulaire : cela pose de sérieux problèmes avec les structures arborescentes de la Programmation Génétique.

Nous avons donc conçu un modèle spécifique permettant de mettre en œuvre les principes de l'Éthogénétique. Ce modèle a été implémenté dans un framework, ATNoSFERES (LANDAU et al. 2001b)<sup>4</sup>.

Nous allons maintenant décrire ce modèle en détail. Les choix que nous y avons faits tentent de trouver un compromis entre :

- un génotype à grain fin (comme dans les Algorithmes Génétiques) ;
- un fort pouvoir expressif (comme dans la Programmation Génétique) ;
- la possibilité d'une évolution complètement autonome du système.

### 10.3.1 Principes d'ATNoSFERES

L'idée principale du modèle utilisé dans ATNoSFERES pour produire un comportement à partir d'un génotype repose sur une métaphore biologique (cf. figure 10.1) :

1° Dans une cellule, c'est la *forme* d'une protéine qui contribue en grande

---

4. Le noyau d'ATNoSFERES a d'abord été écrit en SMALLTALK à des fins de modélisation et de test, puis intégré en C++ au framework SFERES (LANDAU et al. 2001a), qui fournit des techniques évolutionnistes classiques et un environnement de simulation multi-agent, à des fins d'expérimentation à large échelle. De plus amples détails techniques sur le framework ATNoSFERES pourront être trouvés dans (PICAULT et LANDAU 2001a) disponible sur : <http://miriad.lip6.fr/atnosferes>

### 10.3. ATNOSFERES, UN FRAMEWORK POUR L'ÉTHOGÉNÉTIQUE

partie à sa fonction, et non simplement la séquence des acides aminés dont elle est constituée (VOGEL et ANGERMANN 1994)).

2° Cette forme dépend d'*interactions chimiques locales* (où interviennent parfois des enzymes) entre les acides aminés qui la constituent ;

3° La séquence d'acides aminés résulte de la traduction d'une autre chaîne, un acide nucléique (généralement l'ARN<sub>t</sub>, obtenu par copie de portions de l'ADN).

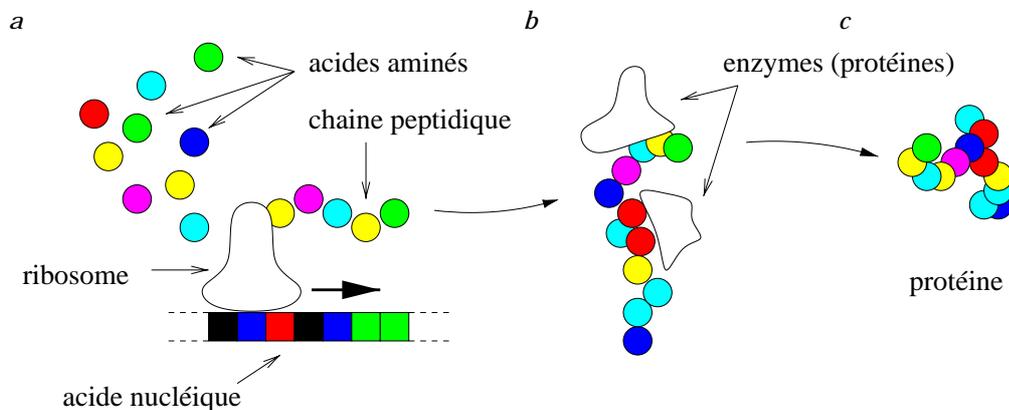
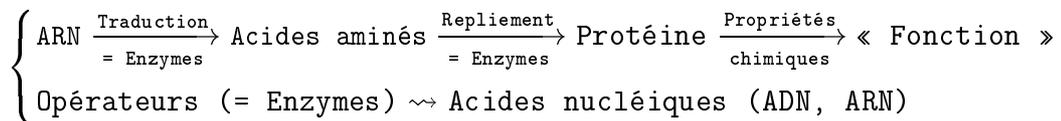


FIG. 10.1 – Mécanisme de la synthèse des protéines dans une cellule. (a) Le génotype, sous la forme d'un acide nucléique (ARN) est d'abord traduit sous l'action d'un organelle, le ribosome : à chaque triplet de bases de l'ARN est associé un acide aminé spécifique, ce qui permet de constituer une chaîne d'acides aminés ou chaîne peptidique. (b) Sous l'action d'enzymes et sous l'effet des interactions chimiques locales entre les acides aminés, la chaîne peptidique adopte une certaine forme (c). qui constitue la protéine finale et contribue à son rôle fonctionnel dans la cellule. — d'après (LANDAU et PICAULT 2001b)

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

Par analogie, nous proposons de suivre le schéma suivant<sup>5</sup> :

$$\left\{ \begin{array}{l} \text{Chaîne} \xrightarrow[\text{= Agent}]{\text{Traducteur}} \text{Lexèmes} \xrightarrow[\text{= Agent}]{\text{Interpréteur}} \text{Graphe} \xrightarrow{\text{Parcours}} \text{Comportement (d'agent)} \\ \text{Opérateurs (= Agents)} \rightsquigarrow \text{Chaînes} \end{array} \right.$$

La traduction de la chaîne binaire (appelée par la suite « chromosome ») comme sa manipulation (opérations génétiques) sont effectuées par des agents, dont le comportement peut faire lui-même l'objet d'une construction à partir d'un chromosome.

### 10.3.2 Produire un comportement : les ATN

Pour produire le comportement des agents dans ATNoSFERES, nous avons décidé de faire appel à une structure de graphe orienté et étiqueté.

En l'occurrence, nous utilisons des ATN<sup>6</sup>. Ces graphes ont été utilisés à l'origine dans le cadre du traitement automatique du langage (WOODS 1970; WINOGRAD 1983), mais ils ont également été appliqués à la description des comportements d'agents, par exemple par (BOURON 1993; GUESSOUM 1996).

Les ATN sont constitués de noeuds connectés par des arcs, ceux-ci pouvant être étiquetés par 1<sup>o</sup> un ensemble de conditions, 2<sup>o</sup> une liste d'actions. Le parcours d'un tel graphe suppose qu'à chaque nœud, les arcs dont les conditions ne sont pas vérifiées ne peuvent être franchis. Lorsqu'un arc est franchi (i.e. s'il ne porte aucune condition ou bien si toutes les conditions sont vérifiées simultanément), on exécute la liste d'actions. Nous donnons un exemple d'ATN sur la figure 10.2<sup>7</sup>.

Pour faire de ces graphes étiquetés la description de comportements d'agents, nous avons défini une superclasse des agents, **ATNAgent**, à charge pour chaque sous-classe d'agents de définir un ensemble de lexèmes correspondant d'une part à des *actions* pouvant être exécutées par les instances de cette classe (des primitives comportementales), d'autre part à des *conditions* portant sur l'environnement des agents, leur état interne, leurs représentations, leurs buts. Le graphe sert donc à *organiser des capacités de perceptions et d'actions*.

---

5. Nous emploierons indifféremment les termes « lexème », « symbole » ou « instruction » dans le texte, qui sont remplacés par leur équivalent anglais de *token* sur les schémas et à l'intérieur du framework.

6. ATN est l'abréviation de *Augmented Transition Network*.

7. Sur les graphes donnés en exemple, les actions seront notées **uneAction!** et les conditions **uneCondition?**.

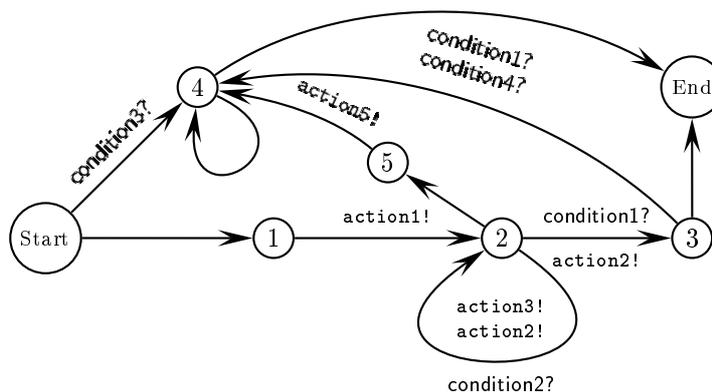


FIG. 10.2 – Un exemple d'ATN. Les arcs sont étiquetés par un ensemble de conditions et une liste d'actions. Deux nœuds jouent un rôle particulier : « Start » est l'état initial de l'agent ; « End » représente un état dans lequel l'agent cesse d'agir.

Le graphe de comportement d'un agent comporte systématiquement un nœud de départ (appelé par la suite *Start*) et un nœud final (*End*). Les autres nœuds sont reliés par des arcs qui peuvent être étiquetés par un ensemble de conditions et par une séquence d'actions (voir l'exemple donné dans la figure 10.2).

Dans ATNoSFERES, le comportement des agents repose systématiquement sur un ATN et consiste à passer de nœud en nœud. Tout agent est initialisé dans l'état *Start*.

L'activité de l'agent consiste alors pour chaque nœud à :

1. Sélectionner, parmi les arcs issus du nœud courant, ceux qui sont *franchissables*, c'est-à-dire sans conditions ou dont les conditions sont vérifiées *simultanément*.
2. Choisir un de ces arcs *aléatoirement*.
3. Le franchir (se placer sur le nœud auquel il aboutit) après avoir exécuté les actions éventuellement associées à l'arc (dans l'ordre).

L'agent cherche en permanence à passer d'un nœud à un autre ; si aucun arc n'est franchissable, il reste dans l'état courant, et tentera de nouveau au « pas de temps » suivant. Il passe en veille lorsqu'il se trouve sur le nœud *End* et cesse alors d'agir.

On notera dès à présent que deux agents possédant un même ATN peuvent adopter des comportements très différents (à partir des mêmes ca-

pacités de perception et d'actions), dans la mesure où ils peuvent emprunter des arcs différents même s'ils sont placés dans les mêmes conditions. Comme nous le verrons dans les expériences présentées au chapitre suivant, ce poly-éthisme peut être rejeté par la sélection darwinienne lorsqu'il n'est pas utile à l'adaptation des agents à leur environnement, mais il constitue une potentialité importante du système.

### 10.3.3 La construction de l'ATN : traduction et interprétation

L'ATN de chaque agent est construit à partir de son génotype, une chaîne de bits appelée « chromosome ». Cette construction s'effectue en deux étapes pour être robuste par rapport aux mutations, s'appuyer sur des relations locales entre les portions de la chaîne, et assurer une continuité dans la production du phénotype (l'ATN) par rapport au génotype. Ces étapes sont résumées sur la figure 10.3.

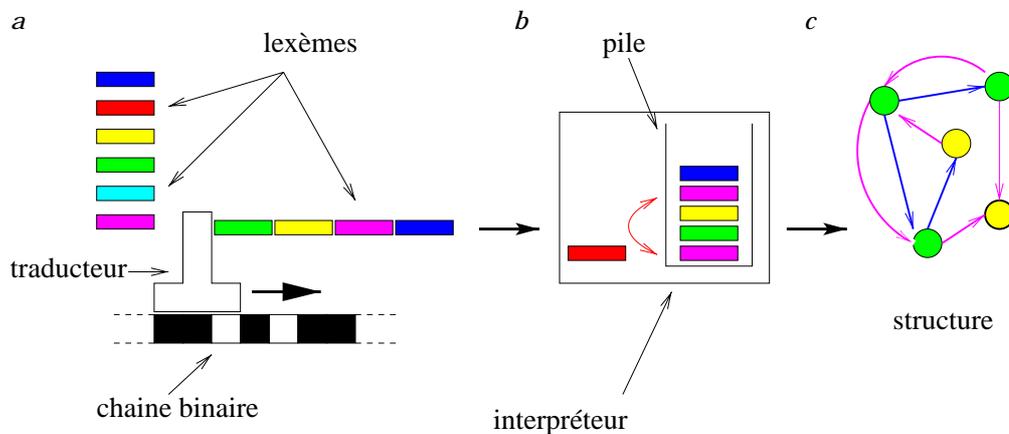


FIG. 10.3 – Mécanisme d'expression génétique « à pile » utilisé dans ATNoS-FERES. (a) Le génotype, sous la forme d'une **chaîne binaire**, est d'abord traduit sous l'action d'un **Traducteur** : à tout  $n$ -uplet de bits est associé un lexème (token) spécifique. (b) La succession de tokens est envoyée ensuite à un **Interpréteur** qui les interprète comme des instructions d'un langage à pile pour construire un graphe (c), l'**ATN**, qui contrôlera le comportement de l'agent. — d'après (LANDAU et PICAULT 2001b)

### Le traducteur

La traduction a pour but de décoder le chromosome d'une classe d'agents donnée en une succession de lexèmes (envoyée à un interpréteur) ; elle s'appuie pour cela sur un *code génétique*, i.e. une fonction de traduction qui permet d'associer des lexèmes à des sous-chaînes du chromosome appelées « codons ». Ces lexèmes représentent, d'une part, les actions et conditions réalisables par la classe d'agents considérée, et d'autre part, des instructions destinées à l'interpréteur (nous détaillerons ce point plus bas).

Les codons étant des  $n$ -uplets binaires, le code génétique est généralement surjectif : il y a généralement strictement plus de codons que de lexèmes (il ne peut y en avoir moins) et l'on a donc de la *redondance* (un même lexème est codé par plusieurs codons).

Cette redondance peut aider à assurer la quasi-continuité entre le génotype des agents et leur phénotype (comportement). Ainsi, si l'on impose une distance de Hamming de 1 entre tous les codons traduits en un même lexème, on disposera d'une résistance accrue aux mutations. On peut également jouer sur les fréquences d'occurrence de chacun des lexèmes si cela s'avère efficace.

Le but de cette manipulation n'est pas de « freiner » nécessairement l'évolution mais de la contraindre de façon pertinente selon des besoins bien identifiés (par exemple il peut être utile de favoriser les mutations transformant une action en une autre, ou au contraire une action en une condition ; ou encore, de favoriser la complexité comportementale en multipliant le nombre d'instructions de création de nœuds, etc.).

Toutefois, dans le cas général, on pourra se contenter de générer un code génétique d'une façon automatique, en le complétant de façon mécanique. C'est ce que nous ferons dans nos expériences.

Le rôle du traducteur est simple : il parcourt le chromosome et le traduit, codon par codon, en lexèmes. Ces lexèmes sont envoyés en séquence à l'interpréteur ; lorsqu'il n'y a plus de codon à lire la production de lexèmes s'interrompt et l'interpréteur construit l'ATN.

### L'interpréteur

L'interpréteur reçoit une succession de lexèmes. Ceux-ci vont être pour une part des actions et des conditions propres à la classe d'agent considérée. D'autres devront comporter un « mode d'emploi » décrivant comment intégrer

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

ces conditions et ces actions à une structure d'ATN : il faut donc pouvoir créer des nœuds et des arcs.

Enfin, si l'on souhaite assurer les propriétés imposées par l'Éthogénétique, telles que la robustesse vis-à-vis des mutations et la quasi-continuité entre le génotype et le phénotype, il faut garantir :

- que la substitution d'un lexème à un autre (consécutive à une mutation) n'empêche pas la construction d'un ATN cohérent ;
- qu'une mutation ait un effet aussi *local* que possible, c'est-à-dire n'affecte *pas systématiquement* ou seulement *faiblement* la sémantique des lexèmes suivants.

Aussi nous avons choisi d'utiliser un langage faisant appel à une *pile* (cf. tableau 10.2 et figure 10.4 page 184). Les lexèmes de ce langage se répartissent en plusieurs catégories (cf. tableau 10.1) : données (lexèmes d'actions et de conditions pour une classe d'agent), structuration des données (création et connexion de nœuds) et manipulation des données dans la pile.

- Les lexèmes d'actions et de conditions sont simplement empilés.
- Les lexèmes *node*, *startNode*, *endNode* entraînent l'empilement d'un objet nœud (les deux derniers spécifiant explicitement que le nœud sera raccordé respectivement à *Start* et à *End*).
- Le lexème *connect* (dont le fonctionnement est illustré par la figure 10.4 page 184) provoque l'établissement d'un arc entre les deux premiers nœuds rencontrés dans la liste, en étiquetant l'arc avec les actions et conditions trouvées jusqu'au deuxième nœud. Par exemple, *connect* appliqué à la liste : (**a1!**, **c1?**, Nœud  $N_i$ , **c2?**, **a2!**, Nœud  $N_j$ , **a3!**, ...) transformera la liste en : (Nœud  $N_i$ , Nœud  $N_j$ , **a3!**, ...) et créera un arc de  $N_j$  vers  $N_i$ , étiqueté par les conditions **c1?** et **c2?**, et par la séquence d'actions **a2!–a1!**.
- Quant aux lexèmes de manipulation de données, ils peuvent agir soit spécifiquement sur des objets (i.e. des nœuds), soit sur une donnée quelconque de la liste.

À cette définition du langage à pile, il faut ajouter que les instructions manipulant des objets (i.e. *connect* et *dupObject*) sont *rémanentes*, c'est-à-dire que si elles ne peuvent pas s'exécuter correctement faute de paramètres adéquats dans la liste, elles se placent dans la liste en attendant l'arrivée d'un objet susceptible de les déclencher.

### 10.3. ATNOSFERES, UN FRAMEWORK POUR L'ÉTHOGÉNÉTIQUE

TAB. 10.1 – *Les catégories de lexèmes.*

Notation	Signification	Exemples
<b>action!</b>	lexème d'action propre à la classe d'agent considérée	<b>turn!</b> , <b>forward!</b> etc.
<b>condition?</b>	lexème de condition propre à la classe d'agent considérée	<b>lowBattery?</b> , <b>lost?</b> , etc.
<i>*node,</i> <i>connect</i>	lexèmes de structure d'ATN	<i>node</i> , <i>start-</i> <i>Node</i> , <i>endNode</i> , <i>connect</i>
<i>manip. de</i> <i>données</i>	lexèmes de manipulation de données (propres à l'interpréteur)	<i>dup</i> , <i>swap</i> , <i>dupObject</i> , etc.

TAB. 10.2 – *Le langage de construction des ATN.*

lexème	(état de la liste)	→	(liste résultante)
	<i>condition?</i> ( <i>x ...</i> )	→	( <i>condition?</i> <i>x ...</i> )
	<b>action!</b> ( <i>x ...</i> )	→	( <b>action!</b> <i>x ...</i> )
	<i>node</i> ( <i>x ...</i> )	→	( <i>N<sub>i</sub> x ...</i> ) <sup>a</sup>
	<i>startNode</i> ( <i>x ...</i> )	→	( <i>N<sub>i</sub> x ...</i> ) <sup>b</sup>
	<i>endNode</i> ( <i>x ...</i> )	→	( <i>N<sub>i</sub> x ...</i> ) <sup>c</sup>
<i>connect</i> ( <i>x N<sub>i</sub> y c1? z a1! a2! t N<sub>j</sub> u ...</i> )		→	( <i>x N<sub>i</sub> y z t N<sub>j</sub> u ...</i> ) <sup>d</sup>
	<i>swap</i> ( <i>x y ...</i> )	→	( <i>y x ...</i> )
	<i>dup</i> ( <i>x y ...</i> )	→	( <i>x x y ...</i> )
	<i>dupObject</i> ( <i>x y N<sub>i</sub> z ...</i> )	→	( <i>N<sub>i</sub> x y N<sub>i</sub> z ...</i> )
	<i>popRoll</i> ( <i>x y ... z</i> )	→	( <i>y ... z x</i> )
	<i>pushRoll</i> ( <i>x ... y z</i> )	→	( <i>z x ... y</i> )
	<i>forget</i> ( <i>x y ...</i> ) + [ <i>z ...</i> ]	→	( <i>y ...</i> ) + [ <i>x z ...</i> ] <sup>e</sup>
	<i>recall</i> ( <i>x ...</i> ) + [ <i>y z ...</i> ]	→	( <i>y x ...</i> ) + [ <i>z ...</i> ] <sup>e</sup>

<sup>a</sup> création d'un nœud

<sup>b</sup> création d'un nœud connecté au *Start*

<sup>c</sup> création d'un nœud connecté au *End*

<sup>d</sup> création d'un arc entre *N<sub>i</sub>* et *N<sub>j</sub>*

<sup>e</sup> utilisation d'une pile auxiliaire

## ◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

Enfin, lorsqu'il n'y a plus de lexèmes à interpréter, l'interpréteur peut achever la construction de l'ATN. Des connexions implicites sont alors réalisées entre les nœuds restants dans la pile (les actions et conditions situées entre deux nœuds sont utilisées pour étiqueter un nouvel arc entre ces nœuds), puis on assure la cohérence de l'ATN en connectant *Start* aux nœuds qui ne possèdent pas d'arcs entrants (sauf issus d'eux-mêmes), et à *End* les nœuds ne possédant aucun arc sortant.

Ainsi, le langage obtenu possède les caractéristiques souhaitées :

- il est « localement » déclaratif : l'ordre des lexèmes importe relativement peu (on peut notamment permuter des groupes de lexèmes sans que l'ATN résultant soit modifié) ;
- il est économique : une instruction de manipulation de nœud n'est pas « perdue », elle attend de pouvoir s'exécuter ; nombre d'informations de structure peuvent rester implicites (du fait des connexions automatiques) ; on peut spécifier qu'un nœud doit être relié à *Start* simplement au moyen d'un lexème approprié (sans avoir à établir explicitement la connexion) ;
- il est robuste par rapport aux mutations : en effet, la construction de l'ATN fait appel à des interactions locales entre lexèmes, donc la modification de l'un d'entre eux n'a pas nécessairement de répercussions fortes sur l'ensemble du graphe fabriqué.

Nous avons de plus vérifié qu'il permettait la construction de n'importe quel ATN spécifié arbitrairement, ce qui garantit le *pouvoir expressif* du modèle.

### 10.3.4 Les classes d'agents

La superclasse de tous les agents d'ATNoSFERES, baptisée **ATNAgent**, détermine son comportement en cherchant à passer d'un nœud à un autre dans l'ATN.

Contrairement par exemple à l'utilisation d'un arbre en Programmation Génétique, *l'ATN n'est pas parcouru intégralement lorsqu'on cherche à produire une action* : si l'agent fonctionne au sein d'un SMA utilisant une horloge, chaque « pas de temps » correspondra au franchissement d'un arc.

Chaque sous-classe d'**ATNAgent** peut redéfinir les lexèmes d'action et de condition qui lui sont propres : nous en verrons un exemple concret au chapitre suivant.

### 10.3. ATNOSFERES, UN FRAMEWORK POUR L'ÉTHOGÉNÉTIQUE

Tout agent est plongé dans un environnement dans lequel s'exercent ses actions et ses perceptions.

Il possède de plus des attributs caractérisant son activité (en tant que « processus ») : prêt (`#ready`) ou endormi (`#asleep`), ainsi que le résultat de ses actions (`actionResult`) : succès (`#success`) ou échec (`#failure`). Il peut d'ailleurs au besoin faire appel à des conditions réflexives, telles que `success?` ou `failure?` qui testent le résultat des actions effectuées.

Nous avons décrit ici les principales caractéristiques du modèle ATNoS-FERES. Nous allons maintenant donner des exemples de fonctionnement et d'expériences menées avec ce framework, puis nous donnerons quelques extensions en cours d'expérimentation.

◇10. MODÈLES D'AGENTS ÉVOLUTIONNISTES

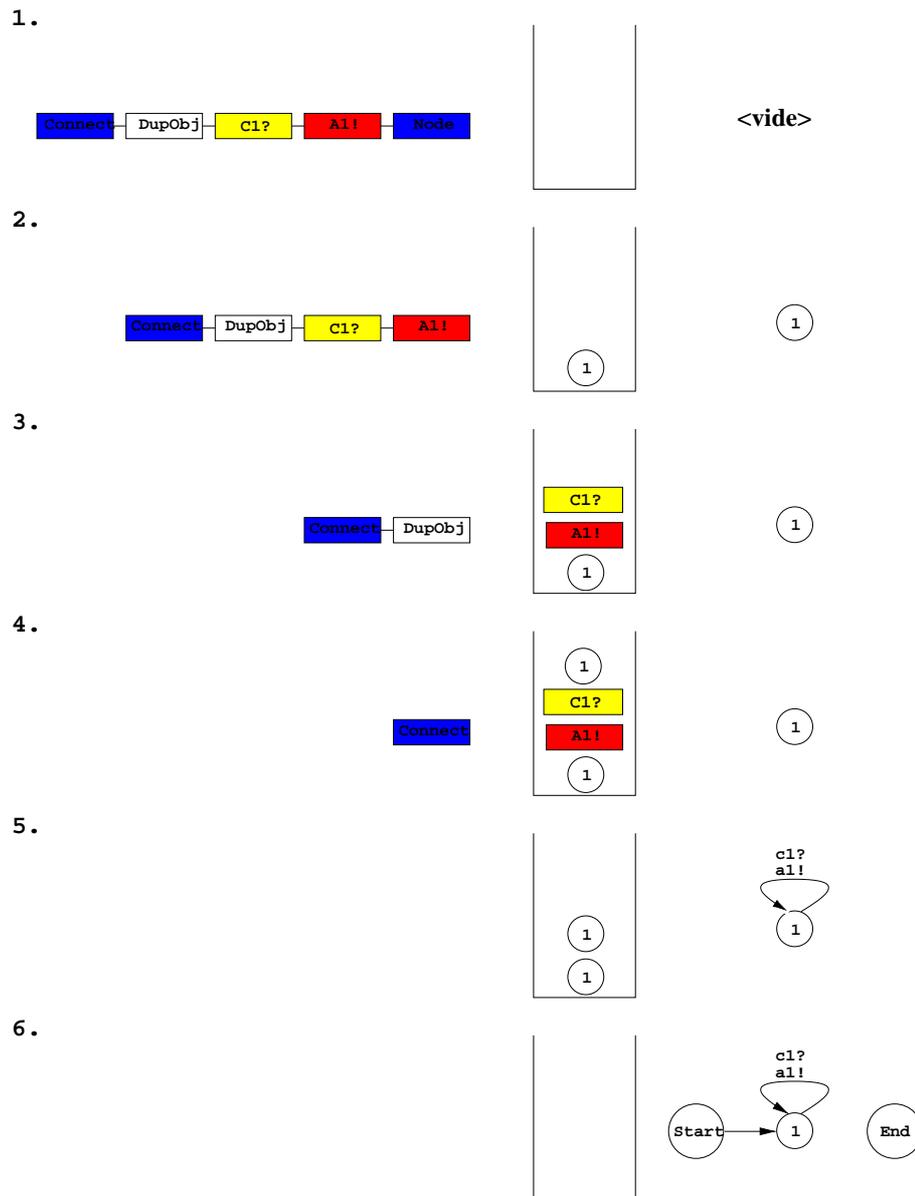


FIG. 10.4 – Exemple d'interprétation de quelques lexèmes dans ATNoSFERES. **1.** Début de l'interprétation : la pile (au centre) et le futur ATN (à droite) sont vides ; une séquence de lexèmes est envoyée à l'interpréteur. **2.** Le lexème node crée un nœud qui reste présent dans la pile et amorce l'ATN. **3.** Les lexèmes d'action et de condition sont seulement empilés. **4.** Le lexème DupObject duplique le premier nœud trouvé dans la pile (la structure en construction n'est pas modifiée). **5.** Le lexème connect provoque la connexion entre les deux premiers nœuds de la pile, en l'occurrence deux copies du nœud 1 ; les actions et conditions situées entre ces nœuds sont utilisées pour étiqueter l'arc et retirées de la pile. **6.** Fin de l'interprétation : il n'y a plus de lexèmes ; l'interpréteur complète l'ATN en établissant des connexions cohérentes avec les nœuds Start et End. — d'après (LANDAU et PICAULT 2001b)

---

# Chapitre onzième.

## Expériences et extensions

---

*Nous avons jusqu'à présent décrit une architecture d'agent évolutive, dans laquelle un génotype à grain fin sert à construire un graphe qui engendre le comportement. Avant de poursuivre la description des autres aspects du framework, nous allons donner les résultats expérimentaux qui montrent la validité de notre approche. Enfin, nous évoquerons les travaux en cours et les perspectives de recherche ouvertes par l'Éthogénétique.*

### 11.1 Une expérience élémentaire

Les expériences que nous allons décrire ici ont d'abord pour but de démontrer la possibilité effective de construction de comportements adaptés à l'environnement dans un cadre classique (opérateurs génétiques fixés, évaluation explicite et individuelle de l'adéquation des agents à leur milieu, etc.).

En effet, les considérations théoriques qui nous ont amené à définir le modèle ATNoSFERES, et plus particulièrement la construction du phénotype à partir du génotype, devaient être mises à l'épreuve au moins sur un problème-jouet pour valider la pertinence de nos hypothèses.

#### 11.1.1 Description de l'expérience

##### Problème posé

L'environnement est un espace simulé discret contenant une lampe qui change aléatoirement de couleur (de vert à rouge et vice-versa) avec à chaque pas de temps une probabilité  $p = 0,05$ .

On définit une sous-classe `Pieton` de `ATNAgent`, disposant des lexèmes de condition et d'action décrits dans le tableau 11.1. Le comportement souhaité

## ◇11. EXPÉRIENCES ET EXTENSIONS

consiste à aller à droite lorsque la lampe est verte, et à gauche quand elle est rouge – problème trivial, mais au moyen duquel nous pourrions néanmoins illustrer quelques spécificités de notre approche (PICAULT et LANDAU 2001a; PICAULT et LANDAU 2001c).

TAB. 11.1 – *Lexèmes d’actions et de conditions de la classe Pieton.*

Actions		Conditions	
N!	aucune action		
R!	aller à droite	g?	vrai si la lampe est verte
L!	aller à gauche	r?	vrai si la lampe est rouge
U!	aller en haut	rand?	vrai avec une probabilité $p' = 0,5$
D!	aller en bas		

Le code génétique associé à la classe `Pieton` comporte :

- les 11 lexèmes de l’interpréteur (cf. tableau 10.2 page 181) ;
- les 8 lexèmes d’action/condition de `Pieton`.

Il faut donc au minimum 32 codons (soit des codons de 5 bits). Dans les expériences présentées ici, un code génétique de 32 codons a été construit automatiquement à partir de la liste de tous les lexèmes (sans souci particulier de robustesse vis-à-vis des mutations). Chaque agent dispose par ailleurs d’un chromosome *aléatoire*, et d’une *taille aléatoire*.

### Évaluation et sélection

On applique une sélection darwinienne sur une population composée de ces individus en les évaluant dans les mêmes conditions environnementales.

Tous les agents sont placés ensemble dans l’environnement pendant 100 pas de temps (il n’y a aucune interaction entre les individus, il s’agit seulement de paralléliser leur évaluation en leur appliquant les mêmes conditions). À chaque pas de temps, chacun recherche un arc à franchir dans son ATN, et quand c’est possible exécute une liste d’actions qui contribuent à sa notation :

$$\left\{ \begin{array}{ll} +1 \text{ point} & \text{si le mouvement est conforme à la contrainte} \\ -1 \text{ point} & \text{s’il est inversé} \\ 0 & \text{dans les autres cas} \end{array} \right.$$

## 11.1. UNE EXPÉRIENCE ÉLÉMENTAIRE

Seul le *premier mouvement* effectivement réalisé pour un pas de temps reçoit une note (toutes les actions sauf **N!** comptent comme un mouvement).

Pour pouvoir donner une note représentative de l'adéquation d'un individu à son environnement (*fitness*), on procède à 10 évaluations indépendantes (ce qui permet de tenir compte d'un éventuel polyéthisme) et on retient leur *fitness* moyenne.

À chaque génération, la population comporte 100 agents. On calcule leur *fitness*, en fonction de laquelle on leur attribue une probabilité de reproduction et de décès. 30 nouveaux agents sont produits par croisement des chromosomes, puis 30 agents de l'ancienne population sont éliminés (la population reste donc constante).

Avant de calculer la *fitness* des individus d'une génération donnée, on leur fait subir des *mutations*. Pour ce faire, deux stratégies ont été envisagées :

- 1° tous les individus de la population (et pas seulement ceux de la nouvelle génération) subissent une mutation aléatoire de leur chromosome : les bits de la chaîne sont inversés avec une probabilité  $r$  (ici  $r = 1\%$ );
- 2° tous les individus subissent soit une mutation aléatoire (de taux  $r$ ), soit (dans  $p\%$  des cas) une insertion ou une délétion aléatoire d'un codon dans leur chromosome.

Cette dernière contrainte est destinée à montrer quelques spécificités de notre modèle : la plupart des approches dans lesquelles le génotype est une chaîne binaire l'utilisent pour coder des paramètres à des positions déterminées de la chaîne : dans ce cas il est hors de question de faire varier sa taille ni de décaler des bits.

### 11.1.2 Premiers résultats et analyse

Les agents étant initialisés sur le nœud *Start*, aucune action n'est effectuée durant le premier pas de temps : le parcours de l'ATN commence avec le choix d'un premier nœud. La plus haute valeur que peut prendre la *fitness* d'un agent au cours d'un test de 100 pas de temps est donc 99 (soit 99 mouvements corrects).

La figure 11.1 montre l'évolution moyenne de la *fitness*, calculée sur 10 expériences, dans la première situation (mutations uniquement avec un taux  $r = 1\%$ ). La figure 11.2 présente l'évolution moyenne de la *fitness* dans le

## ◇11. EXPÉRIENCES ET EXTENSIONS

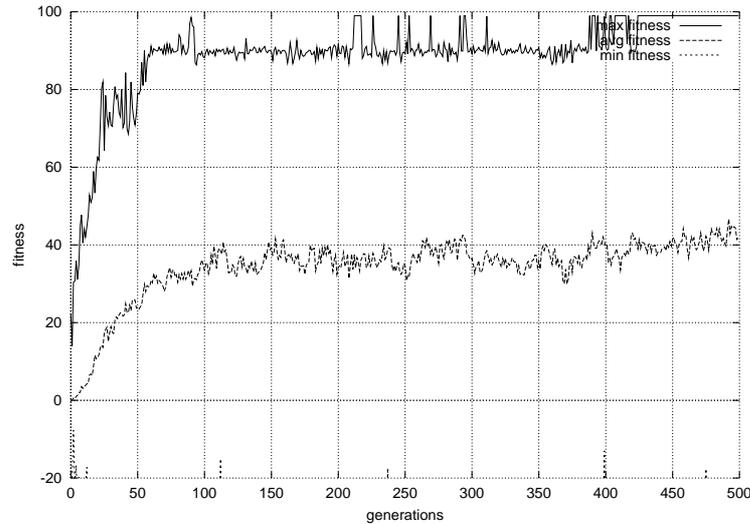


FIG. 11.1 – Évolution moyenne de la fitness (mutations ponctuelles seulement,  $r = 1 \%$ ). Chromosomes initiaux de 50 à 100 bits.

second cas, lorsque  $p = 20 \%$  des individus subissent des insertions ou délétions aléatoires de codons (au lieu de mutations classiques).

Dans quelques cas (notamment avec la première stratégie), aucune bonne solution n'a été trouvée avant les 500 générations qui limitent les expériences.

Bien que le problème à résoudre soit ici extrêmement simple, ces résultats montrent d'emblée que le modèle ATNoSFERES permet bien la construction évolutionniste de comportements à partir d'un substrat génétique de granularité faible (LANDAU et al. 2001b).

Le modèle choisi révèle un certain nombre de propriétés intéressantes. En particulier, il est possible d'effectuer des opérations d'insertion ou de délétion de codons sans nuire à la découverte d'une solution.

Cela n'est bien entendu possible qu'au moyen d'un codage *non positionnel*, ce qui est en grande partie le cas dans ATNoSFERES où les instructions ont un effet essentiellement *local* et où la fabrication de l'ATN par l'interpréteur fait appel à des mécanismes implicites pour assurer la cohérence du graphe.

Nous allons maintenant nous intéresser plus particulièrement aux spécificités de notre modèle en nous livrant à une analyse plus fine des comportements construits par la dynamique sélectionniste.

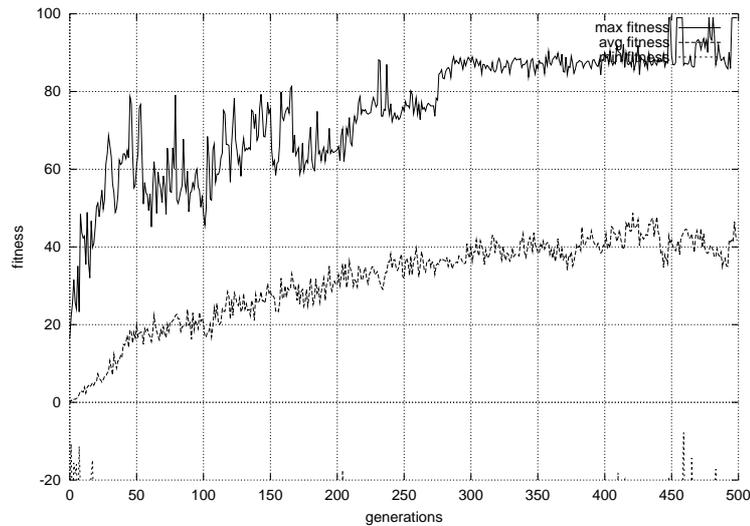


FIG. 11.2 – Évolution moyenne de la *fitness* (avec insertions et délétions aléatoires de codons dans  $p = 20\%$  des cas ; mutations sinon,  $r = 1\%$ ). Chromosomes initiaux de 50 à 100 bits.

### 11.1.3 Optimalité et adaptation

#### Adaptation et taille du chromosome

L'ATN décrit par la figure 11.3 page 190 est une solution « optimale » au problème au sens où il permet d'obtenir les 99 points de la *fitness* maximale avec la structure la plus simple : un seul nœud, deux arcs (si la lampe est verte, aller à droite ; si elle est rouge, aller à gauche).

Pour produire cet ATN, il suffit théoriquement de 35 bits seulement (en faisant un usage maximal du mécanisme de connexion implicite), par exemple pour coder la séquence de 7 lexèmes suivante :

*node - g? - R! - dupObject - L! - r? - dupObject*

Mais dans cette solution, l'*ordre* et la *nature* des lexèmes jouent un rôle crucial ; la séquence et le génotype correspondant sont donc particulièrement vulnérables aux mutations.

On peut donc raisonnablement penser que les individus vont avoir besoin, pour développer des comportements stables vis-à-vis des mutations, de chromosomes un peu plus longs que la solution « ingénieur ».

De plus, il est important que la taille du chromosome puisse varier en fonction des contraintes environnementales : leur changement peut deman-

◇11. EXPÉRIENCES ET EXTENSIONS

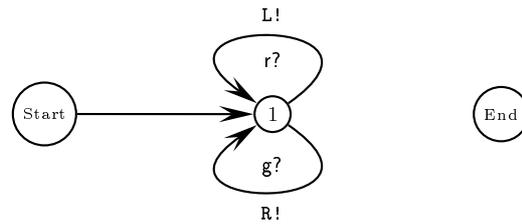


FIG. 11.3 – L’ATN « optimal » (donnant la fitness de valeur maximale avec la structure de graphe la plus simple) ; il peut être construit à partir de l’interprétation de 7 lexèmes seulement.

der l’élaboration d’un comportement plus complexe (donc demandant plus de nœuds, plus d’actions ou plus de conditions dans l’ATN) ou plus simple.

Toutefois, un problème nouveau peut apparaître avec l’allongement du chromosome. Le surcroît de matériel génétique disponible peut fort bien se révéler « nocif », c’est-à-dire donner naissance à des nœuds, des connexions supplémentaires qui peuvent *dégrader* le comportement en constituant une sorte d’« excroissance comportementale » parasite<sup>1</sup>. Autrement dit, l’adéquation du comportement de l’agent à son milieu peut être rendu plus difficile par un allongement des chromosomes.

Qu’en est-il exactement ? On peut formuler un début de réponse au vu des figures 11.4 à 11.7, réalisées alternativement avec les stratégies « mutations seules » ou « insertions-délétions », et pour des tailles de chromosomes initialement comprises entre 200 et 300 bits, puis entre 500 et 700 (soit 20 fois la taille d’un chromosome produit par ingénierie).

D’une part, les résultats (en termes de temps nécessaire pour atteindre la fitness maximale et de fitness moyenne dans la population) s’améliorent nettement lorsqu’augmente le nombre de bits dans le génome, et ce en dépit des difficultés induites par ce surplus d’information génétique. La combinatoire résultant de l’accroissement des chromosomes est donc régulée, biaisée, en faveur de comportements adaptés. Nous allons voir par quels mécanismes.

D’autre part, on voit se réduire la différence entre les deux stratégies de variation. Plus, non seulement les insertions et délétions ne sont pas léthales pour la population, mais on peut même dire qu’elles sont favorables puisqu’en fait, on observe aussi que le nombre de cas de non-convergence diminue dans le cas des insertions-délétions.

1. Cela se traduirait par un ATN très compliqué dans lequel une « bonne solution » se trouverait entremêlée à des portions de graphe inadaptées.

## 11.1. UNE EXPÉRIENCE ÉLÉMENTAIRE

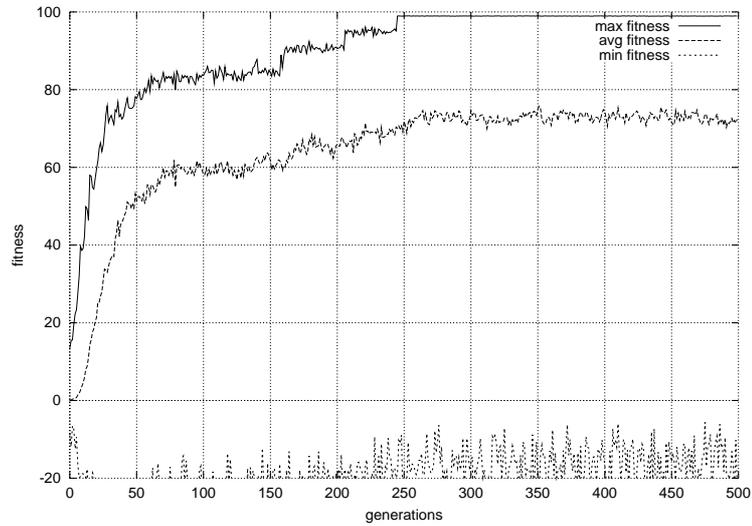


FIG. 11.4 – *Évolution moyenne de la fitness (avec mutations ponctuelles seulement,  $r = 1\%$ ). Chromosomes initiaux de 200 à 300 bits (moyenne sur 20 expériences).*

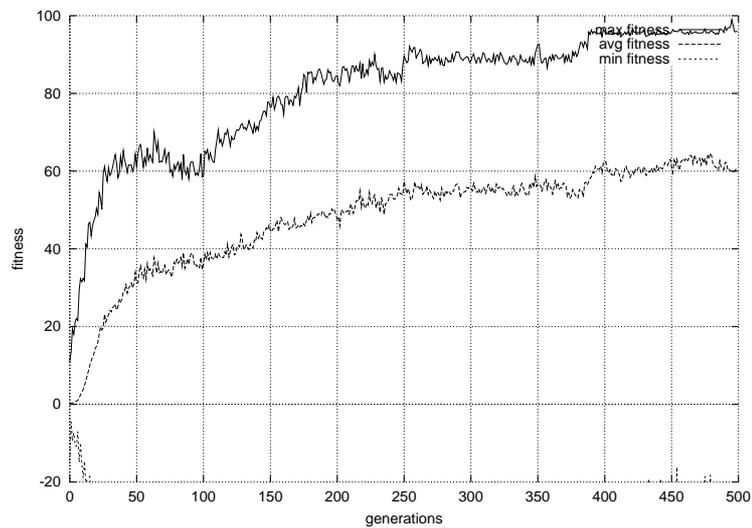


FIG. 11.5 – *Évolution moyenne de la fitness (avec insertions et délétions aléatoires de codons,  $r = 1\%$ ,  $p = 20\%$ ). Chromosomes initiaux de 200 à 300 bits (moyenne sur 20 expériences).*

◇11. EXPÉRIENCES ET EXTENSIONS

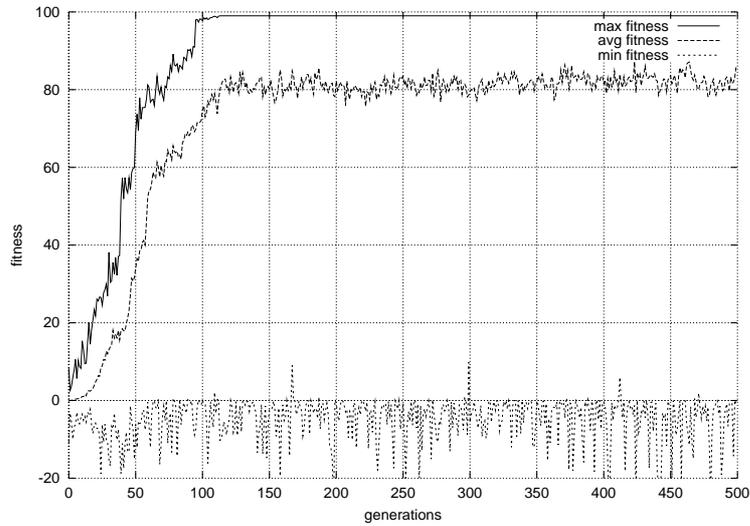


FIG. 11.6 – Évolution moyenne de la fitness (avec mutations ponctuelles seulement,  $r = 1\%$ ). Chromosomes initiaux de 500 à 700 bits (moyenne sur 10 expériences).

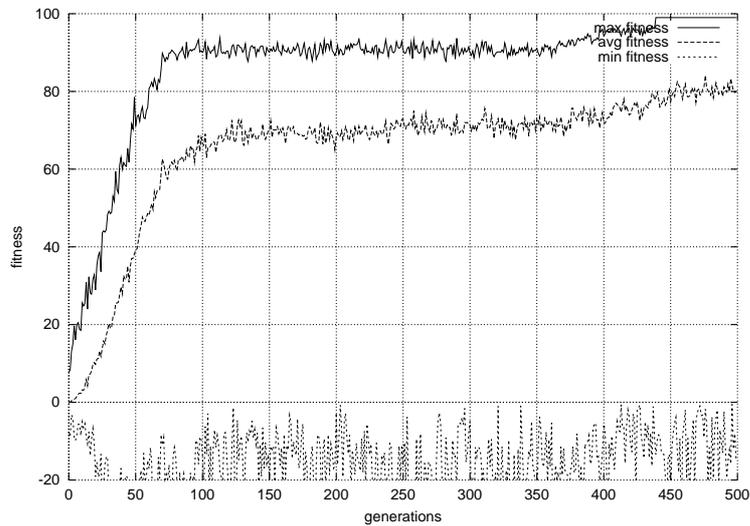


FIG. 11.7 – Évolution moyenne de la fitness (avec insertions et délétions aléatoires de codons,  $r = 1\%$ ,  $p = 20\%$ ). Chromosomes initiaux de 500 à 700 bits (moyenne sur 10 expériences).

### Stabilité génétique

En fait, lorsqu'on augmente la taille d'un chromosome, les individus sont soumis à une double pression sélective : celle issue de leur environnement « fonctionnel » bien sûr (la fonction de *fitness*), mais également une évaluation implicite de la *robustesse* de leur chromosome, c'est-à-dire de la manière dont leur comportement est *construit* (PICAULT et LANDAU 2001c).

Nous pouvons apporter deux arguments dans ce sens. Le premier consiste à observer ce qui se passe expérience par expérience, en particulier dans les cas d'insertions et de délétions, comme par exemple sur la figure 11.8 : certains individus ont trouvé assez rapidement un comportement adéquat, mais celui-ci ne s'est pas maintenu dans la population. Un Algorithme Génétique classique présente, au contraire le défaut de rester aisément piégé dans des « minima locaux » de la fonction de *fitness*. Ici, c'est l'inverse : un comportement ne peut se maintenir que s'il est *génétiquement stable* – la sélection n'opère pas seulement sur un phénotype « idéal » mais aussi sur la manière dont ce phénotype est construit. Mais cette stabilité est parfaitement atteignable, comme nous le constatons de fait.

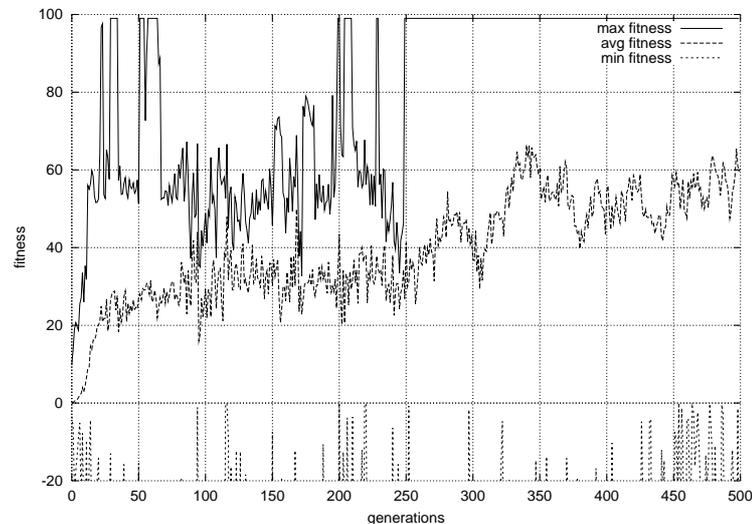


FIG. 11.8 – Évolution de la *fitness* (avec insertions et délétions aléatoires de codons,  $r = 1\%$ ,  $p = 20\%$ ). Chromosomes initiaux de 200 à 300 bits.

Le second argument vient de l'examen qualitatif des comportements adéquats produits. L'analyse des résultats expérimentaux montre que deux stratégies assez différentes peuvent émerger pour produire un comportement adéquat

## ◇11. EXPÉRIENCES ET EXTENSIONS

(donnant la *fitness* maximale) *et stable génétiquement*.

La première consiste à *construire un ATN simple* (proche de l'ATN « optimal » de la figure 11.3), en retardant, lors de l'expression génétique, la création des nœuds, c'est-à-dire en utilisant des lexèmes sans effet (par exemple en plaçant au début de la séquence de lexèmes des instructions de manipulation de la pile qui sont ignorées, la pile étant vide), ou encore en retirant des nœuds de la pile (par exemple avec le lexème *forget*). Ainsi, un des agents représentatifs ce genre de solution possède un chromosome de 207 bits, codant les 41 lexèmes suivants :

L! – dup – swap – popRoll – popRoll – forget – popRoll – swap – forget  
 – recall – recall – R! – L! – recall – rand? – L! – forget – startNode  
 – R! – dup – g? – dupObject – popRoll – pushRoll – L! – forget –  
 pushRoll – L! – r? – startNode – forget – dup – dupObject – r? – R!  
 – forget – dup – R! – dup – g? – popRoll

Ces lexèmes produisent un ATN presque identique à celui de la figure 11.3, avec un arc étiqueté ( $r?$ , L!L!R!R!) et l'autre ( $g?$ , R!R!). C'est un exemple intéressant de l'utilisation des *propriétés du langage* de construction de l'ATN pour aboutir à une structure simple *malgré* la complexité potentielle du graphe.

La seconde stratégie de stabilité génétique consiste à construire un ATN *dont seule une petite partie est réellement utilisée*.

Pour cela, plusieurs solutions sont envisageables : le graphe peut compter *plusieurs composantes connexes* (comme celui de la figure 11.9 page 195) ; il est également possible que les arcs partant des nœuds « utiles » (i.e. contribuant à un comportement adapté) vers des nœuds superflus soient étiquetés par des conditions irréalisables (telles que la conjonction  $\{g?,r?\}$ ).

Ce type d'ATN conduit donc exactement au même comportement que la solution « optimale », mais avec cette fois une utilisation des *propriétés du graphe* lui-même.

On comprend mieux, ainsi, le rôle positif des insertions et délétions aléatoires de codons.

En premier lieu, ces opérations ne sont pas en elles-mêmes plus nuisibles que de simples mutations, puisque notre langage de construction d'ATN ne dépend que faiblement et localement de l'ordre des instructions.

Mais en outre, elles peuvent même *faciliter* l'adaptation aux pressions sélectives. En effet, si l'on s'intéresse à l'évolution de la taille des chromosomes *au*

## 11.1. UNE EXPÉRIENCE ÉLÉMENTAIRE

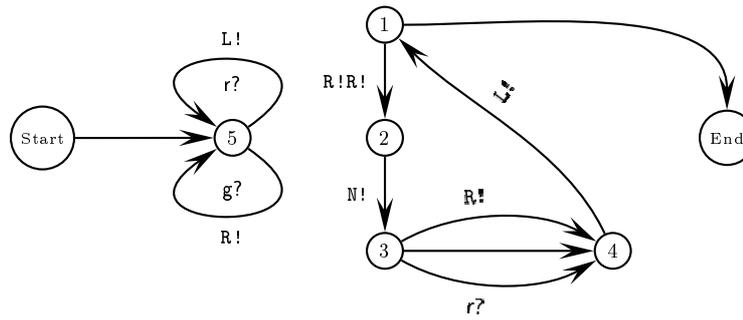


FIG. 11.9 – Un ATN construit par sélection naturelle, donnant un comportement adéquat au moyen de deux composantes connexes, dont une seule est accessible depuis Start. Cet ATN a été construit dans une expérience avec insertions-délétions (chromosomes initiaux de 200 à 300 bits).

cours des expériences, on constate une augmentation (autrement dit, il existe un biais sélectif *en faveur* des chromosomes plus longs), qui est d'autant plus nette que les chromosomes de départ sont courts (cf. figures 11.10 à 11.12). Il semble, pour les populations à chromosomes longs, qu'on obtienne une stabilisation autour de 650 bits (voir notamment l'expérience de la figure 11.13).

Ces expériences, en dépit de la simplicité du problème à résoudre et peut-être justement grâce à cette simplicité, permettent de mettre en lumière un certain nombre de caractéristiques du modèle ATNoSFERES qui se conforment aux prescriptions de l'Éthogénétique :

- l'ATN permet d'exprimer n'importe quel comportement réalisable à partir des capacités de perception et d'action ;
- il constitue une structure modulaire dont la complexité s'adapte automatiquement au problème ;
- les comportements sont immédiatement lisibles sur les ATN ;
- ils sont robustes vis-à-vis des mutations affectant le chromosome ;
- les mutations survenant durant l'existence des individus n'affecte que modérément leur comportement, une fois atteint un certain degré de stabilité génétique.

Le dernier prérequis qui n'apparaît pas ici concerne l'*homogénéité* du système, c'est-à-dire un fonctionnement reposant entièrement sur des agents. Nous allons donc en donner un aperçu à travers une description de quelques autres classes centrales du framework, et notamment celles qui permettent d'intégrer le modèle à une perspective multi-agent, à plusieurs échelles.

◇11. EXPÉRIENCES ET EXTENSIONS

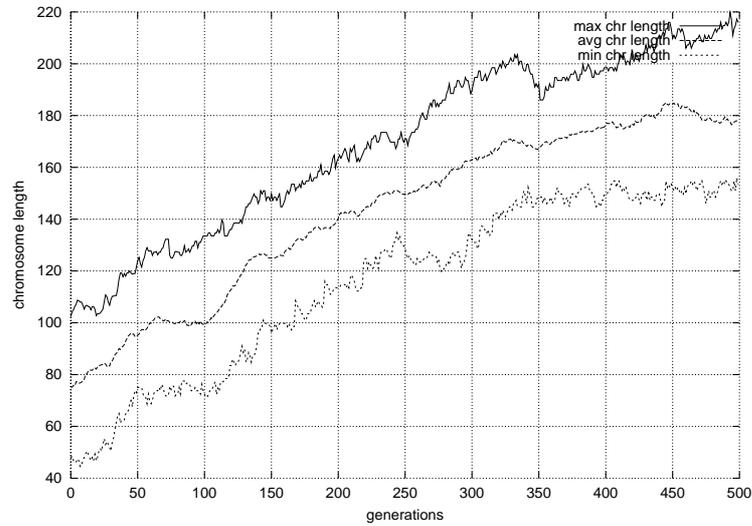


FIG. 11.10 – Évolution moyenne de la taille des chromosomes (avec insertions et délétions aléatoires de codons,  $r = 1 \%$ ,  $p = 20 \%$ ). Chromosomes initiaux de 50 à 100 bits (moyenne sur 10 expériences).

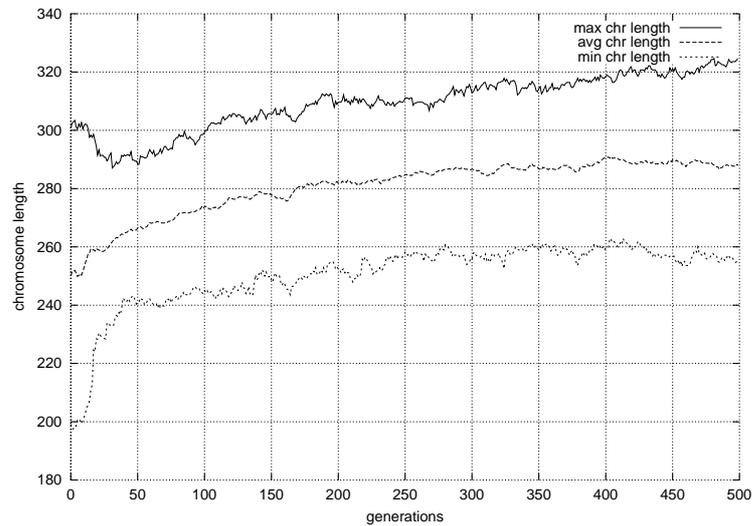


FIG. 11.11 – Évolution moyenne de la taille des chromosomes (avec insertions et délétions aléatoires de codons,  $r = 1 \%$ ,  $p = 20 \%$ ). Chromosomes initiaux de 200 à 300 bits (moyenne sur 20 expériences).

## 11.1. UNE EXPÉRIENCE ÉLÉMENTAIRE

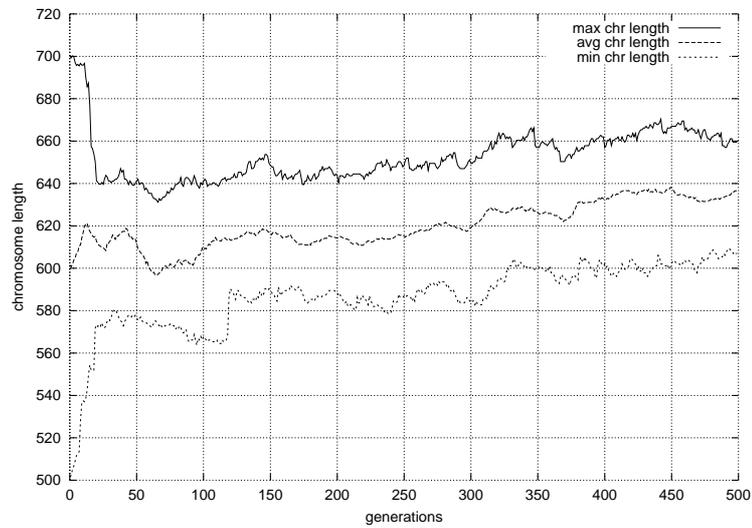


FIG. 11.12 – Évolution moyenne de la taille des chromosomes (insertions et délétions aléatoires de codons,  $r = 1\%$ ,  $p = 20\%$ ). Chromosomes initiaux de 500 à 700 bits (moyenne sur 10 expériences).

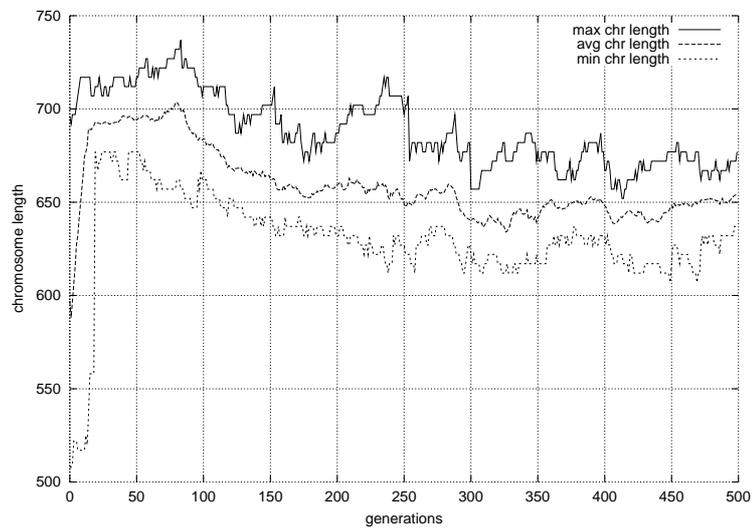


FIG. 11.13 – Évolution de la taille des chromosomes au cours d'une expérience, montrant un cas de diminution (avec insertions et délétions aléatoires de codons,  $r = 1\%$ ,  $p = 20\%$ ). Chromosomes initiaux de 500 à 700 bits.

## 11.2 Extensions élémentaires du modèle

### 11.2.1 La classe `FuzzyATNAgent`

La classe `FuzzyATNAgent` est une sous-classe directe d'`ATNAgent`. Elle permet de traiter des lexèmes de condition renvoyant soit des valeurs booléennes, soit des valeurs numériques.

Les arcs  $\{A_{ij}\}$  partant du nœud  $N_i$  correspondant à l'état de l'agent sont associés à leur poids  $\varpi_{ij} = \prod_k c_{ijk}$  où  $c_{ijk}$  dépend de la valeur retournée par un lexème de condition  $c_k?$  étiquetant  $A_{ij}$ .

$$c_{ijk} = \begin{cases} c_k? & \text{si c'est un nombre} \\ 0 \text{ ou } 1 & \text{si c'est un booléen} \end{cases}$$

Si tous les arcs ont un poids nul, l'agent reste bloqué sur  $N_i$ . Sinon, il procède à un choix probabiliste entre tous les arcs (de poids non nul) :

$$p(A_{ij}) = \frac{\varpi_{ij}}{\sum_j \varpi_{ij}}$$

Cette méthode permet de relâcher les contraintes imposées aux conditions et de faire jouer plus finement une dépendance entre les perceptions de l'agent et son comportement.

### 11.2.2 La classe `CompositeAgent`

Un agent composite est lui-même constitué d'agents dont il peut contrôler la création, l'activation, la mort, etc. (ses « organites »). Il dispose pour cela d'un « cytoplasme » (une instance de la classe `Environment` ou d'une de ses sous-classes) qui sert d'environnement à ses organites.

Par ailleurs, il dispose d'actions et de conditions génériques qui lui permettent de définir des lexèmes équivalents pour toutes les classes d'agents qu'il peut manipuler.

Par exemple, `create:` ou `exists:` donneront, en composition avec une classe `MonAgent` encapsulable dans l'agent composite, les lexèmes `create-MonAgent!` et `existsMonAgent?`.

Nous donnerons plus loin (§ 11.2.5) un exemple d'utilisation d'agents composites.

### 11.2.3 Encapsulation du traducteur

La classe `TranslatorAgent` encapsule un traducteur dans un `ATNAgent`. Elle est prévue pour fonctionner au sein d'un `BuilderAgent` (décrit § 11.2.5).

Pour fonctionner, le `TranslatorAgent` a besoin de trouver dans son environnement un chromosome et un code génétique (celui de la classe pour laquelle il effectue la traduction), c'est-à-dire un tableau de symboles de taille  $2^n$ . Il commence alors la traduction en plaçant dans l'environnement les lexèmes décodés.

Les actions et conditions spécifiques à un `TranslatorAgent` sont récapitulées dans le tableau 11.3, et un ATN « par défaut » pour décrire le comportement des agents de cette classe est présenté figure 11.14.

TAB. 11.2 – *Attributs de la classe TranslatorAgent.*

attribut	type ou classe	rôle
<code>currentLocus</code>	Integer	position courante sur le chromosome à traduire
<code>translator</code>	Translator	le traducteur encapsulé
<code>target</code>	Chromosome	le chromosome à traduire (recherché dans l'environnement)
<code>geneticCode</code>	Array	code génétique permettant le décodage du chromosome (tableau de symboles de taille $2^n$ )

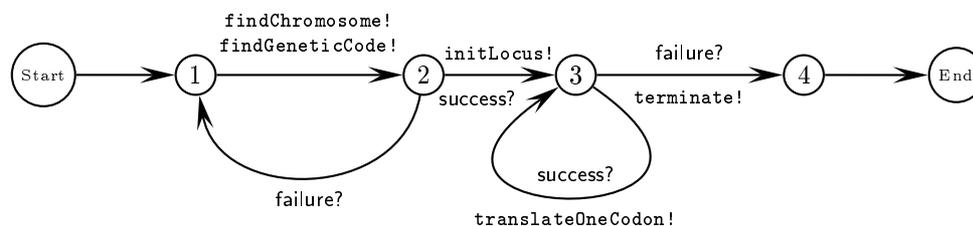


FIG. 11.14 – *L'ATN « par défaut » de la classe TranslatorAgent.*

## ◇11. EXPÉRIENCES ET EXTENSIONS

TAB. 11.3 – *Lexèmes d’actions et de conditions de la classe `TranslatorAgent`.*

<b>Actions</b>	
<code>findChromosome!</code>	recherche un chromosome ( <code>target</code> ) dans l’environnement (le chromosome <i>reste présent dans l’environnement</i> ); si échec, provoque <code>#failure</code>
<code>findGeneticCode!</code>	recherche un code génétique ( <code>geneticCode</code> ) dans l’environnement; si échec, provoque <code>#failure</code>
<code>initLocus!</code>	se place sur le premier locus libre du chromosome <code>target</code> ; si échec, provoque <code>#failure</code>
<code>terminate!</code>	termine la traduction
<code>translateOneCodon!</code>	traduit le codon commençant à <code>currentLocus</code> en un lexème, qui est placé dans l’environnement; si échec (plus de codon à traduire), provoque <code>#failure</code>

<b>Conditions</b>	
<code>success?</code>	vrai si l’agent est dans l’état <code>#success</code>
<code>failure?</code>	vrai si l’agent est dans l’état <code>#failure</code>

### 11.2.4 Encapsulation de l’interpréteur

La classe `InterpreterAgent` encapsule un interpréteur dans un `ATNAgent`. Elle est prévue pour être utilisée au sein d’un `BuilderAgent` (cf. § 11.2.5).

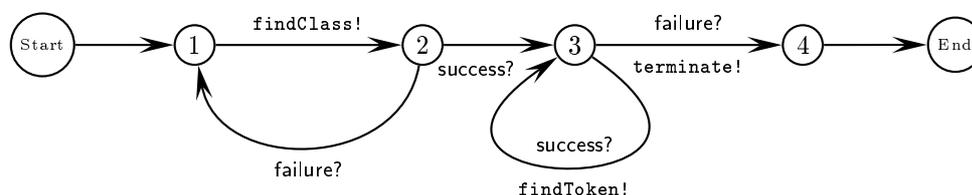
L’`InterpreterAgent` a besoin dans son environnement d’un identifiant de classe d’agent (`client`), celle dont il doit construire une instance. Il fabrique alors un ATN à partir des lexèmes présents dans son environnement.

Lorsqu’il n’en trouve plus, il achève la construction de l’ATN (connexions implicites, vérification de cohérence, etc.), crée une instance de la classe `client` dotée de cet ATN, et la place dans son environnement.

Les actions et conditions manipulées par un `InterpreterAgent` sont récapitulées dans le tableau 11.5, et un ATN « par défaut » pour les agents de cette classe est présenté figure 11.15.

TAB. 11.4 – *Attributs de la classe InterpreterAgent.*

attribut	type ou classe	rôle
currentToken	Symbol	lexème à interpréter (recherché dans l'environnement)
interpreter	Interpreter	l'interpréteur encapsulé
client	ATNAgent Class	classe pour laquelle est effectuée l'interprétation

FIG. 11.15 – *L'ATN « par défaut » de la classe InterpreterAgent.*

### 11.2.5 La classe BuilderAgent

Dans ATNoSFERES, il est possible d'*agentifier* le processus de traduction/interprétation grâce aux agents composites. C'est le rôle de `BuilderAgent`, sous-classe de `CompositeAgent` qui encapsule deux instances des classes `TranslatorAgent` et `InterpreterAgent` décrites précédemment.

Un `BuilderAgent` cherche dans son environnement un identifiant de classe d'agent, un code génétique, et un chromosome à partir duquel il va construire une instance de la classe voulue.

Pour ce faire, il fait appel à des lexèmes d'action et de condition qui lui sont propres ainsi qu'à des lexèmes d'action et de condition génériques qui s'appliqueront aux classes de ses organites (cf. tableau 11.7 page 207), et exécute l'ATN décrit figure 11.16.

On notera que l'ATN présenté permet au `BuilderAgent` de fonctionner « en boucle » et de construire des agents dès qu'il trouve dans son environnement un chromosome et une classe.

◇11. EXPÉRIENCES ET EXTENSIONS

TAB. 11.5 – *Lexèmes d’actions et de conditions de la classe InterpreterAgent.*

<b>Actions</b>	
<b>findClass!</b>	prend dans l’environnement une sous-classe d’ATNAgent ( <b>client</b> ); si échec, provoque <b>#failure</b> , sinon envoie à <b>interpreter</b> les lexèmes d’action et de condition correspondants.
<b>findToken!</b>	prend dans l’environnement un lexème à interpréter; si échec, provoque <b>#failure</b> , sinon l’envoie à <b>interpreter</b> .
<b>terminate!</b>	achève la construction de l’ATN, crée une instance de <b>client</b> pourvue de cet ATN, la place dans l’environnement; si échec, provoque <b>#failure</b>
<b>Conditions</b>	
<b>success?</b>	vrai si l’agent est dans l’état <b>#success</b>
<b>failure?</b>	vrai si l’agent est dans l’état <b>#failure</b>

La classe **BuilderAgent** permet d’illustrer deux caractéristiques importantes d’ATNoSFERES :

- 1° La multiplicité des échelles auxquelles peuvent opérer les agents : tout agent peut être remplacé par un agent composite, et ses comportements remplacés par l’action collective d’un SMA de granularité plus faible. Ce point permet de maintenir une approche incrémentale dans laquelle certaines parties du comportement et de l’organisation du système sont d’abord spécifiées, puis font l’objet d’une évolution.
- 2° L’*homogénéité* du système : toutes les entités capables de *comportement* sont agentifiées, *y compris les opérateurs génétiques et la machinerie assurant l’expression du phénotype.*

TAB. 11.6 – *Attributs de la classe BuilderAgent.*

attribut	type ou classe	rôle
client	ATNAgent Class	classe dont il faut construire une instance (recherchée dans l'environnement)
target	Chromosome	chromosome de l'agent à construire (recherché dans l'environnement)
cytoplasm	ListEnvironment	environnement de liste dans lequel opèrent les organites (instances des classes TranslatorAgent et InterpreterAgent)

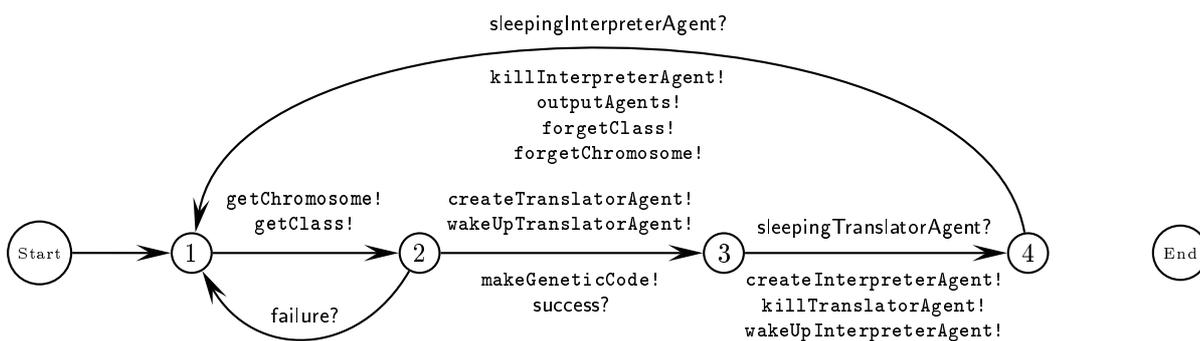


FIG. 11.16 – *L'ATN « par défaut » de la classe BuilderAgent.*

## 11.3 L'Éthogénétique et ATNoSFERES : quelques perspectives

### 11.3.1 Expériences en cours

Les principes du modèle ATNoSFERES ayant été validés, des expériences plus approfondies sont actuellement en cours. Elles visent à explorer les deux facettes complémentaires de l'Éthogénétique : la coévolution de comportements d'agents dans un écosystème d'une part, le développement de la corporité d'autre part.

#### Les écosystèmes : proies et prédateurs

Nous supervisons actuellement des expériences de coévolution entre proies et prédateurs, en vue d'une comparaison entre des stratégies obtenues au moyen d'ATNoSFERES et d'autres produites par des approches plus classiques (Programmation Génétique en l'occurrence).

Proies et prédateurs sont des agents placés dans un environnement toroïdal. Ils peuvent se déplacer à des vitesses variables, percevoir la présence et la nature d'autres agents dans un certain secteur angulaire, s'alimenter et se reproduire.

Plutôt que d'utiliser une fonction de *fitness* puis une sélection des individus les mieux notés, nous laissons la sélection opérer sur des critères comportementaux : les agents doivent maintenir un certain niveau de nourriture pour survivre, et ne se reproduisent que les agents qui disposent des comportements adéquats. Cela nécessite une phase d'amorçage lorsqu'on part de comportements aléatoires, mais il est également possible de commencer l'expérience avec des comportements réactifs de reproduction.

Ces expériences, qui font l'objet du stage de maîtrise de Julien POUDADE (mai–septembre 2001) permettront de préciser les caractéristiques multi-agents de nos modèles et d'explorer des problématiques liées à la coévolution, comme la question d'une « course à l'armement » entre proies et prédateurs (DAWKINS 1989; DAWKINS 1982) et pourront ensuite être complexifiées progressivement pour aborder d'autres problématiques liées aux écosystèmes.

### Développement : la navigation réactive

En parallèle, dans le cadre de sa thèse, Samuel LANDAU a entamé des expériences relatives à la navigation réactive, en commençant par la capacité à sortir d'un labyrinthe.

L'un des enjeux est de vérifier que les agents sauront « compiler » une solution pour la réduire à un ensemble comportemental minimal.

Le modèle ATNoSFERES sera ensuite utilisé pour gérer la navigation réactive et l'évitement d'obstacles sur les robots du projet MICRobES ; cette application permettra une comparaison avec les approches déjà utilisées en Robotique Évolutionniste (NOLFI et FLOREANO 2000, chapitre 4).

Lorsque les caractéristiques de modularité du modèle seront confirmées plus amplement, il sera alors possible d'aborder incrémentalement la question de la proprioception.

#### 11.3.2 Vers une plus grande autonomie

D. FLOREANO a introduit la notion d'« Espace de Fitness<sup>2</sup> » pour caractériser la manière dont opère la sélection dans les divers paradigmes évolutionnistes (FLOREANO et URZELAI 2000).

La fonction de fitness présidant à la sélection se positionne selon les trois axes suivants :

- L'opposition *fonctionnelle-comportementale*. Une fitness fonctionnelle, lorsqu'elle n'est pas simplement une fonction au sens mathématique utilisée pour noter les individus, consiste à évaluer la manière dont l'architecture de contrôle fonctionne (présence d'oscillations spécifiées ou d'un certain motif reconnaissable), par opposition à une fitness comportementale qui consiste à évaluer le résultat des actions de l'individu.
- La dimension *explicite-implicite* : elle mesure le nombre de contraintes, de variables, de paramètres utilisés par la fitness, le cas extrême de fitness implicite correspondant simplement à l'opposition mort/survie-reproduction.
- L'axe *externe-interne* met en jeu la capacité de l'individu à calculer lui-même (fitness interne) ou non les variables et contraintes qui interviennent dans la fitness (et par suite concourent à sa sélection). Les

---

2. En anglais, *Fitness Space*.

## ◇11. EXPÉRIENCES ET EXTENSIONS

variables externes sont un point crucial dans le passage de la simulation au réel (un agent simulé peut tout calculer par lui-même, pas un robot).

Le positionnement d'une fonction de fitness suivant ces trois axes varie considérablement selon le point de vue adopté : une approche orientée ingénierie préférera une évaluation fonctionnelle, explicite et externe, tandis qu'une démarche inspirée par la Vie Artificielle passe par l'utilisation d'une fitness essentiellement comportementale, implicite et interne (NOLFI et FLOREANO 2000, chapitre 3).

En raison des prérequis de l'Éthogénétique, le modèle ATNoSFERES se prête facilement à la première approche.

Toutefois, parce qu'il est constitué entièrement d'agents, dont les comportements sont tous décrits par une structure d'ATN susceptible d'évoluer, ATNoSFERES permet également de donner au système de plus en plus d'autonomie en lui permettant de faire évoluer les opérateurs génétiques, les mécanismes d'expression génétique, de parcours de l'ATN, etc.

Cela ne se fera certainement pas d'emblée. Les êtres vivants, capables d'une autonomie gigantesque, ne font pas « n'importe quoi » : ils se sont dotés au cours de leur histoire de dispositifs de protection, de verrouillage, qui évitent de sombrer dans le chaos, c'est-à-dire la mort.

Pour un système informatique, le chaos se traduit au mieux par l'impossibilité de comprendre ce qui se passe dans le système, de donner sens à ses actions, et par suite de l'utiliser.

Aussi, l'introduction d'autonomie dans un modèle éthogénétique est possible, mais elle devra être progressive. Ce sera la prochaine étape dans nos recherches.

TAB. 11.7 – *Lexèmes d'actions et de conditions de la classe BuilderAgent.*

<b>Actions</b>	
getChromosome!	prend dans l'environnement un chromosome à traduire ( <b>target</b> ); si échec, provoque <b>#failure</b> , sinon le place dans cytoplasm.
getClass!	prend dans l'environnement une classe d'agent dont on veut construire des instances ( <b>client</b> ); si échec, provoque <b>#failure</b> , sinon la place dans cytoplasm.
makeGeneticCode!	lit le code génétique associé à la classe trouvée ( <b>client</b> ) et le place dans cytoplasm.
outputAgents!	recherche des agents dans cytoplasm (qu'un <b>InterpreterAgent</b> y a normalement placés), et les transfère dans son propre environnement.
forgetClass!	« oublie » la classe <b>client</b> .
forgetChromosome!	« oublie » le chromosome en le rejetant dans l'environnement.
<b>Conditions</b>	
success?	vrai si l'agent est dans l'état <b>#success</b>
failure?	vrai si l'agent est dans l'état <b>#failure</b>
<b>Actions génériques</b>	
create:	crée comme organite une instance de la classe passée en argument.
wakeUp:	fait passer (si possible) dans l'état <b>#ready</b> un organite instance de la classe passée en argument.
kill:	détruit si possible un organite instance de la classe passée en argument.
<b>Conditions génériques</b>	
sleeping:	indique si les organites instances de la classe passée en argument sont dans l'état <b>#asleep</b> .

## ◇11. EXPÉRIENCES ET EXTENSIONS

---

# CONCLUSION

---

*« Il se peut dire, avec apparence, qu'il y a une ignorance abécédaire, qui va devant la science ; une autre, doctorale, qui vient après la science : ignorance que la science fait et engendre, tout ainsi comme elle défait et détruit la première »*

MONTAIGNE. *Essais*, I.LIV



---

# Synthèse et perspectives

---

Les recherches que nous avons présentées dans cette thèse doivent être vues comme un point de départ, plutôt que comme un aboutissement. Au terme de notre itinéraire, il nous reste en définitive plus de questions que nous n'en avions au commencement. C'est aussi que ces questions ne sont plus les mêmes, preuve peut-être que notre travail aura quand même été fructueux.

## Des modèles de construction sociale

En résumé, nous avons proposé plusieurs modèles permettant à des groupes d'agents de se construire une organisation sociale adaptée à leurs besoins. Les premiers d'entre eux se sont montrés efficaces pour reproduire des comportements sociaux observés effectivement chez les primates, dont la socialité ressemble tant à celle de l'homme. Nous avons montré qu'ils permettaient une reconnaissance des relations entre agents et pouvaient être appliqués à des situations assez différentes – par exemple la construction collective d'un lexique.

Mais alors que nous pensions possible une transposition assez générale de ces modèles à des applications *situées* dans un environnement physique, dans le cadre du projet de robotique ouverte MICRobES, nous avons dû finalement constater que les contraintes spécifiques à une telle situation remettaient en question la pertinence d'une simple transposition. L'immersion d'un système informatique, distribué ou non, dans un objet physique lui-même placé dans un environnement complexe pose un problème qui ne relève plus de la simple ingénierie : il faut lui donner les caractéristiques adaptatives d'un être vivant.

Nous avons donc été amené à revoir nos hypothèses et à proposer un modèle reposant sur de nouvelles métaphores, elles aussi empruntées à l'éthologie, celles de *développement* et d'*évolution* au sein de SMA situés. Les expériences menées sur ce modèle (ATNoSFERES) en sont encore à leurs débuts et, comme nous l'avons signalé, il reste encore de nombreux points à explorer. Toutefois, nous avons déjà montré que nous pouvions obtenir les propriétés souhaitées.

## Le rôle de la métaphore

« Une science qui accepte les images est, plus que tout autre, victime des métaphores. Aussi l'esprit scientifique doit-il sans cesse lutter contre les images, contre les analogies, contre les métaphores. »

BACHELARD. *La formation de l'esprit scientifique*

Comment ! Mais, au cours de nos travaux, nous n'avons cessé d'utiliser des métaphores : métaphore des comportements de primates pour rechercher des modes d'organisation sociale intéressants ; métaphore des phéromones pour concevoir les modèles utilisés en simulation ; métaphore du corps pour les robots ; métaphore encore de l'organisme qui se développe ; métaphore toujours de l'écosystème qui évolue ; métaphore enfin de la synthèse des protéines et de la cellule. Et combien d'autres, que nous n'avons pas nous-même inventées, n'avons-nous pas rencontrées en chemin, à commencer par la première d'entre elles : *la métaphore de l'agent* !

Comment comprendre le mot de BACHELARD ? Nous sentons bien qu'il a raison, que la physique ne serait pas *la* science empirique par excellence si elle faisait usage de concepts flous et d'analogies, que certaines disciplines comme la biologie (avec sa métaphore de l'ADN-programme) peuvent effectivement s'enfermer dans des images erronées (ATLAN 1999; KUPIEC et SONIGO 2000) et gagneraient à leur abandon.

Mais alors, qu'est-ce qui nous autorise à penser qu'il en va différemment de l'IAD ?

Il nous semble que l'informatique a un statut ambivalent. *Elle relève à la fois des sciences empiriques et des mathématiques*. En effet, comme les mathématiques, elle *construit ses propres objets d'étude* ; mais en même temps elle étudie ces objets non pas sur leurs propriétés intrinsèques mais sur leur *comportement*, à travers une démarche expérimentale.

Aussi, lorsque l'informatique utilise des métaphores, elle ne se contente pas d'importer une image au contenu vague – ce que BACHELARD appelle une « rêverie ». Elle *construit un concept calculable* qu'elle baptise par analogie avec un phénomène naturel. La métaphore au sens informatique est donc avant tout une reconstruction du réel. Elle est manipulable à volonté, les mécanismes qui la sous-tendent (algorithme) sont explicites, elle permet d'abord de donner un

sens intuitif à un modèle formel destiné à se déployer par le calcul. C'est à ce titre qu'elle constitue un outil fécond et non comme dans d'autres disciplines un obstacle épistémologique.

Au reste, BACHELARD critique avant tout l'usage *explicatif* des métaphores : « *on ne se borne pas à décrire par un mot, on explique par une pensée* » (*op. cit.*) et n'est pas hostile à leur usage « mathématisé » : « *La résistance électrique est une résistance épurée par une définition précise, elle est incorporée dans une théorie mathématique qui en limite toute extension abusive* » (*op. cit.*). L'informatique peut à l'évidence parvenir aux mêmes résultats.

## Un « discours de la méthode »

« *Franz Kafka l'a dit avant nous : il y a un but, mais il n'y a aucun parcours ; nous nommons parcours nos dubitations* »

GEORGES PEREC. *La disparition*

Plus encore que les modèles informatiques, c'est sans doute la *démarche méthodologique* qui a évolué au cours de nos travaux. Nous pouvons distinguer quelques principes généraux qui restent féconds en toutes circonstances : *parcimonie, incrémentalité*, et *démarche expérimentale*.

Cela tient au fait que l'IAD telle que nous l'avons abordée (c'est-à-dire essentiellement dans une perspective constructiviste) apparaît bien comme une science empirique en rapports étroits avec l'Éthologie. Le lien entre IAD et Éthologie est triple. Il tient d'abord évidemment à un *objet d'étude* commun : le *comportement* d'entités soit naturelles, soit artificielles, dans leur environnement.

Il est ensuite un *dialogue interdisciplinaire* qui est à l'origine de la définition d'un « cercle vertueux » (DROGOUL 2000) des recherches sur les SMA, qui consiste à rechercher dans les systèmes naturels des phénomènes évoquant des problématiques d'IAD, concevoir des modèles pour les simuler, transposer ces modèles à d'autres applications distribuées, et enfin réutiliser les nouveaux concepts informatiques pour l'étude des phénomènes naturels.

Nous avons montré que la *transposition* n'est possible que si un certain nombre d'hypothèses implicites restent vérifiées dans les deux situations. Cela

ne remet pas en cause le « cercle vertueux » en lui-même, en tant que dynamique de confrontation et d'échange des deux disciplines.

Enfin, le lien est également *épistémologique et méthodologique*. Les schémas explicatifs éthologiques (cf. tableau 2.1 page 19) sont de puissants outils car ils permettent de passer d'une interprétation *finaliste et intentionnelle* d'un comportement à ses *causes effectives*, et vice-versa. A. DROGOUL a déjà souligné comment cette dialectique entre *mécanisme* et *fonction* pouvait être exploitée d'un point de vue méthodologique pour la conception d'applications multi-agents (DROGOUL 2000).

Nous avons pour notre part montré l'intérêt des schémas « historiques » de l'Éthologie (reposant sur les concepts d'ontogenèse et de phylogenèse) pour la « conception automatique » de systèmes adaptés à un environnement complexe. Bien souvent, le choix d'une architecture (réactive ou cognitive) tient plus à l'appartenance à une école qu'à un besoin justifié. L'utilisation d'un processus sélectionniste place le concepteur à un tout autre niveau : ce sont les agents eux-mêmes qui, progressivement, vont adopter une architecture *qui répondra à une nécessité*.

Il serait toutefois réducteur de ne voir dans la dialectique IAD-Éthologie qu'un intérêt méthodologique. En fait, la physique, puis la biologie, se sont construites en s'arrachant à *une vision aristotélicienne du monde*, processus qui n'est pas encore tout à fait achevé (KUPIEC et SONIGO 2000). Le cosmos aristotélicien, s'exprimant en termes téléologiques, avait comme toute interprétation intentionnelle l'avantage d'être plus intuitif, plus directement intelligible qu'un monde galiléen reposant sur des lois universelles.

À l'inverse, la modélisation par objets (PERROT 1998) et *a fortiori* par agents recrée un monde aristotélicien. Mais celui-ci n'existe que dans la mesure où nous le construisons : le double schéma mécanisme/fin permet alors de concilier deux visions du monde. Simuler des phénomènes naturels au moyen d'agents, c'est leur donner une intelligibilité nouvelle (la fonction du comportement) tout en garantissant leur ancrage empirique (les causes efficientes de ce comportement).

## Vers une informatique non-déterministe?

La question de l'organisation d'un système est l'une des plus complexes qui se pose depuis quelques années aux sciences empiriques et humaines : physique, biologie, sociologie, économie... et depuis moins longtemps à l'informatique. Pourtant, avec le mouvement qui a progressivement conduit à l'informatique répartie et aux SMA, on a vu s'accroître l'imprévisibilité des systèmes. Pendant des années l'informatique a cherché à s'en affranchir, par un recours assez systématique à la Spécification et autres méthodes contribuant à *figer* le comportement d'un système en vue d'un nombre restreint de fonctions.

Pourtant, il est vraisemblable qu'il faudra à un moment ou à un autre se donner les moyens de comprendre le fonctionnement de systèmes non déterministes : d'abord, parce c'est une tendance forte de l'informatique, notamment en raison du développement d'applications « situées », en interaction les unes avec les autres dans des environnements imprévisibles ; mais aussi parce que c'est sans doute le seul moyen d'aborder la modélisation de phénomènes naturels complexes (CARDON 1999).

Nous espérons que des modèles comme ceux issus de l'Éthogénétique contribueront à cette évolution, à travers la recherche expérimentale des moyens de donner aux systèmes informatiques une autonomie toujours croissante.

## Perspectives de recherche

Les modèles et principes que nous avons définis, mis en œuvre et expérimentés nous amènent à nous pencher sur les orientations à donner désormais à nos travaux.

L'Éthogénétique leur fournira sans doute le cadre le plus fructueux dans les prochaines années, au prix certainement d'une étude plus précise de ses caractéristiques et propriétés. Issue du projet MICRobES, elle a naturellement vocation à être expérimentée dans le cadre de la Robotique Collective Ouverte, pour poursuivre les travaux sur la socialité et la corporéité.

Ceci dit, il serait envisageable de l'appliquer à nos travaux de simulation : ainsi, pour poursuivre la dynamique générale du « cercle vertueux », nous pourrions par exemple nous interroger sur la phylogénèse des comportements sociaux chez les primates et étudier, au moyen d'ATNoSFERES, les conditions de leur apparition.

## SYNTHÈSE ET PERSPECTIVES

D'une manière plus générale, l'Éthogénétique concilie deux paradigmes souvent distincts : d'un côté, les algorithmes évolutionnistes qui permettent d'aborder la résolution de problèmes et l'adaptation de comportements ; de l'autre, l'approche distribuée, multi-agent, approche privilégiée pour la simulation dans les sciences de la vie.

Ces deux voies sont complémentaires et nous devons les explorer en parallèle au cours des prochaines années pour approfondir et élargir les questionnements esquissés dans cette thèse.

---

## Références bibliographiques

---

- ACKLEY, D. H. ET LITTMAN, M. L. (1991). Interactions between learning and evolution. In (LANGTON et al. 1991).
- ADAMI, C., BELEW, R., KITANO, H., ET TAYLOR, C., éditeurs (1998). *Artificial Life*, volume 6. MIT Press, Cambridge, Massachusetts.
- AGRE, P. E. ET CHAPMAN, D. (1987). Pengi: an implementation of a theory of activity. In *Proceedings of AAAI'87*, pp. 268–272.
- ALTMANN, S. A., éditeur (1967). *Social communication among primates*. University of Chicago Press, Chicago.
- ARKIN, R. C. (1998). *Behavior-Based Robotics*. MIT Press, Cambridge, Massachusetts.
- ARKIN, R. C. ET BALCH, T. (1997). *Cooperative Multiagent Robotic Systems*.
- ASADA, M., KUNIYOSHI, M., DROGOUL, A., ASAMA, H., MATARIC, M., DUHAUT, D., STONE, P., ET KITANO, H. (1998). The RoboCup Physical Agent Challenge: Phase-I. *Applied Artificial Intelligence (AAI) Journal*, 12(2–3):251–265.
- ATLAN, H. (1979). *Entre le cristal et la fumée. Essai sur l'organisation du vivant*. Éditions du Seuil, Paris.
- ATLAN, H. (1999). *La fin du « tout génétique » ? Vers de nouveaux paradigmes en biologie*. Sciences en questions. INRA Éditions, Paris.
- AXELROD, R. (1996). *Comment réussir dans un monde égoïste ?* Collection « Opus ». Odile Jacob, Paris.
- AXELROD, R. ET HAMILTON, W. D. (1981). The evolution of cooperation. *Science*, 211:1390–1396.
- BACHELARD, G. (1938). *La formation de l'esprit scientifique. Contribution à une psychanalyse de la connaissance*. Librairie philosophique J. Vrin, Paris.
- BÄCK, T. ET SCHWEFEL, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23.

## BIBLIOGRAPHIE

- BACON, F. (1620). *Novum Organum*. Réédition aux Presses Universitaires de France de 1986.
- BARLOW, G. ET SILVERBERG, J., éditeurs (1980). *Sociobiology: Beyond Nature/Nurture*. Westview Press, Boulder, Colorado.
- BARON-COHEN, S., LESLIE, A. M., ET FRITH, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21:37–46.
- BARTHÈS, J.-P., CHEVRIER, V., ET BRASSAC, C., éditeurs (1998). *Systèmes Multi-Agents – de l’interaction à la socialité. Actes des JFIADSMA’98*. Hermès, Paris.
- BEASLEY, D., BULL, D. R., ET MARTIN, R. R. (1993). An overview of genetic algorithms: Part 1, fundamentals. *University Computing*, 15(2):58–69.
- BEDEAU, M., MCCASKILL, J., PACKARD, N., ET RASMUSSEN, S., éditeurs (2000). *Artificial Life*, volume 7. MIT Press, Cambridge, Massachusetts.
- BILLARD, A. (2000). Play, Dreams and Imitation in Robota. In (DAUTENHAHN 2000c), pp. 9–12.
- BILLARD, A., DAUTENHAHN, K., ET HAYES, G. (1998). Experiments on human-robot communication with Robota, an imitative learning and communicating doll robot. *Workshop on Socially Situated Intelligence, in SAB’98. Technical report of Centre for Policy Modelling, Manchester Metropolitan University*.
- BOBBÉ, S., éditeur (1998). *Grands singes. La fascination du double*, volume 106 de *Monde/Nature extrême*. Éditions Autrement, Paris.
- BOESCH, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour*, 41:530–532.
- BOESCH, C. (1994). The question of culture. *Nature*, 379:207–208.
- BOESCH, C. ET BOESCH, H. (1989). Hunting behaviour of wild chimpanzees in the Taï National Park. *American Journal of Physical Anthropology*, 78:547–73.
- BOESCH, H. (1983). Optimization of nut cracking with natural hammers by wild chimpanzees. *Behaviour*, 83:265–286.
- BOESCH, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica*, 54:86–99.
- BONABEAU, É., DORIGO, M., ET THERAULAZ, G. (1999a). L’intelligence en essaim. In (GLEIZES et MARCENAC 1999), pp. 25–38.
- BONABEAU, É. ET THERAULAZ, G., éditeurs (1994). *Intelligence collective*. Hermès, Paris.

- BONABEAU, É., THERAULAZ, G., ET DENEUBOURG, J.-L. (1999b). Dominance Orders in Animal Societies: The Self-organization Hypothesis Revisited. *Bulletin of Mathematical Biology*, 61(4):727–757.
- BOURJOT, C., CHEVRIER, V., BERNARD, A., ET KRAFFT, B. (1999). Coordination par le biais de l'environnement: une approche biologique. In (GLEIZES et MARCENAC 1999), pp. 237–250.
- BOURON, T. (1993). *Structures de communication et d'organisation pour la communication dans un univers multi-agent*. Thèse de doctorat, Université Paris VI.
- BRAZIER, F., VAN STEEN, M., ET WIJNGAARDS, N. (2001). Distributed Shared Agent Representations. In (MAŘÍK et al. 2001b), pp. 230–237.
- BREAZEL, C. ET FITZPATRICK, P. (2000). That Certain Look: Social Amplification of Animate Vision. In (DAUTENHAHN 2000c), pp. 18–22.
- BREAZEL, C. ET SCASSELLATI, B. (1998). Infant-like Social Interactions Between a Robot and a Human Caretaker. In DAUTENHAHN, K., éditeur, *Adaptive Behavior*, Special issue on Simulation Models of Social Agents.
- BREDÈCHE, N. ET ZUCKER, J.-D. (2000). From distributed robot perception to human topology: a learning model. In (PARKER 2000).
- BROOKS, R. A. (1991). Intelligence without Reason. In MYLOPOULOS, J. ET REITER, R., éditeurs, *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI'91)*, pp. 569–595. Morgan Kaufmann.
- BROOKS, R. A. (1996). Behavior-based humanoid robotics. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'96)*, pp. 1–8.
- BROOKS, R. A. (1999). *Cambrian Intelligence*. MIT Press, Cambridge, Massachusetts.
- BROOKS, R. A. ET MAES, P., éditeurs (1994). *Artificial Life*, volume 4. MIT Press, Cambridge, Massachusetts.
- BURGARD, W., CREMERS, A. B., FOX, D., HÄHNEL, D., LAKEMEYER, G., SCHULZ, D., STEINER, W., ET THRUN, S. (1999). Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*.
- BYRNE, R. W. (1993). A formal notation to aid analysis of complex behaviour: Understanding the tactical deception of primates. *Behaviour*, 127(3–4):231–246.

## BIBLIOGRAPHIE

- BYRNE, R. W. (1995). *The Thinking Ape: Evolutionary Origins of Intelligence*. Oxford University Press.
- BYRNE, R. W. ET WHITEN, A. (1988a). *Machiavellian Intelligence: social expertise and the evolution of intellect in monkeys, apes and humans*. Clarendon Press, Oxford.
- BYRNE, R. W. ET WHITEN, A. (1988b). Tactical deception in primates. *The Behavioral and Brain Sciences*, 11:233–273.
- CALLAOS, N., ESQUIVEL, S., ET BURGE, J., éditeurs (2001). *Proceedings of the 5th World Multi-Conference on Systemics, Cybernetics and Informatics (SCI'01)*, volume III. ISBN: 980-07-7543-9.
- CANAÑAMERO, L. D. ET FREDSLUND, J. (2000). How Does It Feel? Emotional Interaction with a Humanoid LEGO Robot. In (DAUTENHAHN 2000c), pp. 23–28.
- ČAPEK, K. (1920). *R. U. R. (Reson's Universal Robots)*. Collection « Regards croisés ». Éditions de l'aube. Traduit du tchèque par Jan Rubeš, 1997.
- CARDON, A. (1999). *Conscience artificielle & systèmes adaptatifs*. Eyrolles, Paris.
- CARDON, A. (2001a). The Approaches of the Concept of Embodiment for an Autonomous Robot. Towards Consciousness of its Body. In (MAŘÍK et al. 2001b), pp. 218–229.
- CARDON, A. (2001b). Étude de la conception et du comportement d'une organisation d'agents massive. Article en préparation.
- CARLE, P., BENHAMOU, P., ORNATO, M., ET DOLBEAU, F.-X. (1998). La reconnaissance d'intentions comme dynamique des organisations. In (BARTHÈS et al. 1998), pp. 311–326.
- CASTELFRANCHI, C. (1990). Social Power: a Point Missed in Multi-Agent, DAI and HCI. In (DEMAZEAU et MÜLLER 1990).
- CASTELFRANCHI, C., MICELI, M., ET CESTA, A. (1992). Dependence Relations among Autonomous Agents. In (DEMAZEAU et WERNER 1992).
- CASTELFRANCHI, C. ET WERNER, E., éditeurs (1994). *Artificial Social Agents. Proceedings of the 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW'92)*, volume 830 de *Lecture Notes in Artificial Intelligence*. Springer Verlag, Berlin-Heidelberg.
- CHAUVIN, R. (1982). *Les Sociétés animales*. Seghers, Paris.
- CHENEY, D. L. ET SEYFARTH, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behavior*, 30:362–367.

- CLOCKSIN, W. F. ET MELLISH, C. S. (1979). *Programming in Prolog*. Springer Verlag, New York.
- COLLINOT, A. ET DROGOUL, A. (1998). La méthode de conception multi-agent Cassiopée : application à la robotique collective. *Revue d'Intelligence Artificielle*, 12(1):125–147.
- COLLINOT, A., DROGOUL, A., ET BENHAMOU, P. (1996a). Agent-Oriented Design of a Soccer Robot Team. In (TOKORO 1996).
- COLLINOT, A., PLOIX, L., ET DROGOUL, A. (1996b). Application de la méthode CASSIOPÉE à l'organisation d'une équipe de robots. In (MÜLLER et QUINQUETON 1996), pp. 136–152.
- COQUILLARD, P. ET HILL, D. (1997). *Modélisation et simulation d'écosystèmes*. Masson, Paris.
- CYRULNIK, B. (1983). *Mémoire de singe et paroles d'homme*. Pluriel. Hachette, Paris.
- CYRULNIK, B., éditeur (1998). *Si les lions pouvaient parler. Essais sur la condition animale*. Quarto, Gallimard, Paris.
- DARWIN, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, Londres.
- DASSER, V. (1987). *A Social Concept of Monkeys*. Thèse inaugurale, Université de Zürich.
- DASSER, V. (1988). A social concept in Java monkeys. *Animal Behavior*, 36:225–230.
- DAUTENHAHN, K. (1997a). I could be you – the phenomenological dimension of social understanding. *Cybernetics and Systems*, 25(8):417–453.
- DAUTENHAHN, K. (1997b). The role of interactive conceptions of intelligence and life in cognitive technology. In (MARSH et al. 1997).
- DAUTENHAHN, K. (1998). Grounding agent sociality: The social world is its own best model. In (TRAPPL 1998).
- DAUTENHAHN, K. (1999a). Embodiment and interaction in socially intelligent life-like agents. In (NEHANIV 1999), pp. 105–147.
- DAUTENHAHN, K. (1999b). Robots as social actors: Aurora and the case of autism. In *Proceedings of the International Cognitive Technology Conference (ICTC'99)*.

## BIBLIOGRAPHIE

- DAUTENHAHN, K. (2000a). Design issues on interactive environments for children with autism. In *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT'2000)*.
- DAUTENHAHN, K. (2000b). Socially Intelligent Agents and the Primate Social Brain: Towards a Science of Social Minds. In (DAUTENHAHN 2000c), pp. 35–51.
- DAUTENHAHN, K., éditeur (2000c). *Socially Intelligent Agents: the Human in the Loop (SIA'2000)*, volume FS-00-04 de *AAAI Fall Symposium Series*. AAAI Press, Menlo Park, California.
- DAUTENHAHN, K. ET WERRY, I. (2000). Issues of Robot-Human Interaction Dynamics in the Rehabilitation of Children with Autism. In (MEYER et al. 2000), pp. 519–528.
- DAWKINS, R. (1980). Good strategy or evolutionarily stable strategy? In (BARLOW et SILVERBERG 1980).
- DAWKINS, R. (1982). *The Extended Phenotype. The Long Reach of the Gene*. Oxford University Press.
- DAWKINS, R. (1986). *The Blind Watchmaker. Why the Evidence of Evolution Reveals a Universe Without Design*. Penguin Books Ltd., London.
- DAWKINS, R. (1989). *The Selfish Gene*. Oxford University Press, seconde édition.
- DE JONG, K. A. (1975). *An analysis of the behavior of a class of genetic adaptive systems*. Thèse de PhD, Dept. of Computer and Communication Sciences, University of Michigan.
- DE WAAL, F. B. (1982). *Chimpanzee Politics: power & sex among apes*. Jonathan Cape Ltd, London.
- DEMAZEAU, Y., éditeur (1998). *Proceedings of 3rd International Conference on Multi-Agent Systems (ICMAS'98)*. IEEE Press.
- DEMAZEAU, Y. ET MÜLLER, J.-P., éditeurs (1990). *Decentralized AI*. Elsevier (North-Holland), Amsterdam.
- DEMAZEAU, Y. ET WERNER, E., éditeurs (1992). *Decentralized AI 3*. Elsevier (North-Holland), Amsterdam.
- DENEUBOURG, J.-L., GOSS, S., FRANKS, N., SENDOVA-FRANKS, A., DETRAIN, C., ET CHRETIEN, L. (1991). The Dynamics of Collective Sorting, Robot-Like Ants and Ant-Like Robots. In (MEYER et WILSON 1991), pp. 356–366.

- DENNETT, D. C. (1990). *La stratégie de l'interprète: le sens commun et l'univers quotidien*. nrf Essais. Gallimard, Paris. Traduit de l'anglais par Pascal Engel.
- DENNETT, D. C. (1996). *Darwin's Dangerous Idea. Evolution and the Meanings of Life*. Penguin Books Ltd., London.
- DESOR, D. (1999). *Le Comportement social des animaux. De l'art de vivre ensemble chez les fourmis, les rats, les loups et les autres...* « Vies sociales ». Presses Universitaires de Grenoble, Grenoble.
- DESSALLES, J.-L. (1992). Biomimetic Use of Genetic Algorithms. In MÄNNER, R. ET MANDERICK, B., éditeurs, *Parallel Problem Solving from Nature*, volume 2, pp. 127–135.
- DESSALLES, J.-L. (1996a). *L'ordinateur génétique*. Hermès, Paris.
- DESSALLES, J.-L. (1996b). Pourquoi est-on, ou n'est-on pas, pertinent? *Communication et Langages*, 107:69–80.
- DESSALLES, J.-L. (1999). Coalition factor in the evolution of non-kin altruism. *Advanced Complex Systems*, 2:143–172.
- DESSALLES, J.-L. (2000). *Aux origines du langage: une histoire naturelle de la parole*. Hermès, Paris.
- DORAN, J. (2001). Agent-Based Modelling of Ecosystems for Sustainable Resource Management. In (LUCK et al. 2001), pp. 383–403.
- DROGOUL, A. (1993). *De la Simulation Multi-Agents à la Résolution Collective de Problèmes*. Thèse de doctorat, Université Paris VI.
- DROGOUL, A. (2000). *Systèmes multi-agents situés*. Dossier d'habilitation à diriger les recherches, Université Paris VI, Paris.
- DROGOUL, A. ET COLLINOT, A. (1997). Entre réductionnisme méthodologique et stratégie intentionnelle, l'éthologie, un modèle alternatif pour l'IAD? In (QUINQUETON et al. 1997).
- DROGOUL, A. ET COLLINOT, A. (1998). Applying an Agent-Oriented Methodology to the Design of Artificial Organizations: a Case Study in Robotic Soccer. *Journal of Autonomous Agents and Multi-Agent Systems*, 1(1):113–129.
- DROGOUL, A. ET DUBREUIL, C. (1992). Eco-Problem-Solving Model: Results of the N-Puzzle. In (DEMAZEAU et WERNER 1992), pp. 283–295.
- DROGOUL, A. ET DUBREUIL, C. (1993). A Distributed Approach to N-Puzzle Solving. In *Proceedings of the Distributed Artificial Intelligence Workshop*. Seattle.

## BIBLIOGRAPHIE

- DROGOUL, A. ET DUHAUT, D. (1996). MICROB: Making Intelligent Collective ROBOTics. In *Proceedings of MIROSOT'96*.
- DROGOUL, A. ET FERBER, J. (1992). Multi-Agent Simulation as a Tool for Modeling Societies: Application to Social Differentiation in Ant Colonies. In (CASTELFRANCHI et WERNER 1994).
- DROGOUL, A. ET FERBER, J. (1993). From Tom-Thumb to the Dockers: Some Experiments with Foraging Robots. In (MEYER et al. 1993), pp. 451–459.
- DROGOUL, A., FERBER, J., ET JACOPIN, É. (1991). Pengi: Applying Eco-Problem-Solving for Behavior Modelling in an Abstract Eco-System. In *Modelling and Simulation. Proceedings of ESM'91*.
- DROGOUL, A. ET FRESNEAU, D. (1998). Métaphore du fourragement et modèle d'exploitation collective de l'espace sans communication ni interaction pour des colonies de robots autonomes mobiles. In (BARTHÈS et al. 1998).
- DROGOUL, A. ET MEYER, J.-A., éditeurs (1999). *Intelligence artificielle située – cerveau, corps et environnement*. Hermès, Paris.
- DROGOUL, A. ET PICAULT, S. (1999). MICRobES : vers des collectivités de robots socialement situées. In (GLEIZES et MARCENAC 1999).
- DROGOUL, A. ET SERVAT, D. (2001). L'intelligence diffuse. In *Actes des JFIADSMASMA'01*. Hermès (à paraître).
- DROGOUL, A. ET ZUCKER, J.-D. (1998). Methodological Issues for Designing Multi-Agent Systems with Machine Learning Techniques: Capitalizing Experiences from the Robocup Challenge. Rapport interne LIP6, Université Paris VI.
- DUCOURNAU, R., EUZENAT, J., MASINI, G., ET NAPOLI, A., éditeurs (1998). *Langages et Modèles à Objets, état des recherches et perspectives*. Collection Didactique 19. INRIA, Rocquencourt.
- DUNBAR, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, (20):469–493.
- DUNBAR, R. I. (1996). *Grooming, gossiping and the origins of language*. Faber & Faber, London.
- EDMONDS, B. (2000). The Use of Models – making MABS more informative. In *Proceedings of the Workshop on Multi-Agent Based Simulation (MABS'2000)*.

- EGGENBERGER, P. (1997). Evolving morphologies of simulated 3D organisms based on differential gene expression. In HUSBANDS, P. ET HARVEY, I., éditeurs, *Proceedings of the 4th European Conference on Artificial Life (ECAL'97)*. MIT Press, Cambridge, Massachusetts.
- EGGENBERGER, P. (1999). *Evolution of three-dimensional, artificial organisms: simulations of developmental processes*. Ph.D. Dissertation, Medical Faculty, University of Zurich.
- EIBL-EIBESFELDT, I. (1972). Grundriß der vergleichenden Verhaltensforschung. *III. Aufl. München (Piper)*.
- EIBL-EIBESFELDT, I. (1976). *L'homme programmé. L'inné facteur déterminant du comportement humain*. Flammarion, Paris. Traduit de l'allemand par Anneliese Planck.
- EPSTEIN, J. M. ET AXTELL, R. (1996). *Growing Artificial Societies. Social Science from the Bottom Up*. Brookings Institution Press, Washington, D.C.
- ERCEAU, J. ET FERBER, J. (1991). L'Intelligence Artificielle Distribuée. *La Recherche*, 233.
- FERBER, J. (1995). *Les systèmes multi-agents : Vers une intelligence collective*. InterEditions, Paris.
- FIKES, R. ET SANDEWALL, E., éditeurs (1991). *Proceedings of Knowledge Representation and Reasoning (KRR'91)*. Morgan Kaufmann Publishers, Inc.
- FILLIAT, D. (1999). Utilisation de mécanismes de perception active pour la navigation d'un robot mobile. In (DROGOUL et MEYER 1999), pp. 159–174.
- FILLIAT, D. ET MEYER, J.-A. (2000). Active Perception and Map Learning for Robot Navigation. In (MEYER et al. 2000), pp. 246–255.
- FLOREANO, D. ET URZELAI, J. (2000). Evolutionary Robots with on-line self-organization and behavioral fitness. *Neural Networks*, 13:431–443.
- FOISEL, R., DROGOUL, A., CAYROL, O., ATTIA, M., ET CHAUVAT, N. (1999). Des écosystèmes artificiels d'aide à la conception : l'exemple du projet CAROSSE. In (GLEIZES et MARCENAC 1999), pp. 313–326.
- FOUTS, R., HIRSCH, A., ET FOUTS, D. (1982). Cultural Transmission of a Human Language in a Chimpanzee Mother/Infant Relationship. In FITZGERALD, H. E., MULLINS, J. A., ET PAGE, P., éditeurs, *Psychobiological*

## BIBLIOGRAPHIE

- Perspectives: Child Nurturances Series*, volume III, pp. 159–193. Plenum Press, New York.
- GALLUP, G. G. (1970). Chimpanzees: self-recognition. *Science*, 167:86–87.
- GLEIZES, M.-P. ET MARCENAC, P., éditeurs (1999). *Ingénierie des systèmes multi-agents. Actes des JFIADSMA '99*. Hermès, Paris.
- GOLDBERG, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.
- GOODALL, J. (1979). Intercommunity interactions in the chimpanzee population of the Gombe National Park. In HAMBURG ET MACCOWN, éditeurs, *The Great Apes*, pp. 13–53. Benjamin-Cummings Publisher.
- GOSS, S. ET DENEUBOURG, J.-L. (1992). Harvesting by a Group of Robots. In (VARELA et BOURGINE 1992), pp. 195–204.
- GOURICHON, S. (1999). Catégorisation distribuée et hiérarchisée de son environnement par un robot mobile. In (DROGOUL et MEYER 1999), pp. 87–108.
- GRASSÉ, P.-P. (1959). La reconstruction du nid et les coordinations inter-individuelles chez *Bellicositermes natalensis* et *Cubitermes sp.*, la théorie de la stigmergie: essais d'interprétation du comportement des termites constructeurs. *Ins. Soc.*, 6:41–84.
- GRUAU, F. (1994). *Neural Network Synthesis Using Cellular Encoding and the Genetic Algorithm*. Thèse de doctorat, ENS Lyon – Université Lyon I.
- GUESSOUM, Z. (1996). *Un environnement opérationnel de conception et de réalisation de systèmes multi-agents*. Thèse de doctorat, Université Paris VI.
- GUTKNECHT, O. ET FERBER, J. (1999). Vers une méthodologie organisationnelle des systèmes multi-agents. In (GLEIZES et MARCENAC 1999), pp. 81–92.
- HARA, F. ET PFEIFER, R. (2000). On the relation among morphology, material and control in morpho-functional machines. In (MEYER et al. 2000), pp. 33–40.
- HEITKÖTTER, J. ET BEASLEY, D. (1998). The hitch-hiker's guide to evolutionary computation (faq for comp.ai.genetic).
- HEMELRIJK, C. K. (1998a). Risk sensitive and ambiguity reducing dominance interactions in a virtual laboratory. In (PFEIFER et al. 1998), pp. 255–262.
- HEMELRIJK, C. K. (1998b). Spatial Centrality of Dominants without Positional Preference. In (ADAMI et al. 1998), pp. 307–315.

- HEMELRIJK, C. K. (2000). Towards the integration of social dominance and spatial structure. *Animal Behaviour*, 59:1035–1048.
- HEWITT, C. ET INMAN, J. (1991). Distributed Artificial Intelligence betwixt and between: From “Intelligent Agents” to Open Systems science. *IEEE Transactions on Systems, Man and Cybernetics*, 21(6):1409–1419.
- HIGUCHI, T., NIWA, T., TANAKA, T., IBA, H., DE GARIS, H., ET FURUYA, T. (1993). Evolving hardware with genetic learning: A first step to build a darwinian machine. In (MEYER et al. 1993).
- HOFSTADTER, D. (1985). *Gödel, Escher, Bach: les Brins d’une Guirlande Éternelle*. InterEditions, Paris. Version française de Jacqueline Henry et Robert French.
- HOGEWEG, P. ET HESPER, B. (1983). The Ontogeny of the Interaction Structure in Bumble Bee Colonies: a MIRROR Model. *Behavioural Ecology and Sociobiology*, 12:271–283.
- HOGEWEG, P. ET HESPER, B. (1985). SocioInformatic Processes; MIRROR Modeling Methodology. *Journal of Theoretical Biology*, 113:311–330.
- HOLLAND, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press.
- HOMÈRE (IX<sup>e</sup> s. A.C.). *L’Iliade*. Actes Sud–Babel. Traduit du grec par Frédéric Mugler, 1989.
- HUGUES, L. (2000a). Collective Grounded Representations for Robots. In (PARKER 2000).
- HUGUES, L. (2000b). Grounded Representations for a Robot Team. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS’2000)*.
- HUGUES, L. ET DROGOUL, A. (2001a). Grounding of Robot Behaviors. In *Proceedings of AAAI Fall Symposium on Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems*.
- HUGUES, L. ET DROGOUL, A. (2001b). Robot Behavior Learning by Vision-Based Demonstrations. In *Proceedings of the 4th European Workshop on Advanced Mobile Robotics (EUROBOT’01)*.
- HUGUES, L. ET DROGOUL, A. (2001c). Shaping of Robot Behaviors by Demonstrations. In *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*.

## BIBLIOGRAPHIE

- HUTZLER, G. (1995). *Simulation de l'apparition de hiérarchies dans les sociétés polygynes de fourmis avec MANTA*. Rapport de stage de DEA, Université Paris VI.
- JARVIS, J. U. M. (1981). Eusociality in a mammal: cooperative breeding in naked-mole rat colonies. *Science*, 4494:571–573.
- JEAN, M. R. (1997). Émergence et SMA. In (QUINQUETON et al. 1997), pp. 323–342.
- JENNINGS, N. R. (1993). Specification and implementation of a belief desire joint-intention architecture for collaborative problem solving. *Journal of Intelligent and Cooperative Information Systems*, 2(3):298–318.
- JOAS, H. (1999). *La créativité de l'agir*. Les Éditions du Cerf. Traduit de l'allemand par Pierre Rusch.
- KAPLAN, F. (2000a). Free creatures: The role of uselessness in the design of artificial pets. In CHRISTALLER, T., INDIVERI, G., ET POIGNE, A., éditeurs, *Proceedings of the 1st Edutainment Robotics Workshop*. GMB-AiS.
- KAPLAN, F. (2000b). *L'émergence d'un lexique dans une population d'agents autonomes*. Thèse de doctorat, Université Paris VI.
- KAPLAN, F. (2000c). Semiotic schemata: Selection units for linguistic cultural evolution. In (BEDEAU et al. 2000).
- KAPLAN, F. (2001). *La naissance d'une langue chez les robots*. Hermès, Paris.
- KAPLAN, F. ET OUDEYER, P.-Y. (2000). From Talking AIBO to Talking Head: Morphing sensory spaces between two robotic bodies. In *Proceedings of Sony Research Forum*.
- KIMURA, M. (1990). *Théorie neutraliste de l'Évolution*. Nouvelle Bibliothèque Scientifique. Flammarion, Paris. Traduit de l'anglais par Claudine Montgelard.
- KING, M. C. ET WILSON, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science*, 188:107–116.
- KITANO, H. ET ASADA, M. (1997). RoboCup: a Challenge Problem for AI. *AI Magazine*, 18(1).
- KODJABACHIAN, J. ET MEYER, J.-A. (1998). Evolution and Development of Neural Controllers for Locomotion, Gradient-Following, and Obstacle-Avoidance in Artificial Insects. *IEEE Transactions on Neural Networks*, 9:796–812.

- KORF, R. E. (1991). A Simple Solution to Pursuit Games. In *Proceedings of the Distributed Artificial Intelligence Workshop*, pp. 183–194.
- KOZA, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, Massachusetts.
- KUMMER, H. (1967). *Tripartite relations in hamadryas baboons*. In (ALTMANN 1967).
- KUPIEC, J.-J. ET SONIGO, P. (2000). *Ni Dieu ni gène. Pour une autre théorie de l'hérédité*. Science ouverte. Éditions du Seuil, Paris.
- LALANDE, S., DROGOUL, A., ET THIERRY, B. (1995). MACACA: a Multi-Agent Computer simulation of Animal Communities based on Alliances. In *Proceedings of the Simulating Societies Symposium (SIMSOC'95)*.
- LANDAU, S., DONCIEUX, S., DROGOUL, A., ET MEYER, J.-A. (2001a). SFERES, un framework pour la conception de systèmes multi-agents adaptatifs. Rapport technique, LIP6, Paris.
- LANDAU, S. ET PICAULT, S. (2001a). Modeling Adaptive Multi-Agent Systems Inspired by Developmental Biology. In (MAŘÍK et al. 2001b), pp. 238–246.
- LANDAU, S. ET PICAULT, S. (2001b). Stack-Based Gene Expression. *Journal of Genetic Programming and Evolvable Machines*. Soumission.
- LANDAU, S., PICAULT, S., ET DROGOUL, A. (2001b). ATNoSFERES: a Model for Evolutive Agent Behaviors. In *Proceedings of the AISB'01 Symposium on Adaptive Agents and Multi-Agent Systems*.
- LANGTON, C., éditeur (1988). *Artificial Life*, volume 1. Addison-Wesley, London.
- LANGTON, C., TAYLOR, C., FARMER, J. D., ET RASMUSSEN, S., éditeurs (1991). *Artificial Life*, volume 2. Addison-Wesley, London.
- LECOINTRE, G. ET LE GUYADER, H. (2001). *Classification phylogénétique du vivant*. Belin, Paris.
- LESAGE, F. (2000). *Interprétation adaptative du discours dans une situation multiparticipants : modélisation par agents*. Thèse de doctorat, Université du Havre.
- LESTEL, D. (1998a). Des animaux-machines aux machines animales. In (CYRULNIK 1998).
- LESTEL, D. (1998b). Les singes parlent-ils vraiment? In (CYRULNIK 1998), pp. 990–1009.

## BIBLIOGRAPHIE

- LESTIENNE, R. (1993). *Le hasard créateur*. Sciences et société. Éditions La Découverte, Paris.
- LORENZ, K. (1984). *Les Fondements de l'éthologie*. Flammarion, Paris. Traduit de l'allemand par Jeanne Étoré.
- LUCK, M., MAŘÍK, V., ŠTĚPÁNKOVÁ, O., ET TRAPPL, R., éditeurs (2001). *Multi-Agent Systems and Applications. 9th ECCAI Advanced Course (ACAI'01) and AgentLink's 3rd European Agent Systems Summer School (EASSS'01)*, volume 2086 de *Lecture Notes in Artificial Intelligence*. Springer Verlag, Berlin-Heidelberg.
- MACLENNAN, B. (1991). Synthetic Ethology: An Approach to the Study of Communication. In (LANGTON et al. 1991), pp. 631–657.
- MAŘÍK, V., ŠTĚPÁNKOVÁ, O., KRAUTWURMOVÁ, H., ET BRIOT, J.-P., éditeurs (2001a). *Multi-Agent Systems and Applications. Proceedings of ACAI'2001 & EASSS'2001 Student Sessions*. Czech Technical University, Prague.
- MAŘÍK, V., ŠTĚPÁNKOVÁ, O., KRAUTWURMOVÁ, H., ET BRIOT, J.-P., éditeurs (2001b). *Proceedings of the Workshop on Adaptability and Embodiment Using Multi-Agent Systems (AEMAS'2001)*. Czech Technical University, Prague.
- MARSH, J. P., NEHANIV, C. L., ET GORAYSKA, B., éditeurs (1997). *Proceedings of the 2nd International Cognitive Technology Conference (ICTC'97)*. IEEE Press.
- MARSH, S. (1992). Trust in Distributed Artificial Intelligence. In (CASTELFRANCHI et WERNER 1994).
- MATARIC, M. (1993). Designing Emergent Behaviors: From Local Interactions to Collective Intelligence. In (MEYER et al. 1993), pp. 432–441.
- MATARIC, M. (1995). Learning in Multi-Robot Systems. In *Workshop on Adaptation and Learning in Multi-Agent Systems, 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*. AAAI Press.
- MATARIC, M. (1997). Learning Social Behavior. *Robotics and Automation Systems*, 20:191–204.
- McFARLAND, D., éditeur (1990). *Dictionnaire du comportement animal*. Bouquins. Robert Laffont, Paris. Traduit de l'anglais par Guy Schoeller.
- MEURISSE, T. ET VANBERGUE, D. (2001). Et maintenant à qui le tour? Aperçu de Problématiques de Conception de Simulations Multi-Agents. In

*Actes du Colloque ALCAA (Agents Logiciels, Coopération, Apprentissage et Activité Humaine)*. IUT de Bayonne, Université de Pau.

- MEYER, J.-A., BERTHOZ, A., FLOREANO, D., ROITBLAT, H. L., ET WILSON, S. W., éditeurs (2000). *From Animals to Animats 6. Proceedings of the 6th International Conference on Simulation of Adaptive Behavior (SAB'2000)*. MIT Press, Cambridge, Massachusetts.
- MEYER, J.-A. ET GUILLOT, A. (1990). From animals to animats: Everything you wanted to know about the simulation of adaptive behavior. Technical report 90-1, Groupe de Bioinformatique de l'école Normale Supérieure, Paris.
- MEYER, J.-A., ROITBLAT, H. L., ET WILSON, S. W., éditeurs (1993). *From Animals to Animats 2. Proceedings of the 2nd International Conference on Simulation of Adaptive Behavior*. MIT Press, Cambridge, Massachusetts.
- MEYER, J.-A. ET WILSON, S. W., éditeurs (1991). *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*. MIT Press.
- MICHAUD, F. (2000). Social Intelligence and Robotics. In (DAUTENHAHN 2000c), pp. 127-130.
- MICHAUD, F. ET CARON, S. (2000). Roball – an autonomous toy-rolling robot. In *Proceedings of the Workshop on Interactive Robotics and Entertainment*.
- MICHAUD, F., CLAVET, A., LACHIVER, G., ET LUCAS, M. (2000). Designing toy robots to help autistic children – an open design project for electrical and computer engineering in education. In *Proceedings of the American Society for Engineering Education*.
- MITCHELL, M. ET TAYLOR, C. E. (1999). Evolutionary computation: An overview. *Annual Review of Ecology and Systematics*, 20:593-616.
- MONTAIGNE (1580). *Essais*. Simon Millanges, Bordeaux. Édition moderne établie et annotée par Pierre Michel. Le Livre de Poche.
- MORGAN, L. C. (1894). *An Introduction to Comparative Psychology*. Scott, London.
- MORRIS, D. (1977). *La clé des gestes*. Grasset, Paris.
- MOSS, S. ET DAUTENHAHN, K. (1998). Hierarchical organisation of robots: a social simulation study. In (ZOBEL et MOELLER 1998).
- MOUTON, D. (2001). *Le Cataphile*. CyLibris, Paris.

## BIBLIOGRAPHIE

- MÜLLER, J.-P. (1998). Vers une méthodologie de conception des systèmes de résolution de problèmes par émergence. In (BARTHÈS et al. 1998), pp. 355–371.
- MÜLLER, J.-P. ET QUINQUETON, J., éditeurs (1996). *Intelligence Artificielle Distribuée et Systèmes Multi-Agents. Actes des JFIADSMA '96*. Hermès, Paris.
- MUÑOZ-MELÉNDEZ, A. (2001). Cooperation in Collective Robotics. In (MAŘÍK et al. 2001a), pp. 64–71.
- MUÑOZ-MELÉNDEZ, A. ET DROGOUL, A. (2000). Towards the Design of Self-Vigilant Robots using a Principle of Situated Cooperation. In *Proceedings of the RoboCup Rescue Workshop of ICMAS'2000*.
- NEHANIV, C. L., éditeur (1999). *Computation for Metaphors, Analogy and Agent*, volume 1562 de *Lectures Notes in Artificial Intelligence*. Springer Verlag.
- NISHIDA, T., UEHARA, S., ET NYUNDO, R. (1979). Predatory behaviour among wild chimpanzees of the Mahale Mountains. *Primates*, 20:1–20.
- NOLFI, S. ET FLOREANO, D. (2000). *Evolutionary Robotics. The Biology, Intelligence, and Technology of Self-organizing Machines*. MIT Press, Cambridge, Massachusetts.
- NOLFI, S., MIGLINO, O., ET PARISI, D. (1994). Phenotypic Plasticity in Evolving Neural Networks. In NICLOUD, J.-D. ET GAUSSIER, P., éditeurs, *Proceedings of the Conference From Perception to Action*. IEEE Press.
- NOURBAKHSI, I. (1998). The failures of a self-reliant tour robot with no planner or, what might a planner buy you? Internal report. Robotics Institute, Carnegie Mellon.
- PARDI, L. (1948). Dominance order in Polistes wasps. *Physiological Zoology*, 21:1–13.
- PARKER, L. E., éditeur (2000). *Proceedings of the International Conference on Distributed Autonomous Robotic Systems (DARS'2000)*.
- PARUNAK, H. V. D. (1997). “Go to the Ant”: Engineering Principles from Natural Multi-Agent Systems. *Annals of Operations Research*.
- PEREC, G. (1969). *La Disparition*. L’Imaginaire. Gallimard, Paris.
- PERROT, J.-F. (1998). Objets, classes, héritage : définitions. In (DUCOURNAU et al. 1998).
- PFEIFER, R. (2000). On the role of morphology and materials in adaptive behavior. In (MEYER et al. 2000), pp. 23–32.

- PFEIFER, R., BLUMBERG, B., MEYER, J.-A., ET WILSON, S. W., éditeurs (1998). *From Animals to Animats 5. Proceedings of the 5th International Conference on Simulation of Adaptive Behavior (SAB'98)*. Cambridge, Massachusetts.
- PICAULT, S. (1998). A Multi-Agent Simulation of Primate Social Concepts. In (PRADE 1998).
- PICAULT, S. ET COLLINOT, A. (1998a). Designing Social Cognition Models for Multi-Agent Systems through Simulating Primate Societies. In (DEMAZEAU 1998).
- PICAULT, S. ET COLLINOT, A. (1998b). La socialité: étude, enjeux et applications dans les Systèmes multi-agents. In (BARTHÈS et al. 1998), pp. 327–340.
- PICAULT, S. ET DROGOUL, A. (2000a). The MICRobES Project, an Experimental Approach towards “Open Collective Robotics”. In (PARKER 2000).
- PICAULT, S. ET DROGOUL, A. (2000b). Robots as a Social Species: the MICRobES Project. In (DAUTENHAHN 2000c), pp. 139–141.
- PICAULT, S. ET LANDAU, S. (2001a). ATNoSFERES: un Modèle d'évolution Darwinienne pour le Comportement Individuel ou Collectif d'Agents. Rapport technique, LIP6, Paris. En préparation.
- PICAULT, S. ET LANDAU, S. (2001b). Ethogenetics: an Evolutionary Approach to Agents Organization. In *Actes du Colloque ALCAA (Agents Logiciels, Coopération, Apprentissage et Activité Humaine)*. IUT de Bayonne, Université de Pau.
- PICAULT, S. ET LANDAU, S. (2001c). Ethogenetics and the Evolutionary Design of Agent Behaviors. In (CALLAOS et al. 2001), pp. 528–533.
- PICAULT, S., SERVAT, D., ET KAPLAN, F. (1997). EDEN: un système évolutif endosémantique. Rapport technique, École Nationale Supérieure des Télécommunications, Paris.
- PICQ, P. ET COPPENS, Y., éditeurs (2001). *Aux origines de l'humanité: le propre de l'homme*. Librairie Arthème Fayard, Paris.
- POINCARÉ, H. (1893). *Les méthodes nouvelles de la mécanique céleste*. Gauthier Villars, Paris.
- POPPER, K. (1963). *Conjectures et réfutations*. Payot.
- PRADE, H., éditeur (1998). *Proceedings of 13th European Conference on Artificial Intelligence (ECAI'98)*. John Wiley & Sons, Ltd., Brighton.

## BIBLIOGRAPHIE

- PRIGOGINE, I. (1996). *La fin des certitudes*. Collection « Opus ». Odile Jacob, Paris.
- PYNADATH, D. V., TAMBE, M., ARENS, Y., CHALUPSKY, H., GIL, Y., KNOBLOCK, C., LEE, H., LERMAN, K., OH, J., RAMACHANDRAN, S., ROSENBLOOM, P. S., ET RUSS, T. (2000). Electric Elves: Immersing an agent organization in a human organization. In (DAUTENHAHN 2000c), pp. 150–154.
- QUINQUETON, J. ET HAMADI, Y. (1999). Communication et émergence : une épidémie chez les termites. In (GLEIZES et MARCENAC 1999), pp. 226–235.
- QUINQUETON, J., THOMAS, M.-C., ET TROUSSE, B., éditeurs (1997). *Intelligence Artificielle et Systèmes Multi-Agents. Actes des JFIADSMA '97*. Hermès, Paris.
- RAO, A. S. ET GEORGEFF, M. P. (1991). Modeling rational agents within a BDI-architecture. In (FIKES et SANDEWALL 1991), pp. 473–484.
- RAY, T. S. (1991). An evolutionary approach to synthetic biology, zen and the art of creating life. In (LANGTON et al. 1991).
- RECHENBERG, I. (1973). *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Frommann–Holzboog, Stuttgart.
- RESNICK, M. (1994). *Turtles, Termites, and Traffic Jams. Exploration in Massively Parallel Microworlds*. Complex Adaptative Systems. MIT Press, Cambridge, Massachusetts.
- REYNOLDS, C. (1987). Flocks, Herds, and Schools: a Distributed Behavioral Model. In *Proceedings of the SIGGRAPH'87 Conference*.
- ROBERT, P. (1993). *Le Petit Robert. Dictionnaire alphabétique et analogique de la langue française*. Dictionnaires Le Robert, Paris. Nouvelle édition sous la direction de Josette Rey-Debove et Alain Rey.
- RUSSELL, J. (1997). *Autism as an Executive Disorder*. Oxford University Press.
- SANSONNET, J.-P. (1999). Magenta : représentation intensionnelle des interactions pour le raisonnement causal dans les SMA. In (GLEIZES et MARCENAC 1999), pp. 151–164.
- SAVAGE-RUMBAUGH, E. S., McDONALD, K., SEVCIK, R., HOPKINS, W., ET RUBERT, E. (1986). Spontaneous Symbol Acquisition and Communicative

- Use by Pygmy Chimpanzees (*pan paniscus*). *Journal of Experimental Psychology: General*, 115(335):311–335.
- SAVAGE-RUMBAUGH, E. S., MURPHY, J., SEVCIK, R., BRAKKE, K., WILLIAM, S., ET RUMBAUGH, D. (1993). Language comprehension in Ape and Child. *Monographs of the Society for Research in Child Development*, 223(58):1–221.
- SCERRI, P., TAMBE, M., LEE, H., ET PYNADATH, D. (2000). Don't cancel my Barcelona trip. Adjusting autonomy of agent proxies in human organizations. In (DAUTENHAHN 2000c), pp. 169–173.
- SCHJELDERUP-EBBE, T. (1922). Beiträge zur Sozialpsychologie des Haushuhns. *Zeit. Psych.*, 88:225–252.
- SCHJELDERUP-EBBE, T. (1935). Social behavior in birds. In MURCHISON, C. A., éditeur, *Handbook of Social Psychology*. Clark University press, Worcester Mass, USA.
- SCHOONDERWOERD, R., HOLLAND, O., BRUTEN, J., ET ROTHKRANZ, L. (1997). Ant-based load balancing in telecommunications networks. *Adaptive Behavior*, 5:169–207.
- SCHWEFEL, H.-P. (1975). *Evolutionstrategie und numerische Optimierung*. Dr.-Ing. Thesis, Technical University of Berlin, Department of Process Engineering.
- SERVAT, D. (2000). *Modélisation des dynamiques de flux par agents. Application aux processus de ruissellement, infiltration et érosion*. Thèse de doctorat, Université Paris VI.
- SIMS, K. (1995). Evolving 3D morphology and behavior by competition. *Artificial Life*, 1(1):353–372.
- SOLER, L. (2000). *Introduction à l'épistémologie*. Ellipses.
- STEELS, L. (1989). Cooperation between Distributed Agents through Self-Organisation. In (DEMAZEAU et MÜLLER 1990).
- STONE, P. ET VELOSO, M. (1997). A layered approach to learning client behaviors in the RoboCup soccer server. *Applied Artificial Intelligence (AAI) Journal*.
- STÜMPKE, H. (1962). *Bau und Leben der Rhinogradentia*. Gustav Fischer Verlag, Stuttgart. Traduction française par Robert Weill sous le titre « Anatomie et biologie des Rhinogrades », préface de P.-P. Grassé, avec une post-face de Gérolf Steiner, aux éditions Masson.

## BIBLIOGRAPHIE

- TE BOEKHORST, I. J. A. ET HOGEWEG, P. (1994). Effect of tree size on travelband formation in Orang-utans: Data analysis suggested by a model study. In (BROOKS et MAES 1994).
- THERAULAZ, G. (1991). *Morphogenèse et Auto-Organisation des Comportements dans les Colonies de Guêpes Polistes Dominilus (Christ)*. Thèse de troisième cycle, Université d'Aix-Marseille I.
- THIERRY, B. (1985). Patterns of agonistic interactions in three species of macaque (*Macaca mulatta*, *M. fascicularis*, *M. tonkeana*). *Aggressive Behavior*, 11:223–233.
- THIERRY, B. (1990). Feedback Loop between Kinship and Dominance: the Macaque Model. *Journal of Theoretical Biology*, 145:511–521.
- THRUN, S. (1998). When robots meet people: Research directions in mobile robots. *IEEE Intelligent Systems*.
- THRUN, S., BENNEWITZ, M., BURGARD, W., CREMERS, A. B., DELLAERT, F., FOX, D., HÄHNEL, D., ROSENBERG, C., ROY, N., SCHULTE, J., ET SCHULZ, D. (1999). MINERVA: A second-generation museum tour-guide robot. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'99)*.
- TOKORO, M., éditeur (1996). *Proceedings of the Second International Conference on Multi-Agent Systems (ICMAS'96)*. AAAI Press, Menlo Park, California.
- TOMASELLO, M. (2000). Social cognition and the origins of linguistic communication. In DESSALLES, J.-L. ET GHADAKPOUR, L., éditeurs, *The Evolution of Language. 3rd International Conference*. École Nationale Supérieure des Télécommunications, Paris.
- TRAPPL, R., éditeur (1998). *Proceedings of the 14th European Meeting on Cybernetics and Systems*.
- TURING, A. (1950). Computing machinery and intelligence. *Mind*, LXI(236):433–460.
- TURING, A. ET GIRARD, J.-Y. (1995). *La machine de Turing*. Éditions du Seuil, Paris.
- VARELA, F. J. ET BOURGINE, P., éditeurs (1992). *Towards a Practice of Autonomous Systems. Proceedings of the First European Conference on Artificial Life (ECAL'92)*. MIT Press, Cambridge, Massachusetts.
- VAUCLAIR, J. (1992). *L'intelligence de l'animal*. Éditions du Seuil, Paris.

- VISETTI, Y. M. (1996). Émergence et hétérogénéité : le problème des passages. In *Journées thématiques de l'ARC (émergence et explication)*.
- VOGEL, G. ET ANGERMANN, H., éditeurs (1994). *Atlas de la biologie*. La Pochothèque. Le Livre de Poche.
- WHITLEY, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, 4:65–85.
- WILKES, D. M., PACK, R. T., ALFORD, A., ET KAWAMURA, K. (1997). Hudl, a design philosophy for socially intelligent service robots. In *Socially Intelligent Agents (SIA '97)*, numéro FS-97-02 in AAAI Fall Symposium Series, pp. 140–145. AAAI Press.
- WILSON, E. O. (1975). *Sociobiology: the new synthesis*. Belknap Press of Harvard University Press, Cambridge.
- WINOGRAD, T. (1983). *Language as a Cognitive Process*. Addison-Wesley.
- WOLFE, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychology Monographs*, 12:1–72.
- WOODRUFF, G. ET PREMACK, D. (1979). Intentional communication in the chimpanzee: the development of deception. *Cognition*, (6):333–362.
- WOODS, W. A. (1970). Transition networks grammars for natural language analysis. *Communications of the Association for the Computational Machinery*, 13(10):591–606.
- WORDEN, R. P. (1996). Primate Social Intelligence. *Cognitive Science*, 20:579–616.
- ZAHAVI, A. ET ZAHAVI, A. (1997). *The handicap principle*. New York.
- ZOBEL, R. ET MOELLER, D., éditeurs (1998). *Proceedings of the 12th European Simulation Multiconference*. Society for Computer Simulation International.
- ZWANG, G. (2000). *Les comportements humains. Éthologie humaine*. Masson, Paris.

## BIBLIOGRAPHIE

---

# Index

---

## A

acide nucléique, voir ADN  
ADN, 33, 155, 166, 168, 173, 175  
affiliation, 34–36, 44, 62  
AIBO, 9, 123, 128  
algorithme évolutionniste, **159–160**,  
169  
Algorithme Génétique, **169–170**, 173,  
174  
animat, **16**, 159  
*ape*, 33  
apprentissage, 37, 43, 45, 54, 55, 63,  
68, 121, 123, 132, 138–140,  
154, 156, 167  
arbre, 54, 170, 174  
ATN, **176–178**, 178–183, 189  
ATNoSFERES, **174–183**, 204  
AURORA, 122  
autisme, 122

## B

*Bombus terrestris*, 56

## C

C, 143  
C++, 143, 174  
canon de MORGAN, **18**, 20  
cartographie, 134–135  
CASSIOPÉE, 68–71, 91  
catarhiniens, 33

cause

efficiente, **19**, 20, 21, 23, 89, 151,  
162, 164  
finale, voir finalisme  
proximale, voir c. efficiente  
« cercle vertueux », 10, **24–26**, 90,  
91, 147, 153  
*Cercopithecus aethiops*, 36  
chasse, 17, 39  
chromosome, 168, 176, **178**, 179, 186  
code génétique, **179**, 186  
codon, **179**  
COG, 152  
colophon, 244  
comportement  
affiliatif, 34, 61, 70  
agonistique, 61, 70  
collectif, **3**, 164  
consommatoire, 61  
d'un agent, **3**  
contre-tromperie, 41  
corporéité, 147, **150–154**, 161, 164,  
204  
corps  
d'un robot, 121, 147, 150  
définitions, **150**  
CREATURES, 123  
culture des primates, 46

## INDEX

### D

développement, voir ontogénèse  
dilemme du prisonnier, 37  
dominance, 33, 63, 70  
    chez les primates, 34  
    réelle/formelle, 72, 86

### E

écosystème, 124, 141, 204  
**EDEN, 173**  
*embodiment*, voir corporéité  
embryologie, 155, 164  
EMF, 59, 144, 165, 166  
E-EMF, 59, 73  
émotions, 120, 121  
empathie, 121, 122, 153, 155  
endo-/exo-socialité, **127**, 140, 150  
environnement, 4, 16, 20, 31, 60, 96,  
    152, 159, 164  
    clos, 5, 93, 116, 124  
    d'un agent, **3**, 12, 81, 115  
    ouvert, 8, 11, 65, 118, 124, 131  
    réel, 115–117, 121, 131, 151, 158  
    simulé, 14, 93, 118, 185  
    social, 9, 55, 88, 122, 131, 138,  
    158  
« Espace de Fitness », 205  
Éthogénétique, **165–183**, 204  
éthologie, **17–24**, 126, 141, 151, 158,  
    162  
éthologie synthétique, **23**, 57, 124  
eusocialité, 31  
Évolution, voir sélection  
expérience  
    de pensée, **13**, 24  
    données sensibles, **13**  
    scientifique, **13**

expressions faciales

    chez l'homme, 120  
    chez les primates, 34  
    chez les robots, 120, 128, 153

### F

FEELIX, 120  
finalisme, **19**, 114, 155, 162, 164  
fonction, 19, 21, 31, 68, 69, 89–91,  
    123, 147, 151, 162, 168, 175  
    adaptative, 19, 89, 158, 160, 162,  
    166  
    de fitness, 159, 187, 189, 190,  
    193, 204–206  
FURBY, 123

### G

génotype, **159**, 166, 168, 178, 180  
*Golem*, **114**  
groupes  
    typologie, **30**

### H

*Heterocephalus glaber*, 31  
hiérarchie, 34, 59, 63, 64, 87, 90, 91,  
    97, 100–102, 140  
    linéarité, 34, 63, **72–73**, 100  
hominidés, 33  
hominiens, 33  
*Homo sapiens*, 32, 33, 141  
hypothèse  
    de l'intelligence machiavélique,  
    40, **42–44**, 44  
    de séparation des rôles, **69**, 78,  
    89, 91, 153, 161

**I**

IAD

définition, **4**formes, **4–5**

imitation, 121, 153

incrémentalité, **21**, 65, 66, 92, 116,  
125, 139, 141, 166, 168

indéterminisme, 15, 158

indice de PARDI, voir indice de do-  
minanceindice de dominance, **72**intelligence machiavélienne, voir hy-  
pothèseintentionnalité, 9, 19, 66, 68, 128,  
158

interprétation, 19, 53, 66, 158, 162

**J**« jeu de langage », **94–95**, 97–104**K**

Kanzi, 46

KISMET, 120

**L**

langage

artificiel, 94–96

chez les primates, 46

leurre social, **37**, 86

lexique conventionnel

seelangage artificiel, 94

LISP, 97

**M**

MACACA, 58–63, 66, 82

métaphores, 7, 14

écosystème, 164

cellule, 173

organisme, 155, 164

phéromones, 77

programme, 156

robots, 115

société humaine, 4

sociétés animales, 5, 29, 150

synthèse des protéines, 174

MICRobES, 9, **131–147**, 152, 160,  
165

MINERVA, 119

MiroSot, 116

MIRROR, **56**, 80, 89mise en situation, **117**, 125, 126, 147,  
152sociale, **126–127**

modèle, 118

construction, 5

d'agent, 12

définition, **11**

mathématique, 12

*monkey*, 33**N**

navigation, 133–135

nombre d'or, voir OuLiPo

**O**

ontogenèse, 19, 154, 161–165

ontophylogenèse, **164**, 174organisation, **6–8**

OuLiPo, 151

**P**

P3, 123

## INDEX

*Pan paniscus*, 33  
*Pan troglodytes*, 33  
phénotype, **159**, 160, 166, 168, 180  
phéromones, **77–78**, 96  
    sociales, 78  
phylogénèse, 19, 160–165  
PIONEER, 132, 142, 143  
platyrhiniens, 33  
polyéthisme, 31, 187  
polymorphisme, 31  
*Pongo*, 33  
posture intentionnelle, 19, 68  
primates, **32–50**  
primitive comportementale, 59  
principe  
    d'incrémentalité, voir incrémentalité  
    de parcimonie, **20**, 52, 66, 78, 88, 125  
Programmation Génétique, **170–171**, 173, 174, 182, 204  
PROLOG, 54  
proprioception, **139**, 165  
protéine, 174  
« proto-concept », **36**, 66

## Q

quasi-continuité, 168, 171, 174

## R

R-100, 123  
RCO, voir Robotique Collective Ouverte  
rang, 34, 73  
réciprocité, 37, 88  
réductionnisme méthodologique, **20**  
règles de production, 52

réseau de neurones, 165, 166, 169  
rétroaction positive, 61, 67, 85, 89  
représentation, 61  
    ancrage, 55, 63  
    sociale, 53, 56, 63–87, 91, 140  
réseau de neurones, 156  
RHINO, 119  
ROBALL, 122  
RoboCup, 116–117, 124  
robot  
    définitions, **113–114**, 118  
    footballeur, 116  
ROBOTA, 121  
robotique  
    collective, 9, 32, 115–117  
    comportementale, 115, 129  
    de compagnie, 9, 123  
    et expérimentation, 114, 124, 126  
    évolutionniste, **156**  
    sociable, 118–123  
Robotique Collective Ouverte, 9, **124–129**, 131  
rôles, **69–70**, 91–92, 146, 153

## S

SAGE, 119  
SAPHIRA, 143, 144  
schéma explicatif, **18–20**, 21, 89–90, 162  
script, 53  
SDR-3, 123  
sélection naturelle, 17, 19, 20, 139, **157–159**, 161, 164, 166  
SFERES, 174  
simulation, **13**, 118, 139, 147  
    informatique, 13  
    multi-agent, **13–17**

- passage au réel, 117–118
- singe XML, 144
- anthropoïde, 33
  - grand singe, 33
  - singe vert, 36
- situatedness*, voir mise en situation
- SMA
- définition, **3–4**
- SMALLTALK, 59, 73, 74, 97
- société
- animale, 29
  - d'insectes, 31
- sociobiologie, 29
- stigmergie, **31**, 96
- SUGARSCAPE, 93, 95
- système ouvert, 8, 124
- Systèmes Physiques Symbiotiques,  
131

**X****T**

- TAMAGOCHI, 123
- téléologie, 19, 158
- téléonomie, 20, 22, 158, 162, 164,  
165
- TAMAGOCHI, 123
- territoire, 39
- théorie de l'esprit, 47, 51, 122
- théorie des jeux, 23, 37, 42, 88
- TIERRA, **172**, 173
- transposition, 24, 68, 90, 91, 105,  
146, 151, 153, 158, 161
- tromperie sociale, 39–41
- Typogénétique, **171–172**, 173

**V**

- Vie Artificielle, 16, 30, 172, 206

Ce mémoire a été écrit  
.: sous  $\text{\LaTeX}2_{\varepsilon}$  .:  
entre Noyal-sur-Vilaine,  
Châtellerault, Prague,  
Doué-la-Fontaine et Paris.

PARIS, 2001