



université de bretagne
occidentale



THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE

sous le sceau de l'Université européenne de Bretagne

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Mention : Informatique

École Doctorale SICMA

présentée par

Kristen Manac'h

Préparée au Centre Européen
de Réalité Virtuelle

Vers la notion d'agent éactif virtuel Application à l'approche dynamique évolutionnaire

Thèse soutenue le 10 Janvier 2011

devant le jury composé de :

Salima HASSAS

Professeur, Université de Lyon 1 / *rapporteuse*

Arnaud REVEL

Professeur, Université de La Rochelle / *rapporteur*

Pierre DE LOOR

Professeur, ENIB / *directeur*

René DOURSAT

Directeur de Recherche, Institut des Systèmes Complexes/ *examinateur*

Agnès GUILLOT

Maître de Conférences, Université Pierre et Marie Curie/ *examinateur*

Jacques TISSEAU

Professeur, ENIB / *examinateur*

UNIVERSITÉ EUROPÉENNE DE BRETAGNE

— Mémoire de thèse —

Spécialité : Informatique

Vers la notion d'agent éactif virtuel: Application à l'approche
dynamique évolutionnaire

KRISTEN MANAC'H

Soutenue le 10 Janvier 2011 devant la commission d'examen :

Rapporteurs

Salima	HASSAS	Professeur, Université Claude Bernard, Lyon 1
Arnaud	REVEL	Professeur, Université de La Rochelle

Examineurs

Doursat	RENÉ	Directeur de Recherche, Institut des Systèmes Complexes
Guillot	AGNÈS	Maître de Conférence, Université Pierre et Marie Curie
Tisseau	JACQUES	Professeur, École Nationale d'Ingénieurs de Brest

Directeur de thèse

De Loor	PIERRE	Professeur, École Nationale d'Ingénieurs de Brest
---------	--------	---

**Vers la notion d'agent énaactif virtuel:
Application à l'approche dynamique
évolutive**

Mémoire de thèse

KRISTEN MANAC'H

**Laboratoire d'Informatique des SYstèmes Complexes
EA 3883 UBO-ENIB**

KRISTEN MANAC'H
e-mail : manach@enib.fr
tel : +33 (0)2 98 05 89 62



Centre Européen de Réalité Virtuelle (CERV)

Technopôle Brest-Iroise
25, rue Claude Chappe
BP 38 F-29280 Plouzané
([http ://www.cerv.fr](http://www.cerv.fr))



Remerciements

Je remercie tout d'abord Jacques TISSEAU, qui m'a accueillie au CERV pendant mon stage de master Recherche. Il a permis à cette thèse de commencer en soutenant le sujet. J'adresse un grand merci à Pierre DE LOOR qui a proposé le sujet, dirigé cette thèse et a fait preuve de beaucoup de patience et de disponibilité. Il s'est beaucoup investi dans ce travail, je lui en suis très reconnaissante et j'espère ne pas l'avoir déçu.

Je tiens à remercier Salima HASSAS et Arnaud REVEL pour avoir rapporté cette thèse. J'adresse également mes remerciements à Agnès GUILLOT et René DOURSAT qui ont accepté de faire partie du jury de thèse.

Je remercie l'ensemble des personnes rencontrées au CERV au cours de ces années ainsi que mes amis. Merci à Nicole, Matthieu, Nico, Poulpo, Patoche, Fabien, Haï, Camille, Zak, Steevy, PDL, Pierre, Jacques, Fabrice, Cédric, Ronan, Champion, Goby, Gireg, les Cy(k)ril(le), Franck, Karen, Marc, Eric, Alexis, Alexandra, Romain, Gwen, Mathieu, Elen, Marie, Camille, Gilles, Fabien, Jim, Alain, Thierry et Julie Léna Mathilde Morgane .

Pour finir, je remercie ma famille. J'ai une pensée particulière pour Fañch.

Table des matières

Remerciements	iii
Table des matières	v
Table des figures	vii
Liste des tableaux	xi
Introduction	1
1 Autonomie constitutive : de la biologie à l'artificiel	9
1.1 Fondements biologiques de l'autonomie constitutive	9
1.2 L'autonomie constitutive simulée	15
1.3 Vers une autonomie constitutive comportementale	20
1.4 Synthèse	27
2 Choix techniques pour l'autonomie constitutive	33
2.1 Autonomie et architectures d'agent	33
2.2 Réseaux de neurones récurrents à temps continu	37
2.3 Approche Évolutionnaire	39
2.4 Synthèse	47
3 Argumentation pour la génération d'agents éactifs avec des CTRNN	49
3.1 Exemples de travaux pour l'étude de l'autonomie	49
3.2 Faire évoluer un système autonome	55
3.3 Synthèse	62
4 Proposition d'une démarche pour l'apprentissage ontogénétique	65
4.1 Définition d'un invariant sensorimoteur	65
4.2 Acquisition d'invariants sensorimoteurs	69
4.3 Synthèse	79
5 Évolution phylogénétique incrémentale de comportements sensorimoteurs	81

5.1	Comportement sensorimoteur dans un espace en 2 dimensions	81
5.2	Comportement sensorimoteur à travers un exemple minimaliste	90
5.3	Discussion	98
5.4	Synthèse	105
6	Apprentissage associatif ontogénétique guidé	107
6.1	Méthode	107
6.2	Analyses	112
6.3	Synthèse	118
7	Apprentissage ontogénétique guidé d'invariants sensorimoteurs	121
7.1	Méthode	121
7.2	Analyses	126
7.3	Synthèse	131
	Conclusions	133
	Publications	141
	Références bibliographiques	143
A	Comportement et invariants sensori-moteurs	151
B	Simulations dans un environnement en deux dimensions	159
	Résumé de thèse	165
	Abstract	171

Table des figures

1.1	Attracteur étrange de Lorenz.	14
1.2	Automate de tessellation.	16
1.3	Simulation 3D d'un automate de tessellation.	17
1.4	Embryogenèse artificielle. Modèle de Dellaert.	18
1.5	Embryogenèse artificielle. Modèle de Beurier.	20
1.6	Embryogenèse artificielle. Modèle de Doursat.	21
1.7	La marche passive.	26
1.8	Continuité dans la complexité des capacités cognitives en lien avec les possibilités de médiation entre un besoin et sa satisfaction.	30
2.1	Architecture d'un réseau de neurones artificiels récurrent standard à temps discret.	36
2.2	Architecture d'un réseau de neurones artificiels complètement récurrent.	37
2.3	Activité d'un CTRNN.	38
2.4	Principe d'un algorithme évolutionnaire.	40
2.5	Fronts de Pareto.	44
2.6	Algorithme NSGA-II.	45
3.1	Architecture du contrôleur d'un hexapode.	50
3.2	Agent phototaxique.	52
3.3	Plasticité homéostatique multi-régions.	54
3.4	Préférence comportementale.	54
3.5	Expérience de Yamauchi et Beer.	57
3.6	Expérience de Izquierdo et al. Structure d'un essai.	59
4.1	Dispositifs de suppléance perceptive.	67
4.2	Dispositifs expérimentaux d'inversion visuelle.	68
4.3	Trajectoire d'un agent, motif carré.	69
4.4	Trajectoire d'un agent, motif sablier.	69
4.5	Deux ontogenèses pour un même agent.	71
4.6	Couplage sensorimoteur.	74
4.7	Transformation progressive d'un comportement sensorimoteur par une action indirecte.	75

4.8	Influence du guidage sur le couplage entre l'agent et l'environnement.	76
5.1	Trajectoires correspondant aux invariants sensorimoteurs recherchés.	82
5.2	Un agent et l'environnement.	83
5.3	Trajectoires du meilleur agent pour la troisième étape.	87
5.4	Trajectoires du meilleur agent pour la troisième étape, modification des conditions initiales liées à la balise 2.	88
5.5	Configurations possibles de l'environnement pour une tâche d'approche de cible.	90
5.6	Comportement sensorimoteur minimal : invariant sensorimoteur.	92
5.7	Cible en mouvement : invariant sensorimoteur.	94
5.8	Distance initiale arbitraire : invariant sensorimoteur minimal.	96
5.9	Courbes d'évolution des scores.	98
5.10	Trajectoires de deux agents pour des essais avec une cible mobile.	100
5.11	Activité interne selon la vitesse.	101
5.12	Basculement de l'activité interne.	101
5.13	Activité motrice de l'agent B.	102
5.14	Activité interne de l'agent B.	102
5.15	Influence de la vitesse sur les trajectoires.	103
5.16	Influence du changement de vitesse sur l'activité de l'agent A.	104
6.1	Un agent et l'environnement.	108
6.2	Structure d'un exercice d'apprentissage associatif.	111
6.3	Activité du moteur "innée" provoquée par un guidage.	113
6.4	Protocole de test de la capacité à apprendre une association par guidage.	114
6.5	Évolution de l'activité sensorimotrice résultant du guidage.	114
6.6	Apprentissage de deux associations différentes.	116
6.7	Dynamique sensorimotrice de l'agent.	117
6.8	Dynamique interne de l'agent en fonction du guidage.	117
7.1	Évolutions des scores pour les comportements "innés".	125
7.2	Comportements "innés" par rapport à la cible et au guidage.	125
7.3	Apprentissage ontogénétique : comportement conditionné par la cible.	128
7.4	Apprentissage ontogénétique : comportement conditionné par le guidage.	129
7.5	Apprentissage ontogénétique : comportement transformé.	129
7.6	Dynamique interne d'un agent conditionné par la cible.	130
7.7	Dynamique interne d'un agent conditionné par le guidage.	131
7.8	Dynamique interne d'un agent perturbé par le guidage.	132
A.1	Invariants sensori-moteurs durant la simulation d'un agent réalisant un carré.	153
A.1	Invariants sensori-moteurs durant la simulation d'un agent réalisant un carré.	154
A.2	Invariants sensori-moteurs durant la simulation d'un agent réalisant un sablier.	156
A.2	Invariants sensori-moteurs durant la simulation d'un agent réalisant un sablier.	157
B.1	Trajectoires non lissées du meilleur agent lors de 10 essais dans les conditions de la phase 3.	159
B.2	Activité des moteurs et des capteurs de l'agent pendant l'essai représenté en noir sur la figure B.1.	160
B.3	Trajectoires du meilleur agent pour la troisième étape, modification des conditions initiales liées à la balise 2.	160

B.4	Activité des moteurs et des capteurs de l'agent pendant l'essai représenté en noir sur la figure B.3.	161
B.5	Trajectoires du meilleur agent pour la troisième étape, modification des conditions initiales liées aux deux balises.	161
B.6	Trajectoires du meilleur agent pour la troisième étape, modification des conditions initiales liées aux deux balises.	162
B.7	Trajectoires du meilleur agent pour la troisième étape, essais longs.	163
B.8	Trajectoires du meilleur agent pour la troisième étape, modification des conditions initiales liées à la balise 2.	163

Liste des tableaux

1	Crédibilité d'un comportement	3
1.1	Comparatif entre le paradigme de l'éaction, le computationnalisme et le connexionnisme.	13
1.2	Principales différences entre l'approche animat et l'intelligence artificielle classique.	24
5.1	Évolution incrémentale de la première étape.	85
5.2	Évolution incrémentale de la deuxième étape.	86
5.3	Évolution incrémentale de la troisième étape.	86
5.4	Comportement sensorimoteur minimal : résultats des évolutions.	91
5.5	Cible en mouvement : résultats des évolutions.	93
5.6	Distance initiale arbitraire : résultats des évolutions.	95
5.7	Maintien d'une distance : résultats des évolutions.	98
6.1	Évolution incrémentale : première étape.	110
6.2	Évolution d'une association par guidage.	113
7.1	Évolution incrémentale de la première étape.	124
7.2	Évolution incrémentale de la deuxième étape.	127
A.1	Carré : temps de passage de l'agent.	151
A.2	Sablier : temps de passage de l'agent.	155

Introduction

LE contexte de ce travail est l'autonomisation d'entités artificielles, qui constitue l'un des axes de recherche de l'équipe ARéVi¹. Pour Tisseau (2001), *l'autonomisation des modèles [...] contribue à peupler les environnements virtuels d'une vie artificielle qui renforce l'impression de réalité*. Face à des entités autonomes, l'homme peut diversifier ses interactions avec elles. Ainsi, donner plus d'autonomie aux entités qui peuplent les mondes virtuels participe à donner également plus d'autonomie à l'homme qui y est immergé.

La ligne directrice qui m'a été donnée au début de ce travail était de creuser la notion d'autonomie telle qu'elle est abordée par les sciences cognitives pour dégager des pistes applicables en réalité virtuelle qui renforcent le principe d'autonomie. Autrement dit le questionnement de départ était : Comment exploiter la notion d'autonomie des sciences cognitives pour rapprocher les préoccupations d'interaction et d'autonomie de la réalité virtuelle ?

La question de l'autonomie peut être envisagée selon deux points de vue :

- ▷ **Du point de vue de l'homme :** La machine peut être considérée comme un outil qui augmente les capacités d'interaction de l'utilisateur. De nouveaux modes d'interaction avec un environnement constituent des possibilités d'acquérir de nouveaux savoir-faire, de nouvelles connaissances dans cet environnement. De ce point de vue, la réalité virtuelle offre les possibilités sensorimotrices nécessaires à la mise en place d'un couplage entre la machine et l'homme qui puisse permettre à ce dernier d'acquérir des savoir-faire dans un environnement virtuel (Bossard et al., 2008). Une première piste que nous pouvons envisager est donc de travailler sur l'interaction entre l'homme et la machine de manière à respecter l'autonomie de l'utilisateur. En réalité virtuelle, cette autonomie est limitée par les contraintes que l'application impose en terme de capacités d'action, de sensation et d'évolution du système. Le travail sur les interfaces éactives² consiste à étudier l'influence de ces contraintes sur l'utilisateur et à

¹ Atelier de Réalité Virtuelle

² Ces interfaces sont le thème des conférences ENACTIVE, organisées depuis 2005 par le réseau ENACTIVE <http://www.enactivenetwork.org/>

améliorer le couplage sensorimoteur entre l'homme et la machine.

- ▷ **Du point de vue des modèles :** La seconde piste envisageable est celle de l'autonomisation des modèles artificiels. Jacques Tisseau décrit cette démarche ainsi :

“ L'autonomisation d'un modèle consiste à le doter de moyens de perception et d'action au sein de son environnement, ainsi que d'un module de décision lui permettant d'adapter ses réactions aux stimuli tant externes qu'internes.”

(Tisseau, 2001) p. 31

En vue d'autonomiser des modèles artificiels, il faut donc les imaginer en tant qu'entités situées dans un environnement virtuel dans lequel l'utilisateur est aussi immergé. L'autonomisation nécessite de leur donner les moyens d'interagir avec l'utilisateur mais également avec les autres entités artificielles. Enfin, elles doivent pouvoir adapter leurs lois de fonctionnement à un environnement dynamique et ouvert en tenant compte à la fois d'états internes et à la fois de cet environnement. à cette condition, l'autonomie des modèles favorisera celle de l'homme.

Nous avons choisi de centrer notre problématique sur cette seconde piste qui est celle de l'autonomisation de modèles artificiels.

Autonomie dans l'interaction

Pour Tisseau (2001), *la réalité nous résiste car elle est autonome par rapport à nous*. Autonomiser un modèle revient à lui permettre de nous résister, de manière à ce *que nous soyons capables de comprendre le modèle (De Loor, 2006)*. En faisant un parallèle avec la phénoménologie, De Loor (2006) tente de préciser cette notion de résistance qui doit faire en sorte que le modèle prenne sens pour l'homme parce qu'il le surprend tout en lui offrant une certaine stabilité. Son comportement se situe alors entre le prédéterminé et l'aléatoire en cela qu'il doit exhiber des régularités lors des interactions tout en étant capable de *désobéir*, de susciter de la curiosité et donc de la motivation.

Le tableau 1 montre comment un comportement autonome *pertinent* est à la frontière entre régularité et variabilité. Sans régularité, il devient insaisissable et il devient impossible de lui donner un sens. à l'inverse, sans variabilité, le modèle ne nous résiste pas et il n'est pas pertinent, en ce sens qu'il ne fait pas de propositions nouvelles³.

Autonomie faible versus autonomie forte

Le *Larousse* donne la définition suivante du terme autonome :

³ Ces considérations concernant la prise de sens dans l'interaction proviennent d'une inspiration phénoménologique et sont développées dans De Loor et al. (2009)

	Comportement prédéterminé	Comportement aléatoire	Comportement pertinent
Variabilité	Nulle	Très forte	Équilibrée
Réponse à l'interaction	Immuable	Absente	Évolutive
Sens	Immédiat	Absent	Saisissable
Sentiment pour l'utilisateur	Lassitude	Lassitude	Engagement
Dynamique	Prévisible	Aléatoire	Chaotique

TAB. 1 – Un comportement pertinent se situe entre l'aléatoire et le prédéterminé (De Loor, 2006). Il doit en effet exister une certaine variabilité dans le comportement, mais qui n'empêche pas de pouvoir l'interpréter, lui attribuant ainsi un sens. L'interaction permet de dégager des régularités qui aident à réaliser cette interprétation. Cependant, elle peut également être source de nouveauté amenant ainsi la créativité propre aux systèmes autonomes.



Autonome : Dont l'évolution est réglée par des facteurs qui lui sont propres.

Dictionnaire Larousse

Cependant, bien qu'en intelligence artificielle, la littérature soit riche de travaux sur les agents autonomes, il n'existe pourtant pas de définition formelle clairement admise de l'autonomie pour un tel agent.

En réponse au manque de consensus sur la notion d'autonomie, Carabelea et al. (2003) proposent une classification d'agents autonomes artificiels établie à partir de travaux issus de la communauté des systèmes multi-agents. Le domaine des systèmes multi-agents est proche de ce qui nous intéresse. En effet, il s'agit de systèmes dans lesquels des agents situés dans un environnement interagissent entre eux et parfois avec un utilisateur. La classification proposée reprend le modèle Voyelles développé par Demazeau (1995), étendu en incluant l'utilisateur par Tisseau (2001). Cette classification distingue des agents :

- ▷ **A-autonomes** Ce sont des agents autonomes par rapport à eux-mêmes, capables de choisir parmi plusieurs comportements le plus adapté à une situation.
- ▷ **E-autonomes** Ce type d'autonomie implique que l'environnement dans lequel l'agent est situé l'influence sans le contrôler.
- ▷ **I-autonomes** La I-autonomie se réfère à l'autonomie sociale. Plus précisément, ici, il s'agit de l'autonomie entre agents dans l'adoption de buts.
- ▷ **O-autonomes** Un agent O-autonome est un agent autonome par rapport à une

organisation dans le sens où il peut violer une norme.

- ▷ **U-autonomes** Un agent est autonome par rapport à un utilisateur s'il peut choisir une action à exécuter sans l'intervention de l'utilisateur.

à partir de cette classification, les auteurs proposent de définir les agents autonomes de la manière suivante :



Agent autonome : Un agent X est autonome par rapport à Y pour p dans le contexte C , si, dans C , son comportement par rapport à p n'est pas imposé par Y .

Carabelea et al. (2003)

L'autonomie est ici donnée comme une propriété relationnelle qui dépend d'un contexte. Si le niveau d'autonomie d'un agent par rapport à Y pour p varie selon le contexte, les auteurs parlent d'autonomie adaptable. Cette définition de l'autonomie propose une perspective externe, celle d'un observateur, car elle ne s'intéresse pas aux mécanismes qui mènent au choix du comportement. Nous l'identifions en tant qu'autonomie faible. Carabelea et al. (2003) évoquent une autre perspective sur l'autonomie qui considère la manière dont les agents sélectionnent un comportement. Cette perspective interne concerne l'architecture de l'agent, le niveau d'autonomie étant ici dépendant du ratio d'influence des facteurs internes ou externes sur le choix du comportement. Bien que les auteurs la développe peu, cette perspective nous semble nécessaire pour clarifier ce qu'est l'autonomie. Les lois de fonctionnement d'un agent entrent ici en jeu dans l'autonomie que l'observateur lui attribue, cependant cette perspective ne s'intéresse pas cette fois aux mécanismes qui régulent ces lois de fonctionnement.

Nous présentons ici une autre définition issue de travaux en robotique. Elle correspond à une troisième perspective, nommée autonomie forte ou autonomie constitutive qui est au cœur du paradigme de l'énaction. Di Paolo et Iizuka (2008) définissent les systèmes autonomes ainsi :



Système autonome : Système composé de plusieurs processus qui génèrent de manière active et maintiennent une identité dans des conditions précaires.

Di Paolo et Iizuka (2008)

Selon cette définition, certains systèmes informatiques, en tant que systèmes complexes adaptatifs composés d'agents humains et d'agents artificiels en interaction peuvent alors être envisagés comme des systèmes autonomes. Par exemple, Hassas (2005) travaille dans le cadre d'applications orientées réseaux sur des systèmes informatiques qui se comportent de manière autonomes en adaptant de manière co-évolutive leur fonctionnement à la structure de l'environnement. L'environnement en question est un réseau dont la topologie oriente l'usage, celui-ci influençant également la topologie du réseau.

Un système a une identité s'il possède les propriétés d'auto-distinction et de clôture opérationnelle. Dans un système opérationnellement clos, chacun des processus est conditionné par au moins un autre processus du système. L'auto-distinction reflète le fait que la

relation de clôture permet de distinguer les processus et les composants qui appartiennent au système. La clôture opérationnelle ne signifie pas que le système est nécessairement isolé de l'environnement. En effet, certains processus du système peuvent être dépendants de conditions extérieures. Dans le cas des systèmes autonomes, l'identité est maintenue dans des conditions précaires. Il s'agit donc de systèmes en interaction avec leur environnement. Ces interactions peuvent perturber le fonctionnement de certains processus et c'est l'existence d'une organisation qui préserve le réseau de processus pour maintenir l'identité.

L'autonomie constitutive ne peut donc pas être abordée en terme de détermination interne versus détermination externe. En effet, un système qui génère ses propres lois, c'est-à-dire qui est autonome, doit se construire lui-même à un certain niveau d'identité. Un système dont l'identité est spécifiée par le concepteur ne peut agir sur lui-même par ses propres moyens et reste donc soumis aux lois de fonctionnement inhérentes à sa conception. Ceci interroge alors la possibilité de construire un système artificiel autonome et rejoint ce que Froese et Ziemke (2009) nomment le *hard problem* de l'intelligence artificielle énaïve, qui se base sur la biologie. Un système autonome est organisé autour du maintien de son identité, qui émerge par l'interaction de ses composants. Quand son identité est menacée, le système se réorganise pour la maintenir. Il a la capacité de modifier ses lois de fonctionnement pour continuer à exister en tant que système distinct de l'environnement. Les questions de l'émergence d'une identité et de sa potentielle dissolution semblent donc être fondamentales dans l'étude de l'autonomie.

De ce point de vue, la synthèse de systèmes artificiels autonomes paraît impossible. En effet, comment concevoir un système autonome que l'on ne peut pas construire de manière externe sans lui retirer cette propriété d'autonomie ? Pour Di Paolo, cette difficulté peut être dépassée en opérant un changement d'attitude dans la conception de systèmes artificiels. Ainsi, en abandonnant la spécification fonctionnelle du système au profit de la recherche de conditions appropriées à la constitution d'une identité autonome par elle-même, la démarche de conception d'une autonomie artificielle constitutive devient envisageable.

Froese et al. (2007) distinguent les approches s'intéressant à une autonomie comportementale, de celles qui s'intéressent à l'autonomie constitutive. L'approche comportementale comprend des travaux en robotique dans lesquels l'autonomie peut se résumer à la capacité à fonctionner sans intervention humaine, des travaux sur des systèmes capables de modifier leurs plans pour satisfaire un but dans des environnements imprévisibles ou encore des travaux sur l'auto-organisation pour la production d'un comportement à la fois robuste et flexible. L'approche constitutive se base sur l'autopoïèse pour définir l'autonomie. Froese et al. (2007) y classent les travaux sur l'autopoïèse computationnelle et ceux sur l'autopoïèse chimique. Ils proposent d'envisager l'autonomie comme un continuum incluant les dimensions comportementale et constitutive et de les lier en les considérant comme deux aspects inter-connectés d'un même concept.

Nous avons vu qu'il y avait deux visions de l'agent autonome : l'autonomie attribuée par un observateur et l'autonomie constitutive. Nous rejoignons Rohde et Stewart (2008) sur le fait que le principe qui permet de produire le comportement est aussi important que le comportement lui-même pour qu'un observateur lui attribue un caractère autonome. De ce point de vue, la validation d'une autonomisation d'agents artificiels nécessite qu'un observateur émette un jugement sur le caractère autonome d'un comportement, mais ce

jugement aura plus de valeur si l'observateur est informé des mécanismes à l'origine du comportement. La conception d'agents autonomes selon cette démarche relève plus des sciences cognitives que de l'ingénierie, dans la mesure où elle cherche à valider des modèles et des hypothèses avec l'ambition de créer des systèmes *les plus autonomes possibles*, plutôt que de concevoir des systèmes avant tout fonctionnels.

Ce premier questionnement sur l'autonomie pour la réalité virtuelle nous amène donc à nous intéresser à un problème qui relève à la fois des sciences cognitives et de l'intelligence artificielle, à savoir la synthèse d'agents artificiels dotés d'une autonomie constitutive. En effet, ce type d'autonomie semble, à terme, être une piste prometteuse pour accéder aux propriétés comportementales que nous avons identifiées dans le tableau 1 pour un comportement autonome pertinent. Cependant, ceci ne sera pas immédiat et le présent travail ne donnera vraisemblablement pas lieu à la construction d'un système de réalité virtuelle peuplé d'agents constitutivement autonomes. Il peut toutefois constituer un pas vers un objectif encore très éloigné d'autonomisation de modèles pour la réalité virtuelle et plus largement pour les systèmes artificiels interactifs.

Organisation du mémoire

Le chapitre 1 s'intéresse à l'autonomie constitutive, à travers son étude théorique, sa simulation dans le domaine de la vie artificielle et comme un principe directeur dans le domaine comportemental. En particulier, nous justifierons par le contexte de la réalité virtuelle notre intérêt pour le comportement, au détriment d'autres propriétés propres à l'autonomie.

Le chapitre 2 énonce et justifie les choix que nous faisons pour la conception d'agents artificiels, dotés d'une autonomie constitutive, à savoir l'utilisation de réseaux de neurones artificiels dynamiques comme contrôleurs d'agents situés, paramétrés selon une approche évolutionnaire.

Le chapitre 3 illustre les possibilités comportementales de l'approche dans laquelle nous nous inscrivons en présentant des travaux directement liés à l'étude de l'autonomie constitutive. Nous développons la question de la possibilité d'évolution du comportement par le biais d'une interaction et nous identifions le besoin d'une étude sur la capacité d'apprentissage de comportements à l'échelle de l'individu.

Le chapitre 4 développe notre proposition pour cette étude. Elle repose sur une transformation d'invariants sensorimoteur provoquée par une interaction, mais ne reposant pas, si possible, sur un objectif prédéfini d'invariant. Ainsi, l'autonomie de l'agent est reflétée par le fait qu'il adoptera potentiellement autant d'invariants qu'il y aura de schémas d'interaction. Cette interaction sera pour nous un guidage.

Le chapitre 5 illustre la notion d'invariant sensorimoteur et démontre la possibilité de complexifier un comportement sensorimoteur en utilisant une approche phylogénétique. Nous faisons évoluer de manière incrémentale des populations d'agents artificiels qui adoptent un

comportement sensorimoteur graduellement plus complexe.

Le chapitre 6 présente une expérience dans laquelle nous initions un apprentissage associatif en utilisant le guidage. L'expérience consiste à faire associer un comportement sensorimoteur simple à la présence d'un stimulus.

Le chapitre 7 décrit une expérience visant à valider notre proposition à travers la synthèse des deux expériences précédentes. Il s'agit ici de complexifier le comportement sensorimoteur d'un agent en utilisant une approche ontogénétique, c'est-à-dire grâce au guidage.

Les expériences présentées n'aboutissent pas à des résultats très concluants et montrent la difficulté à utiliser, dans le cadre de tâches complexes, l'approche évolutionnaire pour paramétrer des CTRNN⁴ que nous avons identifié comme le modèle dont les caractéristiques se rapprochent le plus d'une autonomie constitutive.

⁴ Pour *Continuous Time Recurrent Neural Networks* soit réseaux de neurones récurrents à temps continu.

Chapitre 1

Autonomie constitutive : de la biologie à l'artificiel

L' AUTONOMIE constitutive ou autonomie forte, qui est à l'origine du paradigme de l'énaction, est celle de systèmes qui construisent et maintiennent une identité. Nous avons déjà émis des réserves quant à la possibilité d'obtenir des systèmes artificiels dotés d'une autonomie forte. Cependant, cette perspective sur l'autonomie peut nous guider dans une démarche d'autonomisation de modèles artificiels. L'autonomie constitutive à laquelle nous pouvons nous référer est celle dont nous faisons l'expérience, celle des organismes vivants. Nous nous intéressons dans la section 1.1 aux fondements biologiques de l'autonomie. L'usage de cette notion en réalité virtuelle ne sera sans doute pas immédiat et direct, mais nous verrons dans les sections 1.2 et 1.3 que le domaine de la vie artificielle et certaines branches de l'intelligence artificielle ont créés des liens avec la biologie qui vont dans le sens d'une plus grande autonomie des agents artificiels inspirés de l'autonomie constitutive.

1.1 Fondements biologiques de l'autonomie constitutive

En biologie, l'étude de l'organisation du vivant est classiquement abordée par la génétique. C'est en effet dans l'étude de l'influence des paramètres génétiques sur le développement des organismes vivants que se concentrent la majorité des travaux sur l'organisation des êtres vivants. Cependant, le génotype, c'est-à-dire l'information génétique d'un individu, ne détermine pas de manière directe son phénotype, c'est-à-dire l'ensemble de ses caractères. En effet, les facteurs environnementaux jouent également un rôle dans le développement de l'organisme. Ceci constitue l'objet d'étude de l'épigénétique (Waddington, 1940). Kupiec (2008) propose le concept d'ontophylogenèse qui postule que l'individu se forme comme l'espèce en interaction avec le milieu extérieur et sous la pression du darwinisme. Les questions de l'organisation du vivant et de sa relation avec l'environnement sont au centre des travaux de Varela (1989) et vont l'amener à formuler une théorie du vivant et de la connais-

sance qui constituera le paradigme de l'énaction. Nous trouvons un intérêt particulier dans cette théorie car elle repose sur la notion d'autonomie de l'organisme, c'est pourquoi nous exposons dans la suite quelques uns de ses concepts clés qui nous amèneront à la définition que donne Varela d'un système autonome.

Les systèmes sont des machines dont le fonctionnement dépend d'une organisation. Cette organisation est définie par les relations entre les composants du système indépendamment des propriétés de ces composants. Les relations effectives entre les composants d'une unité constituent ce que Varela appelle la structure.



Organisation : Ensemble des relations qui définissent une machine comme une unité.

Varela (1989)



Structure : Ensemble des relations effectives entre les composants présents sur une machine concrète dans un espace donné.

Varela (1989)

L'étude d'un système passe par la définition de son organisation et de sa structure. De plus, une machine n'ayant pas de but, la finalité d'un système est introduite par un observateur pour l'aider à en comprendre le fonctionnement. Le but n'est pas nécessaire pour décrire la fonction d'une classe de machines, c'est-à-dire d'un type d'organisation.

Les systèmes vivants sont des systèmes stables qui maintiennent des paramètres mobiles. Maturana et Varela (1973) proposent de décrire cette classe particulière de machines en soulignant la circularité de leur organisation. Ceci débouche sur le concept d'autopoïèse, du grec *auto* (soi-même) et *poiësis* (production, création), dont Varela donne la définition suivante :



“Un système autopoïétique est organisé comme un réseau de processus de production de composants qui régénèrent continuellement par leurs transformations et leurs interactions le réseau qui les a produits, et qui constituent le système en tant qu'unité concrète dans l'espace où il existe, en spécifiant le domaine topologique où il se réalise comme réseau. Il s'ensuit qu'une machine autopoïétique engendre et spécifie continuellement sa propre organisation. Elle accomplit ce processus incessant de remplacement de ses composants parce qu'elle est continuellement soumise à des perturbations externes, et constamment forcée de compenser ces perturbations.”

Varela (1989)

L'organisation des machines homéostatiques est constituée par leur environnement et par un circuit de rétroaction sur cet environnement. L'organisation de systèmes autopoïétiques est, elle, décrite seulement en référence à des éléments internes du système. L'identité d'un tel système vient de l'interdépendance des processus. Une machine autopoïétique engendre et spécifie sa propre organisation. Le système se distingue de son environnement par son activité

propre.

Les caractéristiques des systèmes autopoïétiques sont les suivantes :

- ▷ Ils sont autonomes, en opposition aux systèmes allopoïétiques, du grec *allos* (autre) et *poiësis* (production, création), ici le fonctionnement est subordonné à l'autoconservation.
- ▷ Ils ont une individualité, car l'organisation est invariante. Leur identité dépend de leur fonctionnement.
- ▷ Ils forment une unité, le fonctionnement spécifie des frontières.
- ▷ Ils n'ont ni entrées, ni sorties. Les relations entre perturbations et transformations internes n'appartiennent pas au domaine de l'organisation, mais au domaine dans lequel la machine est observée.

Selon Varela, tout système autopoïétique est autonome. De plus, le couplage de systèmes autopoïétiques peut engendrer un système autopoïétique de niveau supérieur si son organisation est dirigée vers son maintien et qu'il constitue une unité dans un espace. Des unités sont couplées dans un domaine si le comportement de l'une est fonction du comportement des autres. Si l'espace dans lequel le système autopoïétique existe est l'espace matériel, le système est vivant.

Si tout système autopoïétique est autonome, tout système autonome n'est pas autopoïétique. Avec le concept de clôture opérationnelle, Varela (1989) étend ses conclusions sur l'autonomie du vivant pour proposer une caractérisation de l'autonomie en général.



“Nous dirons d'un système autonome qu'il est opérationnellement clos si son organisation est caractérisée par des processus : a) dépendant récursivement les uns des autres pour la génération et la réalisation des processus eux-mêmes, et b) constituant le système comme unité reconnaissable dans l'espace (le domaine) où les processus existent.”

Varela (1989)

Le terme *comportement propre* désigne pour Varela (1989) un comportement autodéterminé, c'est-à-dire autonome, caractérisé par les propriétés suivantes :



- “
1. *c'est un état global du système autonome étudié ;*
 2. *il est spécifié en rapport avec la clôture du système ;*
 3. *il exprime une cohérence ou une invariance de l'opération du système ;*
 4. *il est inséparable de l'histoire du couplage structurel du système.*
- ”

Varela (1989)

Sans couplage structurel, le comportement propre d'un système est statique. Il est nécessaire que la structure du système puisse changer pour maintenir l'organisation face à des perturbations mais aussi pour que de nouveaux comportements propres apparaissent. La plasticité de la structure du système est donc essentielle pour le développement de

comportements appris à l'échelle ontogénétique. De plus, ce développement est le résultat de la co-évolution du système et de l'environnement.



“En résumé, les interactions continues d'un système structurellement plastique au sein d'un environnement, source de perturbations récurrentes, produiront une sélection continue au sein des structures possibles du système. Cette structure déterminera, d'une part, l'état du système et le domaine des perturbations permises, d'autre part elle lui permettra de fonctionner, sans se désintégrer, au sein de cet environnement. Nous nommons ce processus le couplage structurel. Si l'on peut considérer l'environnement d'un système comme un système structurellement plastique, le système et son environnement doivent alors être considérés dans l'histoire imbriquée de leurs transformations structurelles, où chacun sélectionne les trajectoires de l'autre.”

Varela (1989)

Un agent cognitif autonome de manière constitutive, est en couplage avec son environnement à travers son “corps”. En ce sens, la connaissance qu'il a de son environnement est “incarnée” et ne peut être en lien qu'avec des contacts directs qu'il a eu avec cet environnement. Ainsi, la connaissance repose sur une base sensorimotrice et non pas symbolique. Ceci est particulièrement développé dans la notion de *morphocomputation* sur laquelle nous reviendrons dans la section 1.3.3. De plus, dans ce cadre, sensation et perception se distinguent. Alors que la sensation est un processus physiologique, la perception est un processus psychologique dynamique se basant sur des sensations qui permet d'accéder à une connaissance directe du monde (McGann, 2010).

Varela propose une théorie de la cognition incarnée, dans laquelle le sens émerge de l'histoire des interactions entre un organisme et son environnement. Le tableau 1.1 compare le point de vue sur la cognition relatif à l'éaction à ceux du computationnalisme et du connexionnisme. Dans le paradigme éactionniste, la connaissance repose sur une boucle sensorimotrice et sur l'autonomie. L'organisme, en absorbant les perturbations, met en place un répertoire de savoir-faire qui donnent du sens à l'environnement. Pour Varela, la brique de base de la connaissance n'est donc pas la représentation symbolique, comme le suppose le paradigme cognitiviste, mais la boucle sensorimotrice. Un agent cognitif peut être défini comme un système autonome, c'est-à-dire opérationnellement clos, incarné, c'est-à-dire pour lequel on ne peut pas faire abstraction des relations sensorimotrices.

Nous venons de voir comment un système autonome peut être envisagé comme un système opérationnellement clos doté d'une dynamique interne qui spécifie son unité. Le couplage structurel du système permet de perturber la dynamique interne du système qui modifie son comportement pour compenser ces perturbations. En effet, le comportement est le produit de l'interaction de la dynamique interne et de la dynamique de l'environnement. Si l'on considère alors le système global, composé du système autonome et de son environnement, l'importance de l'histoire du couplage structurel rend la dynamique de cet ensemble difficilement prévisible. Cette propriété peut être expliquée dans le cadre de la théorie des systèmes dynamiques.

Un système dynamique est un système qui évolue avec le temps. Il est caractérisé par

	Computationalisme	Connexionnisme	Énactionnisme
Métaphore pour la cognition	Ordinateur séquentiel et centralisé	Réseau parallèle et distribué	Dynamique des systèmes non linéaires couplés
Vision du monde	Séparé de l'individu Objectif Symboles	Séparé de l'individu Objectif Patterns d'activation	Inséparable de l'individu Comportements
Définition de la cognition	Traitement de symboles	Émergence d'états	Histoire de couplages structurels

TAB. 1.1 – Comparatif entre le paradigme de l'énaction et les deux principaux paradigmes des sciences cognitives : le computationalisme et le connexionnisme.

un ensemble de variables d'état dont les valeurs changent au cours du temps selon une loi dynamique, qui impose que l'état à un instant t soit dépendant des états précédents, dans le cas d'une loi déterministe. Notons que ceci introduit implicitement la notion de mémoire de l'histoire du système dans son comportement.

Bien que leur évolution soit régie par une loi déterministe, les systèmes dynamiques non linéaires peuvent exhiber des comportements chaotiques. Un système est dit chaotique si son comportement à long terme est impossible à prévoir, alors qu'il obéit à une loi dynamique déterministe. Ceci est dû à une grande sensibilité aux conditions initiales. Ainsi, une légère variation dans les conditions initiales donne lieu à long terme à un comportement sensiblement différent. Cependant, il est possible de déterminer des attracteurs pour des systèmes chaotiques. Par exemple, la figure 1.1 présente les lois d'évolution caractérisant le système dynamique de Lorenz et l'attracteur étrange de Lorenz qui représente l'ensemble des trajectoires du système à long terme (les valeurs des paramètres sont données dans la légende). Ainsi, pour presque toutes les conditions initiales, la trajectoire du système rejoint cet attracteur. Et pourtant, pour des conditions initiales voisines, l'état du système à un instant t est très différent.

En considérant un système autonome comme un système dynamique couplé au système dynamique que constitue son environnement, le comportement de ce système global peut s'exprimer par une trajectoire dans l'espace des phases. La théorie des systèmes dynamiques permet alors une analyse qualitative du comportement de tels agents. Les systèmes chaotiques nous intéressent particulièrement car ils sont régis par des lois dynamiques déterministes tout en présentant une évolution qui semble aléatoire. Ainsi, leur comportement se situe entre l'aléatoire et le préprogrammé, condition que nous donnions dans l'introduction pour qualifier un comportement de pertinent dans son autonomie. Notons également que des perturbations peuvent exister dans le couplage entre les systèmes et accentuer la nature chaotique du comportement de l'ensemble. Dans le cas d'un système autonome, ces perturbations influencent le comportement sans le déterminer de manière directe, de la même façon que les conditions initiales influencent le comportement d'un système dynamique chaotique. Le bruit est donc un élément à prendre en compte dans la production du

$$\begin{cases} \frac{dx(t)}{dt} = \sigma(y(t) - x(t)) \\ \frac{dy(t)}{dt} = \rho x(t) - y(t) - x(t)z(t) \\ \frac{dz(t)}{dt} = x(t)y(t) - \beta z(t) \end{cases} \quad (1.1)$$

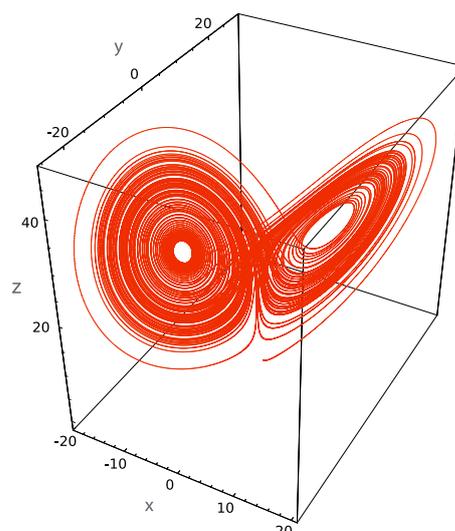


FIG. 1.1 – à gauche : le système dynamique de Lorenz, caractérisé par les variables d'état x , y et z . Dans certaines conditions, le système présente un comportement chaotique. à droite : l'attracteur de Lorenz représente l'ensemble des trajectoires à long terme du système (ici : $\sigma = 10$ $\rho = 28$ et $\beta = 8/3$).

comportement d'un agent autonome basé sur ce type de système.

Nous reviendrons ultérieurement sur cette analogie avec les systèmes dynamiques. Avant cela, nous pouvons désormais proposer une première caractérisation d'agents artificiels dotés d'une autonomie forte, que nous désignerons par le terme *agents énaectifs virtuels*.



Définition AEV-1 : Nous appelons agent énaectif virtuel un système artificiel opérationnellement clos simulé dans un environnement virtuel, auquel il a accès uniquement par une boucle sensorimotrice. Son comportement exprime le maintien d'une organisation interne et est inséparable de l'histoire du couplage structurel du système et de l'environnement virtuel. Il émerge de la dynamique de systèmes non linéaires couplés.

Ceci implique que dans un environnement dynamique ouvert, le comportement de l'agent énaectif virtuel est difficilement prévisible, cette variabilité dans la réponse à l'interaction rend pertinent son comportement pour susciter un engagement chez un utilisateur qui initie l'interaction.

Puisque la théorie de l'énaection repose sur des fondements biologiques de l'autonomie et de la cognition, nous allons dans la section 1.2 aborder son pendant dans le domaine de l'artificiel : la vie artificielle, avant de nous intéresser, dans la section 1.3, à des travaux issus de l'intelligence artificielle et de la robotique qui portent sur la simulation ou la production de comportements sensorimoteurs basés sur une autonomie constitutive.

1.2 L'autonomie constitutive simulée

La synthèse ou la simulation de mécanismes biologiques ont été l'objet de travaux dès les débuts de l'informatique (Turing, 1952; Von Neumann et al., 1966). Néanmoins, Langton (1989) a été l'un d'un premiers à parler de vie artificielle pour désigner ce champ de recherche. Il définit la vie artificielle ainsi :



Vie artificielle : Étude de systèmes vivants naturels pour en trouver des principes afin de recréer sur des supports artificiels des phénomènes biologiques à partir de rien.

Cette section présente différentes approches issues de la vie artificielle qui se réclament ou que nous avons pu rapprocher du courant de l'énaction. En effet, elles respectent les principales contraintes imposées par le paradigme de l'énaction sur l'autonomie qui sont (Varela, 1988) :

1. l'absence de représentations
2. la plasticité du système autonome
3. la co-évolution du système et de son environnement

Nous avons structuré cette partie selon deux axes : la simulation des systèmes autopoïétiques (autonomie par régénération) et l'embryogenèse artificielle (autonomie par évolution). Nous avons retenu ces approches car :

- ▷ Elles proposent un comportement d'entités autonomes qui n'est pas fondé sur une représentation de l'environnement mais dépendant pourtant de ce dernier. C'est le cas des systèmes autopoïétiques artificiels (section 1.2.1).
- ▷ Elles rendent compte du fait que l'environnement perturbe la structure de l'entité qui modifie son comportement en retour. C'est également le cas de certains automates autopoïétiques.
- ▷ Elles proposent des modèles évolutifs. C'est le cas des modèles d'embryogenèse (section 1.2.2), dans lesquels l'entité contient en elle-même tous les mécanismes de son évolution.

1.2.1 Autonomie par la régénération : l'autopoïèse artificielle

Les systèmes autopoïétiques font émerger un phénomène de nature indépendante (le domaine phénoménal) de celle des éléments qui lui ont donné naissance (Laughlin, 2005). Ainsi, il est possible d'utiliser un domaine phénoménal différent de celui de la biologie pour tenter de reproduire certaines caractéristiques d'un système autopoïétique. La simulation offre la possibilité d'utiliser des domaines phénoménaux modelables à loisir par l'homme. C'est pourquoi, pour avancer dans la compréhension, l'identification ou la simulation des mécanismes de la connaissance, une piste consiste à modéliser, simuler et analyser les mécanismes de base de l'autopoïèse (Maturana et Varela, 1984; Beurier et al., 2003; McMullin, 2004; Bourguine et Stewart, 2004; Beer, 2004). Pour cela, les automates ont l'intérêt de reposer sur des mécanismes simples et réactifs stipulant les changements d'état de leurs composants.

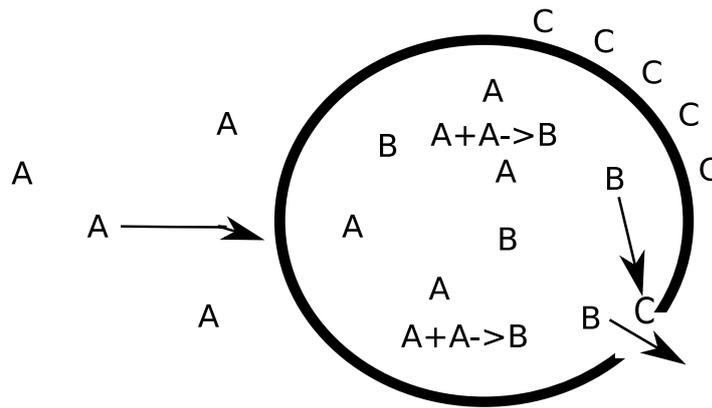


FIG. 1.2 – Fonctionnement d'un automate de tessellation. La membrane est composée d'éléments de type C qui se désintègrent. Les éléments de type B se transforment en C au contact de trous dans la membrane. à l'intérieur du système, deux éléments de type A donnent un élément de type B. *D'après Bourguine et Stewart (2004).*

Pour présenter un système autopoïétique minimal, nous décrivons l'automate de tessellation analysé par Bourguine et Stewart (2004) et décrit par la figure 1.2. Cet automate contient des éléments de trois types : A, B et C. Les éléments de type C forment la membrane du système autopoïétique, les éléments de type B se transforment en C au contact d'un trou de la membrane. La membrane est imperméable aux éléments de type B. Les éléments de type A sont répartis dans tout l'espace et peuvent passer à travers la membrane. à l'intérieur du système, deux A se rencontrant se transforment en B. Le système contient donc un catalyseur pour cette réaction. Enfin, les éléments C ont une durée de vie limitée et se désintègrent, ce qui provoque des trous dans la membrane qui permettent aux éléments de type B de s'échapper. à partir de ce modèle, les auteurs déterminent des solutions d'équilibre où les B générés suffisent à combler les trous. Pour cela, ils formalisent l'automate à l'aide d'un système dynamique stochastique. Le système se maintient à partir du moment où la concentration des éléments ne dépend plus du temps. Le modèle mathématique utilisé repose sur la notion de concentration relative des différents éléments et sur des probabilités de collision entre un élément B et la surface d'un trou. Ce niveau de description rend la résolution mathématique possible. Cependant, si l'on cherche à simuler ces phénomènes, les aspects topologiques et mécaniques rendent bien délicate l'obtention d'un tel résultat.

Ainsi, nous avons simulé l'automate de tessellation dans un espace tridimensionnel (voir figure 1.3). Dans notre simulation, les éléments du système sont des agents situés qui occupent un certain volume dans l'espace et interagissent exclusivement avec leur voisinage proche. Les éléments de type C interagissent entre eux pour former la membrane. Les éléments de type B se déplacent selon un mouvement brownien. S'ils entrent en collision avec un élément de type C, soit ils intègrent la membrane, i.e. ils comblerent un trou, soit ils sont repoussés. En fait, le rapprochement entre le modèle théorique et la simulation est peu aisé.

Les problèmes auxquels nous avons été confrontés dans la simulation sont liés à des notions qui sont implicites dans le modèle de Bourguine et Stewart (2004) et que nous reportons en italique ci-dessous :

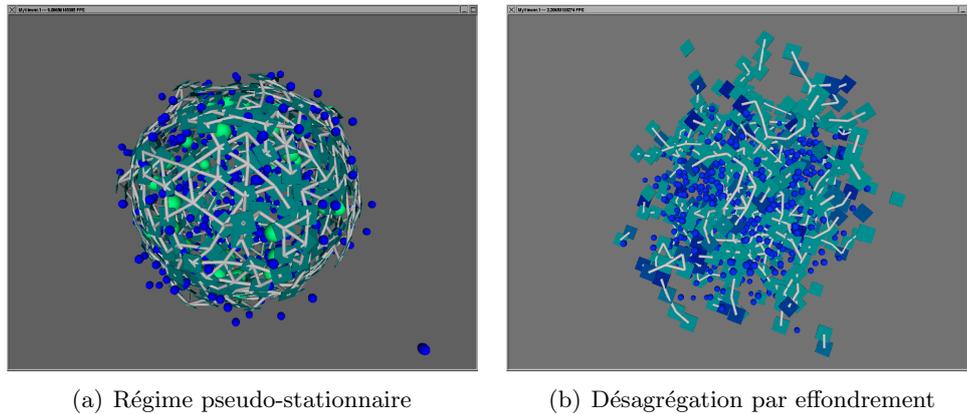


FIG. 1.3 – D’après Manac’h et De Loor (2007). Version tridimensionnelle souple d’un automate de tessellation (a) en régime pseudo-stationnaire, (b) se désagrégant. Les conditions de départ sont : 0 éléments B, environ 480 éléments C. Le nombre d’éléments A reste constant à 500 (a) et 1000 (b).

- ▷ *la réaction $A+A \rightarrow B$ est catalysée à l’intérieur du système. Qu’est-ce que l’intérieur d’un système lorsque sa membrane se désagrège et est constituée d’éléments n’ayant qu’une perception locale ?*
- ▷ *si un élément B entre en collision avec le bord d’un trou de la membrane, celui-ci s’attache à la membrane et répare le trou. Ceci nécessite de détecter des trous ou des ‘non-trous’ puis de positionner les éléments B correctement relativement à leurs voisins qui sont en nombre non défini car dépendant de la taille du trou et de la géométrie de la membrane.*
- ▷ *la membrane est une feuille composée d’éléments C avec un taux d’unité par surface constant. Ajoutons à cela que la membrane forme un cercle. Nous voulons la doter d’une certaine plasticité pour autoriser, par la suite, la possibilité de lui faire subir des déformations et des évolutions. Ceci empêche d’avoir un taux d’occupation constant et l’usage de modèles géométriques simples.*
- ▷ *la membrane intacte est imperméable aux B. Ce point nécessite d’avoir un modèle de surface composée des éléments de type C associé à des mécanismes de détection de collision.*

Pour répondre à l’ensemble de ces problèmes, les agents situés caractérisant les éléments de type C sont reliés à l’aide de ressorts virtuels dont l’orientation et la raideur imposent une forme globale à la membrane. Le passage à la simulation nécessite de donner une forme et une taille aux éléments C. Or, le pavage idéal d’une sphère ne peut se faire qu’avec des triangles sphériques, dont on détermine la forme si on connaît le nombre de triangles et le rayon de la sphère. Pour avoir une bonne approximation d’une sphère de rayon quelconque, on peut jouer sur la taille pour compenser l’imperfection de la forme. Dans notre simulation, les éléments de type C sont des sphères qui peuvent s’intersecter. Nous supposons que la qualité de l’approximation de la sphère, et donc la taille des éléments, influe sur le temps pendant lequel le système se maintient. Pour améliorer la simulation, il faudrait réduire la taille des éléments, et donc augmenter le nombre de ces éléments. Nous sommes cependant

limités par les ressources de calcul disponibles.

Ces choix de modélisation permettent d'obtenir une membrane déformable, mais n'empêchent pas facilement des phénomènes d'agglutination ou de pliage comme cela est illustré figure 1.3. Dans nos simulations, quelles que soient les conditions initiales et les taux de concentration, la membrane finit toujours par se désagréger. Plus précisément, il existe un régime pseudo-stationnaire durant lequel les concentrations se maintiennent voire augmentent. Mais la déformation de la membrane aboutit inexorablement à l'apparition de trous qui ne peuvent être rapidement comblés. Il existe donc un compromis entre le degré de déformation de la membrane et le maintien de l'autopoïèse. Rappelons que dans le modèle de Bourguin et Stewart (2004), la plasticité est ignorée. Pour obtenir cette plasticité, une piste envisageable est de compléter le modèle avec des phénomènes physiques. Le risque est cependant de perdre "l'essence de l'autopoïèse". Il est toutefois possible, dans des environnements en deux dimensions discrétisés, de trouver des règles qui maintiennent l'autopoïèse. Nous pouvons citer par exemple le travail de Beurrier et al. (2003) qui présentent un modèle permettant l'émergence d'une organisation membranaire caractérisant les systèmes autopoïétiques. De plus, Egbert et Di Paolo (2009) ont récemment ouverte une piste intéressante en simulant un automate autopoïétique doté d'un comportement, lié à la réalisation de l'autopoïèse mais pourtant distinct. Dans leur modèle, des réactions permettent la régénération de la membrane qui se désagrège alors que d'autres provoquent des déformations qui entraînent un mouvement.

1.2.2 Autonomie par l'évolution : l'embryogenèse

Dans cette partie nous présentons les approches qui s'intéressent à la modélisation de l'évolution d'un système à partir d'une cellule unique. L'idée générale de ces approches est de trouver un mécanisme qui permette d'engendrer un développement, puis un paramétrage de ce mécanisme pour que celle-ci tende vers un état final connu.

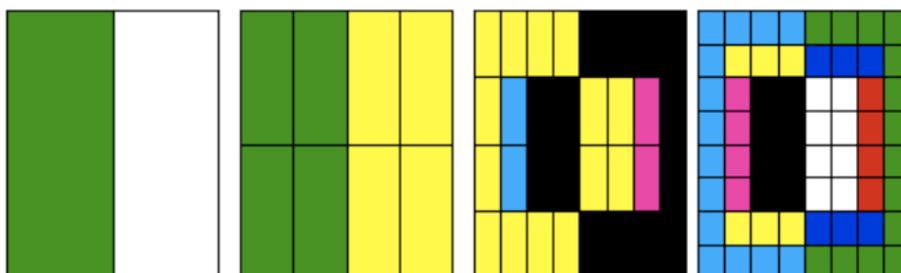


FIG. 1.4 – D'après Dellaert et Beer (1994). Développement de l'organisme le mieux adapté sélectionné par l'algorithme génétique. Les différentes étapes du développement sont présentées. Les cellules se divisent et se spécialisent en fonction de leur voisinage. La couleur d'une cellule représente son état défini par le régulateur génétique.

Dellaert et Beer (1994) annoncent que leur objectif à long terme est la co-évolution du corps du système et de son contrôle. Dans ce modèle, toutes les cellules d'un même organisme partagent un même régulateur génétique mais le comportement issu de ce régulateur est dépendant du voisinage des cellules. Les cellules sont représentées graphiquement par des cases sur une grille dont la couleur est dépendante de leur état défini par le régulateur génétique. La couleur est une abstraction de la spécialisation des cellules : à terme, une couleur représentera le fait que cette cellule soit un capteur, un moteur ou une unité de contrôle. Au niveau de l'organisme, ce sont les couplages entre les cellules qui sont définis (échanges entre régulateurs de cellules voisines). Un algorithme génétique a pour tâche de sélectionner le régulateur qui permet, suite à l'évolution, d'atteindre une configuration particulière. La complexité potentielle de ces régulateurs, décrits à l'aide de réseaux booléens, est très élevée et ouvre une très large gamme de comportements possibles. Cependant, ces comportements restent pour l'instant abstraits. La figure 1.4 présente un exemple d'embryogenèse simulée à l'aide de ce modèle.

Beurier et al. (2006) s'inspirent de l'embryogenèse pour produire des "organismes" multi-agents à partir d'un seul agent. Le modèle est basé sur des agents réactifs capables de se répliquer et d'évoluer. Les gènes des agents sont composés d'une partie codante et d'éléments régulateurs. Les éléments régulateurs influent sur la transcription de la partie codante. Dans le modèle, l'environnement joue le rôle de l'oeuf, dans lequel des gradients de morphogènes sont diffusés, qui vont agir comme facteurs de transcription pour certains gènes. Les agents jouent le rôle de cellules et leur génome est composé de gènes de segmentation et de gènes homéotiques (voire figure 1.5). Bien que tous les agents aient le même génome, leur comportement diffère selon les morphogènes qu'ils perçoivent. Dans cette approche, l'organisme évolue en couplage structurel avec son environnement. Les gradients de morphogènes ont une influence sur le comportement des agents, qui influence en retour les gradients de morphogènes.

Doursat (2009) propose un modèle d'embryogenèse artificielle basé également sur un réseau artificiel de régulation génétique dans lequel un système grandit et s'organise grâce à des instructions génétiques qui contrôlent le déclenchement de mécanismes qui s'influencent circulairement : l'auto-assemblage et la formation de motifs (voir figure 1.6). L'embryogenèse est le fruit de mécanismes d'auto-organisation par la division cellulaire et l'adhésion entre cellules ainsi que de mécanismes de régulation par la diffusion de gradients et la spécialisation des cellules.

Le modèle *Cell2Organ* de Cussat-Blanc et al. (2009) permet de simuler une embryogenèse artificielle qui intègre un métabolisme. Il produit une morphologie et dote également l'organisme de fonctions. La question de la spécialisation des cellules est donc abordée véritablement au niveau de l'effet et non pas simplement au niveau d'une catégorisation.

Les travaux issus de la vie artificielle que nous venons de présenter sont liés plus ou moins directement à l'érection car ils illustrent une manière d'étudier l'autonomie artificielle à travers l'émergence ou le maintien d'un système. Cependant, les systèmes en question ont une influence très limitée sur leur environnement et leur comportement se résume à rester stable dans le cas des automates autopoïétiques ou à se constituer dans le cas des approches embryogénétiques. Les travaux sur l'embryogenèse présentés amorcent des pistes intéressantes vers l'obtention d'organismes dotés de comportements. Récemment, c'est également le cas

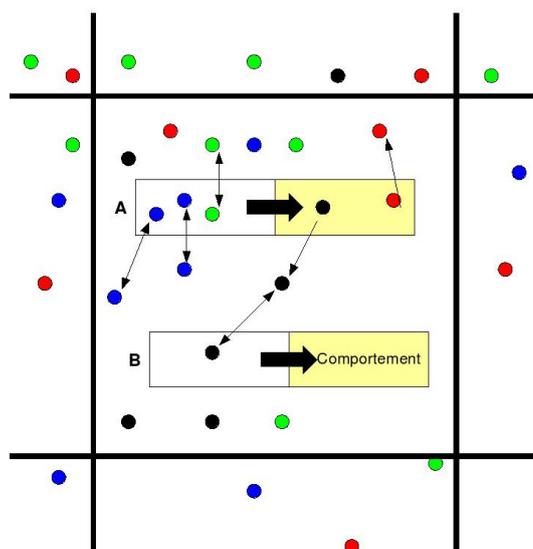


FIG. 1.5 – D’après Beurier et al. (2006). Un gène est composé d’éléments régulateurs (partie gauche) et d’une partie codante (partie droite). Si les morphogènes (points colorés) correspondant aux éléments régulateurs d’un gène sont présents en quantité suffisante, la partie codante est activée. Pour un gène de segmentation (A), de nouveaux morphogènes sont alors libérés. Pour un gène homéotique (B), un comportement est déclenché.

en ce qui concerne les systèmes autopoïétiques avec la proposition de Egbert et Di Paolo (2009). Ces travaux ne permettent néanmoins pas pour l’instant d’obtenir d’agents, au sens d’entités mettant en œuvre des comportements interactifs intégrant l’équivalent de choix, d’apprentissage ou encore de localisation dans l’espace. Or, nous cherchons l’autonomisation d’entités artificielles situées dans un espace virtuel dans le but de pouvoir interagir avec celles-ci. La question du comportement, traitée par le domaine de l’intelligence artificielle, est dès lors essentielle pour nous. C’est pourquoi la section 1.3 s’attache à présenter certains courants de l’intelligence artificielle qui se sont orientés vers la mise en place de comportements sensorimoteurs basés sur l’autonomie forte chez des agents artificiels.

1.3 Vers une autonomie constitutive comportementale

En Intelligence Artificielle, l’approche prédominante issue du cognitivisme est questionnée depuis plusieurs années. Pour Dreyfus (1984), les présupposés sur lesquels elle se base la voue à l’échec. De plus, cette approche, que nous appellerons dans la suite IA classique¹, a délaissé ses objectifs initiaux devenant ainsi une Informatique Avancée concentrée sur la performance de la résolution de problèmes au détriment de la recherche des conditions

¹ Également nommée GOFAI pour *Good Old-Fashioned Artificial Intelligence*

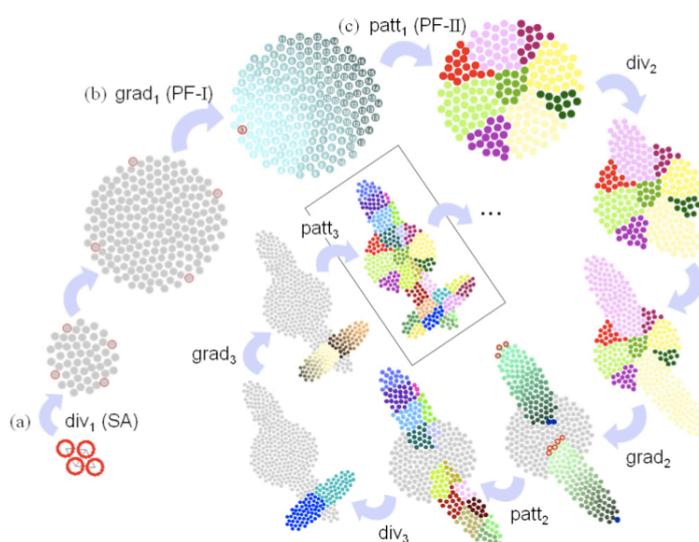


FIG. 1.6 – D’après Doursat (2009). Simulation d’une embryogenèse dans laquelle les mécanismes d’auto-organisation et de régulation s’influencent et permettent la formation d’un organisme complexe. à partir de quelques cellules, trois opérations sont appliquées : (a) “div” représente le comportement d’auto-assemblage (SA) à savoir la division, la prolifération et le ré-arrangement spatial ; (b) “grad” représente une première part du comportement de formation de motif (PF) correspondant à la diffusion des morphogènes ; et (c) “patt” représente la second part du comportement de formation de motif à travers l’extraction et la transformation des signaux de gradient en niveaux d’expression des gènes.

nécessaires à la constitution d'une véritable intelligence. D'autres approches en intelligence artificielle sont cependant apparues qui rejettent les prémisses de l'IA classique et dans lesquelles la notion de cognition est proche de celle de comportement. Selon les orientations, ces approches se qualifient de Nouvelle Intelligence Artificielle, d'approche animat, de Nouvelle Robotique ou encore de robotique inspirée organismes. Nous présentons ici leurs spécificités respectives afin de réaliser une synthèse des liens existants avec la notion d'autonomie constitutive.

1.3.1 La Nouvelle Intelligence Artificielle

La *Nouvelle Intelligence Artificielle* se base sur l'hypothèse de l'*ancrage physique* des représentations, par opposition à l'hypothèse symbolique (Simon, 1996)². Selon Brooks (1990), pour construire un système intelligent, il est nécessaire qu'il utilise des représentations ancrées dans le monde physique. Il préconise d'abandonner les représentations symboliques et l'idée que la prise de décision consiste à manipuler ces symboles par des règles, pour passer à des systèmes incarnés qui utilisent l'environnement comme son propre modèle. Le système est pourvu de capteurs et d'effecteurs qui lui permettent de sentir le monde et d'agir sur celui-ci. Un système ancré physiquement tire ses connaissances directement du monde à travers ses capteurs et exprime ses buts par des actions physiques.

Brooks (1991b) argumente ainsi contre l'utilisation d'une représentation interne du monde. En effet, pour lui, le monde est son meilleur modèle car la représentation interne d'un environnement par nature dynamique ne peut pas refléter de manière fidèle les changements qu'il subit. Le système se retrouve alors dans des situations où le modèle avec lequel il compose diffère de l'environnement dans lequel il est, ce qui entraîne des comportements inadaptés. En prenant l'environnement qu'elle perçoit comme modèle, une créature peut faire face de manière immédiate et donc plus adaptée aux changements que cet environnement subit.

L'absence de représentations doit cependant être nuancée. Pour Brooks, l'intelligence artificielle va trop vite et ne s'attaque pas aux fondements de l'intelligence. Il est pourtant indéniable que certaines capacités cognitives font appel à des représentations. La question de la nature et de l'origine de ces représentations est plus problématique. Gigliotta et Nolfi (2008) montrent par exemple que des robots peuvent développer des formes de représentations spatiales sans utiliser de symboles, simplement en intégrant une information sensorimotrice. Dans leur expérience, la capacité pour des robots à se localiser et à distinguer différents lieux dans leur environnement provient du couplage des dynamiques internes et environnementales des robots. Il y a bien représentation spatiale, mais sous forme d'invariants sensorimoteurs plutôt que sous forme symbolique.

La Nouvelle IA se distingue aussi de l'IA Classique dans sa manière de concevoir la production d'un comportement global³. Cette approche a l'avantage de renoncer à l'hypothèse

² Cette hypothèse (en anglais *Physical symbol system hypothesis*) considère que tout système symbolique physique a les moyens nécessaires et suffisants pour agir de manière intelligente.

³ D'un côté, la vision classique décompose l'intelligence en modules fonctionnels, qui traitent de l'information et vont pouvoir être combinés de manière centralisée pour produire un comportement global. De

symbolique permettant ainsi au comportement d'être le reflet direct des connaissances du système dans le sens où il n'y a de signification que dans la relation entre les sensations et les actions du système. Cependant, si Brooks n'envisage plus d'entrées/sorties typées (c'est-à-dire avec un pouvoir représentatif), le couplage est toujours pensé en terme d'entrées et de sorties. La clôture opérationnelle propre à une autonomie constitutive n'apparaît pas dans le fonctionnement du système.

1.3.2 L'approche animat

L'approche animat (Meyer (1996)) partage les critiques de Brooks face à l'IA classique. Cependant, elle se présente comme une approche complémentaire. Elle n'a pas pour ambition de s'attaquer à la résolution de problèmes ou à la planification, mais se concentre sur les interactions corps/"esprit" ou organisme/environnement. L'approche animat consiste à étudier chez des animaux simulés les mécanismes d'adaptation dans un environnement imprévisible. Le tableau 1.2 résume les principales divergences entre cette approche et l'intelligence artificielle classique.

L'approche animat s'est tout d'abord construite autour du paradigme connexionniste, l'objectif étant de simuler les interdépendances entre le corps, le système nerveux et l'environnement en utilisant des interactions entre multiples unités de codage. Cependant, pour Guillot (1999), la causalité linéaire que comporte le connexionnisme est inadaptée pour la simulation de créatures autonomes, pour qui le sens des informations provenant de l'environnement est donné à la fois par ce qui est capté et à la fois par ce qui est effectué. Il y a donc nécessité de prendre en compte une causalité circulaire, dans laquelle l'action influence la perception. Ce principe de causalité circulaire est à la base de l'approche dynamique⁴, qui constitue pour Guillot (1999) un paradigme adapté à l'approche animat car plus apte à simuler des animats réellement situés.

l'autre côté, la Nouvelle IA propose une décomposition de l'intelligence en modules qui produisent des comportements simples de manière quasi-réactive et font émerger un comportement plus complexe.

⁴ Selon Giraud (2006), l'émergence de l'approche dynamique a permis à la psychologie et aux sciences du vivant en général de s'intéresser de nouveau à la variabilité intra-individuelle. Plus globalement, elle s'accompagne d'un renouveau de l'étude de la complexité aussi bien en physique, qu'en mathématiques, en sciences biologiques, en sciences cognitives ou encore en sciences humaines et sociales. Cette complexité ne peut en effet pas être abordée par une approche analytique. En s'intéressant à la dynamique de fonctionnement d'un système, l'analyse temporelle devient centrale et permet l'étude de systèmes adaptatifs. L'approche dynamique consiste à aborder l'étude d'un système non plus dans une analyse de ses composants et de leurs interactions de manière globale, mais dans une analyse de l'évolution temporelle du système. Par exemple, elle peut s'illustrer en psychologie par le passage d'observations par moyennes à des calculs d'auto-corrélations qui permettent de mettre en évidence une évolution.

Intelligence artificielle classique	Approche animat
Approche descendante : étude des capacités cognitives élaborées propre à l'homme	Approche ascendante : étude des capacités adaptatives basiques des animaux et des robots
Systèmes qui utilisent leurs connaissances pour résoudre un problème donné	Systèmes qui agissent pour résoudre un problème donné
Application à des problèmes limités et bien définis	Application à des problèmes généraux et éventuellement mal définis
Contrôle centralisé et calculs séquentiels	Structures décentralisées et calculs parallèles
Fonctionnalités programmées	Fonctionnalités émergentes
Systèmes fermés. Environnements contrôlés par l'expérimentateur qui généralement fournit les données d'entrées	Systèmes ouverts. Agents autonomes qui interagissent en temps réel avec un environnement imprévisible, sans intervention humaine. Approche située.
Intérêt limité pour la perception et l'action	Intérêt majeur pour les problèmes de coordination perceptions / actions
Intérêt limité pour l'apprentissage	Intérêt majeur pour les mécanismes adaptatifs liés à la maturation, au développement, à l'apprentissage et à l'évolution
Intérêt majeur pour les causes proximales (le comment) des comportements intelligents	Intérêt majeur pour les causes ultimes (le pourquoi) et les causes proximales (le comment) des comportements adaptatifs

TAB. 1.2 – Principales différences entre l'approche animat et l'intelligence artificielle classique. D'après Meyer (1996).



“Si l’on envisage un animat et son environnement comme deux systèmes dynamiques séparés, on les considérera comme couplés, chaque système déformant continuellement la trajectoire de l’autre par le biais de liens sensorimoteurs. Les stimulations de l’environnement vont en effet influencer l’animat, mais celui-ci va modifier en retour l’environnement par les actions qu’il aura produites. Ainsi, le comportement d’un animat ne s’expliquera pas simplement par ses mécanismes internes, ni par sa propre dynamique, mais par la dynamique de l’ensemble des deux systèmes couplés.”

Guillot (1999) p.38

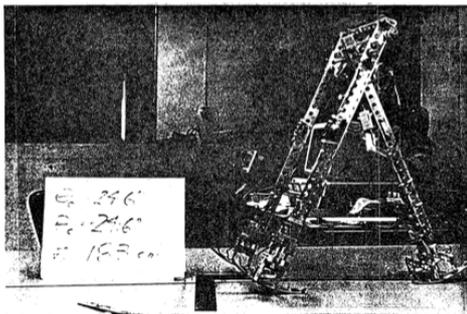
Avec l’approche animat, l’IA fait un pas de plus vers l’autonomie constitutive comportementale. En effet, la boucle sensorimotrice permet comme pour la Nouvelle IA de se passer de représentations symboliques. De plus, l’objet d’étude ici sont les mécanismes adaptatifs. Il est donc question de systèmes capables de se réorganiser pour faire face à des changements de l’environnement. Cependant, la clôture opérationnelle n’est pas posée en tant que principe pour la définition d’un animat, l’unité du système est donc définie de manière extérieure et ne se reflète pas de manière systématique dans son fonctionnement.

1.3.3 La Nouvelle Robotique

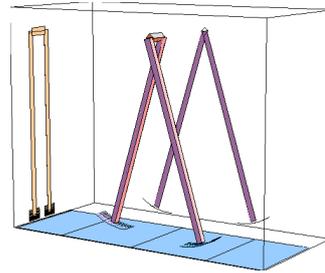
Pfeifer et al. (2005) proposent une synthèse des principes de conception d’agents autonomes incarnés en insistant sur les concepts de balance écologique et de coordination sensorimotrice. Ces auteurs, qui nomment leur approche la *Nouvelle Robotique*, poussent la démarche synthétique plus loin encore en considérant des agents non seulement situés, mais surtout incarnés, dans le sens où le corps prend pleinement part à la production des comportements. Ainsi, la *morphocomputation* participe à diminuer la complexité des comportements. Dans la nature, un corps est adapté à une niche écologique, il contraint les comportements et en facilite certains. Il est également souhaitable de donner une place importante au corps dans la conception d’agents artificiels car cela augmente la coordination sensorimotrice passive et réduit donc les besoins de contrôle actif pour produire un comportement. La marche passive (*cf.* figure 1.7) est une belle illustration de ce principe, car elle permet de démontrer la possibilité d’obtenir un comportement, en l’occurrence la marche, sans contrôle actif et sans avoir besoin de fournir de l’énergie au système, en utilisant uniquement les propriétés physiques du corps.

1.3.4 La robotique inspirée organisme

Les approches inspirées de la biologie ne peuvent pas permettre selon Di Paolo (2003) de construire des robots réellement autonomes. Il manque l’intentionnalité, la génération de buts propres. Dans une posture énonciationniste, Di Paolo propose de s’inspirer non plus de la biologie seulement mais de l’organisation des organismes. Pour lui, ce qui distingue actuellement un robot d’un organisme vivant, c’est le fait que l’organisme génère lui-même ses buts, qu’il



(a) Premier marcheur passif.



(b) Principe d'architecture d'un marcheur passif bipède.

FIG. 1.7 – (a) Première machine exhibant un comportement de marche sans contrôle actif, ni injection d'énergie. D'après McGeer (1990) (b) Illustration de l'architecture d'un marcheur passif bipède. D'après Kuo (1999)

existe une relation réciproque entre ce qu'il est et ce qu'il fait ou subit. L'organisme vivant est un système téléologique, il est autodéterminé, capable de générer et modifier ses buts ; les robots sont des systèmes téléonomiques, qui maintiennent un but prédéfini. Pour Jonas (1966), c'est le métabolisme qui distingue les organismes vivants d'autres systèmes, dans le sens où c'est le fonctionnement de l'organisme qui définit son identité.



Téléologie : Du grec telos : fin, but et logos : discours. Doctrine qui considère que dans le monde tout être a une fin, qui conçoit le monde comme un système de relations, de rapports entre des moyens et des fins.

Téléonomie : Du grec telos : fin, but et nomos : loi. Forme de finalité admise dans les adaptations biologiques en tant que le but concernerait une utilité immédiate sans considération d'un avenir plus ou moins lointain

Trésor de langue française informatisé

La difficulté de la modélisation est de construire des modèles basés sur une identité autodéterminée de l'organisme et non plus sur une identité continue provenant de la permanence des composants. La structure d'un organisme vivant évolue, l'organisation aussi ; la mort d'un organisme est la perte de son identité, il n'y a plus de système mais un ensemble de composants. Pour Di Paolo, la modélisation par la simulation est une réponse possible à ce problème car elle permet de considérer différents niveaux d'organisation.

La finalité intrinsèque d'un organisme est l'auto-conservation. Il donne un sens à une situation par rapport aux conséquences qu'elle a sur la conservation des habitudes de vie. Même si l'auto-conservation est essentielle, la conservation des habitudes est importante aussi, car elle permet d'anticiper une mise en danger de l'auto-conservation. Une hypothèse intéressante pour Di Paolo est que l'adaptation a lieu non pas parce que la survie est menacée directement mais parce que le processus circulaire de génération des habitudes est perturbé. Pour lui, pour générer une intentionnalité chez les robots, il peut suffire de les doter de mécanismes leur permettant d'acquérir des habitudes.

1.4 Synthèse

Dans ce chapitre, nous avons présenté des éléments d'une théorie biologique de l'autonomie : la théorie de l'énaction. La biologie, en tant qu'outil pour mieux comprendre l'autonomie constitutive, peut être intégrée aux sciences cognitives. à partir de ces éléments, nous avons proposé une première définition d'agent énaactif virtuel (définition 1.1), qui nous permet d'établir une liste de propriétés internes pour un agent énaactif virtuel :

▷ **La clôture opérationnelle du système.**

“Nous appelons agent énaactif virtuel un système artificiel opérationnellement clos”

L'agent énaactif est un système autodéterminé, il maintient une identité de manière active. Il possède donc une dynamique propre et une organisation récursive.

▷ **La situation dans un environnement.**

“simulé dans un environnement virtuel,”

Les agents qui nous intéressent sont simulés et à ce titre la notion d'incarnation est problématique. Nous considérons que l'équivalent pour la simulation comprend le fait que l'agent soit localisé dans l'environnement virtuel, qu'il puisse le sentir et qu'il ait des possibilités d'action dans cet environnement. De plus, l'agent n'a pas accès à un modèle global de l'environnement, mais seulement aux éléments de cet environnement avec lesquels il peut interagir.

▷ **L'ancrage sensorimoteur des représentations.**

“auquel il a accès uniquement par une boucle sensorimotrice.”

Un agent énaactif n'utilise pas de modèle de l'environnement. Le “monde” de l'agent émerge à travers l'histoire de ses interactions avec l'environnement. Les sensations sont fournies par des capteurs qui contraignent le type de données senties. L'agent énaactif est doté d'effecteurs qui lui permettent d'agir sur son environnement. Ces actions ont des conséquences sur les perceptions futures. Les interactions ne sont pas abordées en terme d'entrées/sorties mais au travers de l'existence d'une boucle sensorimotrice fermée.

▷ **La capacité d'autorégulation.**

“Son comportement exprime le maintien d'une organisation interne”

L'autonomie forte suppose que le comportement soit le résultat d'une réorganisation d'un système en réponse à des perturbations. La cognition n'est alors pas envisagée en terme de traitement d'une entrée qui produit une sortie, mais plutôt comme un ensemble de savoir-faire qui permettent à un système de se maintenir en interaction avec un environnement.

▷ **La sensibilité aux conditions environnementales qui rend le comportement non prévisible.**

“et est inséparable de l'histoire du couplage structurel du système et de l'environnement virtuel.”

Un comportement résultant d'une autonomie forte résulte de l'interaction entre une dynamique interne et celle d'un environnement complexe. L'interaction entre ces deux

systèmes peut engendrer une dynamique imprévisible et chaotique.

▷ **Le caractère émergent du comportement.**

“Il émerge de la dynamique de systèmes non linéaires couplés.”

Une décomposition fonctionnelle ne peut pas rendre compte du caractère émergent du comportement d'un agent, qui reflète un état global. Elle ne peut être qu'un modèle qui permet l'analyse et la compréhension, mais ne capture pas la manière dont est produit le comportement. L'utilisation de modules spécialisés est donc inadaptée au paradigme de l'énaction, elle ne peut pas à un niveau opérationnel générer une autonomie constitutive.

Nous avons ensuite présenté des travaux de vie artificielle qui simulent la constitution d'organismes en s'appuyant sur des mécanismes biologiques. Cependant, pour rendre compte de comportements sensorimoteurs, nous sommes amenés à mettre de côté la question de l'auto-production de l'identité.

Nous avons vu différents courants qui se situent entre la robotique, la vie artificielle et l'intelligence artificielle et qui ont en commun d'être dans une démarche de synthèse d'agents artificiels tentant de ne pas reposer sur une autonomie faible. Cela permet d'étudier des comportements plus complexes mais impose de faire l'impasse sur la question de la génération de l'identité.

L'intelligence artificielle est ainsi passée d'une approche cognitiviste, proche des modèles introspectifs mais éloignée des neurosciences, à une approche connexionniste, utilisant le cerveau comme guide mais ignorant la nature dynamique et incarnée de la cognition, pour adopter une approche réactive, située et incarnée, dans laquelle l'interaction avec le monde se fait dans une boucle sensorimotrice fermée.

La théorie de l'énaction est souvent cantonnée aux aspects bas-niveau de la cognition, puisqu'elle met le couplage structurel, l'engagement incarné et situé d'un agent au cœur de celle-ci et ne pourrait donc expliquer que les comportements sensorimoteurs. Pour Di Paolo et al. (2010), au contraire elle permet d'envisager la complexité de la cognition dans sa continuité. Ces auteurs mettent en avant une continuité vie-esprit qui peut s'exprimer à travers les mécanismes créateurs de valeurs pour le système.

Une entité cognitive est un système téléologique qui dans la recherche de la conservation de valeurs poursuit ses propres buts. Les possibilités de médiation qui existent pour l'organisme entre un besoin et sa satisfaction lui permettent d'adopter des valeurs liées aux différentes identités qu'il se constitue. L'autopoïèse correspond à un niveau minimal de téléologie, dans lequel la valeur est la conservation du système. La proposition de Weber et Varela (2002) fait le lien entre l'organisation du vivant, à travers l'autopoïèse, et le *sense-making* pour doter les organismes d'une téléologie intrinsèque : les interactions avec le monde ont une signification pour l'organisme à travers une norme établie par le processus ininterrompu d'auto-production.

Selon Di Paolo (2005), ce lien ne peut être fait qu'en tenant compte d'un autre aspect du vivant qui est implicite dans l'interprétation ci-dessus : la capacité d'un organisme à se réguler lui-même en fonction des limites de sa propre viabilité. Pour Weber et Varela, l'autopoïèse

implique une téléologie de l'organisme selon deux sens complémentaires : tout d'abord, une finalité naturelle définie par les relations mutuelles génératives entre les composants et entre ces composants et le tout ; ensuite, l'instauration d'un point de vue et d'une individualité qui, à travers l'auto-production, est un lieu de sensation et d'agentivité. En effet,

1. L'auto-production est un processus qui définit une unité et une norme, qui est d'entretenir l'unité et de la garder distincte.
2. Les rencontres avec le monde extérieur sont évaluées par le système comme contribuant ou non au maintien de l'autopoïèse.
3. L'autopoïèse implique le *sense-making*, une perspective intrinsèque sur le monde lui donnant de la valeur.

Or, Di Paolo remarque que la définition originale de l'autopoïèse ne permet pas d'admettre de manière évidente le second point, car elle ne permet pas de parler de but pour un système autopoïétique qui est, soit adapté et viable, soit inadapté et non viable. La conservation est une conséquence et non un but. Si le but du système est de se conserver, il doit avoir les moyens d'anticiper la menace de sa viabilité et d'y réagir. L'agentivité, ainsi qu'une capacité d'auto-monitoring, sont donc nécessaires pour introduire une finalité. Alors que le système autopoïétique subit un couplage structurel avec son environnement, un système adaptatif qui régule ce couplage de manière active peut maintenir une norme qu'il perçoit comme menacée. Ceci fournit une définition du comportement en tant qu'action sur les paramètres du couplage structurel qui introduit une asymétrie dans les échanges entre le système et l'environnement. Cette asymétrie est nécessaire selon Di Paolo pour parler d'agent.



Agent : Du latin agens, participe présent du verbe ago : agir. Unité auto-construite qui investit le monde en régulant de manière active ses échanges avec lui à des fins adaptatives qui doivent servir sa viabilité.

Traduit de Di Paolo (2005)

La figure 1.8 présente différentes étapes que Di Paolo et al. (2010) identifient comme particulièrement importantes dans la complexité croissante d'une entité cognitive en lien avec des mécanismes qui augmentent les possibilités de médiation entre un besoin et sa satisfaction.

Si un système autopoïétique est capable de résister à des perturbations, il ne surveille cependant pas le respect de la norme et n'adapte par conséquent pas ses lois de fonctionnement s'il s'en écarte. Il n'y a pas de régulation en fonction de l'état du système. La capacité d'adaptation du fonctionnement constitue une étape de transition dans les possibilités de médiation du système avec son environnement, qui rend possible la constitution du sens. Les perturbations ayant un impact sur la conservation d'une norme pour le système font alors sens pour celui-ci car elles vont être favorables ou défavorables. L'étape suivante est la régulation interactive du système qui agit sur l'environnement pour maintenir une norme et utilise les lois de fonctionnement de l'environnement pour suivre un but qui lui est propre. Cette capacité de régulation permet la forme la plus fondamentale de l'agentivité. Le système nerveux apporte la mobilité au système, lui permet d'interagir avec des objets distants et de ressentir des émotions comme la peur ou le désir. La capacité humaine à former des images, c'est-à-dire à construire des représentations, autorise l'examen objectif de situations. Enfin, la constitution d'une image de soi-même donne accès à l'élaboration de projets.

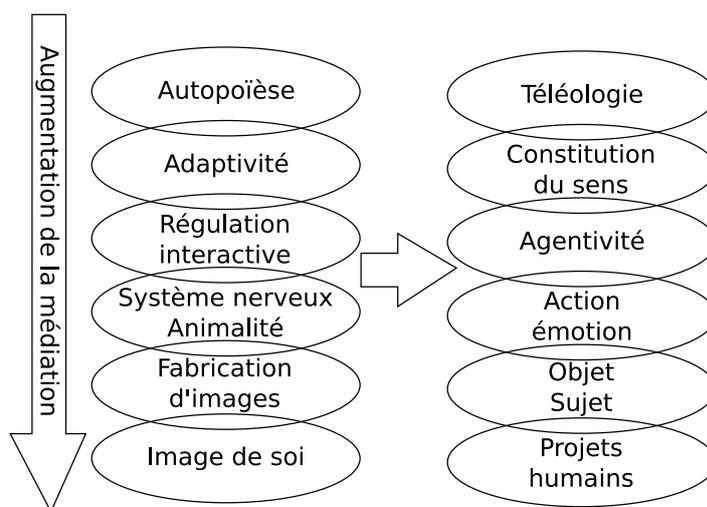


FIG. 1.8 – D'après Di Paolo et al. (2010). Lien entre la complexité des capacités cognitives d'un organisme et les mécanismes dont il dispose pour satisfaire ses besoins. Pour Di Paolo, cette complexité est croissante. Des étapes particulièrement importantes selon lui sont ici représentées.

Notre objectif premier n'est pas d'atteindre des capacités cognitives très élaborées, mais d'avancer dans la synthèse d'agents artificiels dotés d'une autonomie constitutive. Nous chercherons donc à obtenir des entités qui possèdent au minimum la capacité de réguler leur fonctionnement de manière interactive en respectant au maximum les idées issues de la théorie de l'énaction.

En conclusion de ce chapitre, la définition que nous adoptons pour un agent éactif virtuel devient :



Définition AEV-2 : Nous appelons agent éactif virtuel un système dont le fonctionnement le définit comme une unité. Il est en couplage avec son environnement et régule de manière interactive son fonctionnement pour suivre une ou plusieurs normes. Un agent éactif virtuel maintient et génère ses propres normes/buts.

De plus, nous avons mises en évidence des propriétés caractéristiques d'une autonomie constitutive. Certaines, que nous avons données précédemment, sont liées au fonctionnement du système, elles constituent des propriétés internes. D'autres sont des observables qui sont présentes chez des systèmes autonomes sans être la preuve d'une autonomie constitutive. Ce sont des propriétés externes.

Nous reprenons la définition AEV-2 pour en extraire les propriétés externes que nous considérons comme nécessaires chez un agent éactif virtuel :

▷ **La robustesse.**

“Nous appelons agent éactif virtuel un système dont le fonctionnement le définit

comme une unité.”

Un agent éactif virtuel résiste aux perturbations. Des changements mineurs dans l’environnement ne doivent pas entraîner l’effondrement du comportement d’un agent autonome. Un environnement complexe présente de nombreuses petites variations qui sont souvent négligeables dans la réalisation d’une tâche spécifique et par conséquent peuvent être considérées comme du bruit. Un agent autonome doit être capable d’adopter un comportement robuste aux bruits de l’environnement.

▷ **L’adaptativité.**

“Il est en couplage avec son environnement et régule de manière interactive son fonctionnement...”

L’adaptativité est la capacité à adapter son comportement de manière appropriée face à des changements dans son environnement. De plus, cette prise en compte des changements dans l’environnement doit se faire dans une échelle de temps adaptée à la perturbation. Un système adaptatif s’auto-régule, il est donc actif dans l’interaction avec l’environnement, ce qui lui permet de maintenir de manière active une norme, c’est-à-dire de poursuivre un but.

▷ **L’indéterminisme.**

“...pour suivre une ou plusieurs normes.”

Nous nous intéressons ici à la capacité d’un agent à maintenir plusieurs buts et à déterminer par lui-même quel but il adopte à un instant donné. Nous nommerons cette caractéristique l’indéterminisme des comportements dans la mesure où elle reflète l’impossibilité pour un observateur à prévoir totalement le comportement de l’agent autonome. Le comportement est à la fois déterminé par l’environnement qui perturbe le système et à la fois, il est déterminé par l’état interne du système.

▷ **La capacité d’évolution.**

“Un agent éactif virtuel maintient et génère ses propres normes/buts.”

L’adaptativité seule ne permet pas de rendre compte de la capacité d’évolution d’un agent éactif virtuel. L’ontogenèse, c’est-à-dire le développement d’un individu, dépend de ses capacités d’adaptation mais aussi d’apprentissage.

à partir des propriétés internes que nous avons établies ici, nous justifions dans le chapitre 2 le choix des CTRNN comme architecture de contrôle, ainsi que l’utilisation de l’approche évolutionnaire comme méthode de conception. Nous confronterons ensuite ce modèle aux propriétés externes d’un agent éactif virtuel.

Chapitre 2

Choix techniques pour l'autonomie constitutive

LA question que nous abordons dans ce chapitre est la suivante : Quels choix techniques peuvent refléter la dynamique d'un comportement autonome, à la fois sensible à l'environnement et capable d'initiatives propres ? Nous commençons dans la section 2.1 par justifier le choix des réseaux de neurones artificiels récurrents à partir des contraintes que nous identifions pour la conception d'agents autonomes. Nous détaillons ensuite dans la section 2.2 le fonctionnement des réseaux de neurones récurrents à temps continu (CTRNN) qui semblent les plus pertinents pour notre problème. L'objet de la section 2.3 est l'approche évolutionnaire qui est la technique utilisée dans les travaux existants pour paramétrer des réseaux de neurones récurrents utilisés comme contrôleurs d'agents artificiels.

2.1 Autonomie et architectures d'agent

2.1.1 Architectures classiques d'agents autonomes

Au chapitre précédent, nous avons relevé certaines propriétés internes caractéristiques d'un agent éactif virtuel. Celles-ci sont les suivantes :

- ▷ La clôture opérationnelle du système.
- ▷ La situation dans un environnement.
- ▷ L'ancrage sensorimoteur des représentations.
- ▷ La capacité d'autorégulation.
- ▷ La sensibilité aux conditions environnementales qui rend le comportement non prévisible.
- ▷ Le caractère émergent du comportement.

La question se pose alors de savoir si les architectures d'agents autonomes, telles qu'on les trouve habituellement, peuvent être utilisées ou étendues pour rendre comptes des contraintes

que nous venons de préciser. Nous serons synthétiques sur ces architectures en les regroupant selon la terminologie classique réactives, cognitives ou hybrides et en proposant, pour chacune de ces catégories, une analyse globale. Elles ont toutes en commun de ne pas avoir été imaginées dans la perspective de l'autonomie forte. Ainsi, elles reprennent toutes un schéma de la cognition comme traitement d'entrées pour la production de sorties.

L'approche réactive peut être représentée par les modèles de Brooks (1986) ou Braitenberg (1986) par exemple. Ils ont l'avantage de reposer sur la situation d'un agent dans un environnement qui lui sert directement de modèle. Cependant, l'architecture de subsomption de Brooks est basée sur une organisation linéaire des processus de production du comportement. Braitenberg propose lui une approche connexionniste dans laquelle des boucles sont possibles permettant une mémoire. L'organisation réursive que nous donnions parmi nos contraintes apparaît dans ce modèle. Sans exclure totalement ce type d'architectures, elles semblent néanmoins limitées pour rendre compte de caractéristiques telles que l'ancrage sensorimoteur de "connaissances" qui résulteraient de l'histoire des interactions. En d'autres termes, le système semble trop simple pour accumuler une connaissance liée à ses interactions et il nous faut aller vers plus de complexité pour aborder des notions telles que la clôture opérationnelle.

L'approche cognitive pour la conception d'agents est typiquement représentée par les architectures de type BDI ¹(Bratman et al., 1988). Elles sont généralement basées sur un système formel doté d'une axiomatique permettant un raisonnement sur des "états mentaux" comme les croyances ou les intentions. Elle permettent de planifier un comportement et d'optimiser des décisions en fonction d'un but et des évolutions potentielles de l'environnement "prévues" par le modèle. Il va sans dire que ce type d'architecture est clairement d'inspiration computationnaliste (entrée/traitement/sortie) et il semble contradictoire de les retenir pour atteindre notre objectif.

L'approche hybride combine les deux précédentes approches. L'architecture d'un agent est alors composée de modules, chaque module étant soit une composante cognitive avec représentation symbolique des connaissances et capacités de raisonnement, soit une composante réactive. InteRRaP² (Müller et Pischel, 1994) propose une décomposition conceptuelle en couches verticales. L'architecture à sélection d'action de Maes (1989) et celle de Simonin et Ferber (2000) sont également des architectures hybrides, mais qui proposent une décomposition fonctionnelle. L'architecture PerAc (Gaussier et Zrehen, 1995) utilise deux niveaux. Un premier niveau réflexe engendre des propositions d'action à partir de la situation perçue. Un second niveau reconnaît des situations précises et renforce ou inhibe les propositions d'action. De plus, l'algorithme PCR (Revel et Gaussier, 2003) permet à un agent d'émettre des "hypothèses" en adaptant ses règles de renforcement.

Les approches hybrides ont l'intérêt de préserver l'aspect situé de la production du comportement tout en conservant la notion de mémoire ou dynamique interne, permettant ainsi une régulation du comportement. Cependant, toutes ces architectures supposent des hypothèses fortes sur le fonctionnement du système, qui est décomposé soit en modules concep-

¹ Pour *Belief Desire Intention* soit croyances, désirs, engagements

² Pour *Integration of Reactive Behavior and Rational Planning* soit intégration du comportement réactif et planification rationnelle

tuels (planification, communication,...), soit en modules fonctionnels (évitement d'obstacle, phototaxie,...). Là encore, elles sont pensées dans l'optique où le comportement consiste en un traitement d'entrées pour produire des sorties et non pas dans celle où le comportement serait le résultat d'une réorganisation pour se maintenir face à des perturbations. à priori, nous ne les retenons pas mais nous reviendrons sur elles lors de nos perspectives concernant la recherche d'un compromis entre complexité et maîtrise suite aux écueils que nous avons rencontrés. Néanmoins, notre choix s'oriente vers un positionnement plus radical, puisque c'est un des objectifs de cette thèse, mettant la circularité et la clôture à la base de l'architecture candidate à nos expérimentations.

2.1.2 Les modèles récurrents

Clôture, autorégulation, importance de l'histoire du couplage, tous ces attributs mettent en avant l'aspect dynamique et temporel du système qui est d'ailleurs également considéré comme majeur par des spécialistes de la cognition (Elman, 1990; Varela et al., 1993; Beer, 2000). Une première idée pour considérer le temps est de traiter des vecteurs d'évènements, c'est-à-dire de représenter l'ordre des événements en ajoutant une dimension au motif à traiter. Mais le temps a un rôle particulier : il n'est pas juste une dimension supplémentaire à considérer, il a plutôt une influence sur le traitement lui-même. Elman rejette donc les approches qui introduisent le temps comme une dimension supplémentaire. Pour lui, il faut donner au système des propriétés dynamiques de traitement, c'est-à-dire une mémoire.

Les réseaux de neurones artificiels ont été introduits par McCulloch et Pitts (1943). Ce sont des modèles mathématiques dans lesquels des unités de calcul sont inter-connectées pour réaliser des fonctions. Dans un réseau sans cycle, la communication ne se fait que d'une couche à la suivante de manière unidirectionnelle. Ce type de réseau est purement réactif, le système ne maintient pas de trace de son activité. Introduire de la récurrence dans le réseau permet d'obtenir une mémoire. L'état du système fournit alors un contexte. Un même stimulus ne produit plus nécessairement la même sortie en fonction de l'état interne du système. La récurrence des connexions ajoute la dimension temporelle au fonctionnement du réseau³. Alors que des réseaux de neurones artificiels acycliques ne peuvent implémenter que des fonctions, les réseaux de neurones récurrents peuvent modéliser des systèmes dynamiques et c'est pour cela qu'ils constituent un modèle particulièrement intéressant pour la modélisation de contrôleurs d'agents énaactifs virtuels.

Tani (1996) utilise des réseaux de neurones artificiels récurrents standards à temps continu (Jordan, 1986; Elman, 1990) comme contrôleurs d'agents capables d'apprendre des trajectoires. L'agent utilise un niveau de navigation pour produire une commande motrice aux embranchements. Cette activité motrice est associée à d'autres entrées telles que la valeur de capteurs de distance ou un indicateur de la distance parcourue pour nourrir un réseau de

³ Tous les réseaux récurrents n'utilisent cependant pas cette propriété dynamique. Le modèle de Hopfield (1982) par exemple permet de simuler une mémoire associative. En effet, l'apprentissage du réseau consiste à faire correspondre certaines configurations d'entrées à des attracteurs du réseau. Il n'y a cependant pas d'utilisation dynamique, un motif constant est fourni en entrée et le réseau converge vers un des états stables pour lesquels il a été entraîné.

neurones. L'architecture des réseaux utilisés est présentée sur la figure 2.1. La récurrence du réseau va permettre aux agents de prédire l'endroit qu'ils vont atteindre en fonction de leurs actions. Pour cela, les agents fonctionnent d'abord en mode ouvert jusqu'à ce qu'il soient "situés", c'est-à-dire que leurs prédictions soient suffisamment justes. Il est ensuite possible de passer en mode fermé pour qu'ils planifient une trajectoire afin d'atteindre un endroit précis. Dans ce mode, les agents simulent différents programmes moteurs et calculent une séquence optimisée. Ils repassent en mode ouvert pour suivre la trajectoire calculée.

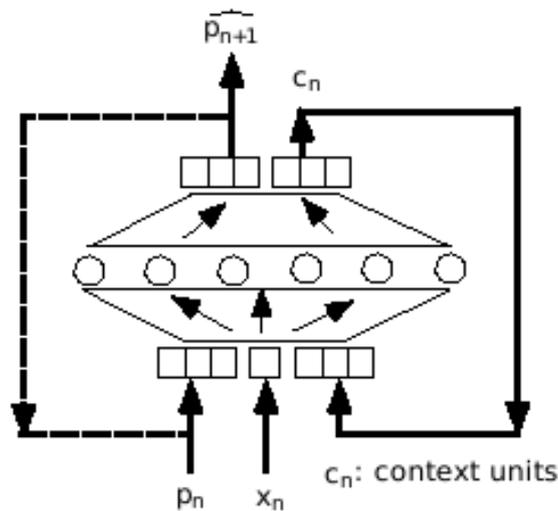


FIG. 2.1 – Architecture d'un réseau de neurones artificiels récurrent standard à temps discret. D'après Tani (1996). p_n est l'entrée sensorielle courante, x_n la commande moteur, c_n est le contexte qui est une copie de la sortie au pas de temps précédent.

La récurrence est un principe important qui peut amener vers l'autonomie constitutive. Cependant, pour retrouver d'autres propriétés de cette autonomie, telle que la capacité à compenser des perturbations par une dynamique propre ou encore la clôture opérationnelle du système, cette notion de récurrence peut être exploitée plus fortement. Pour obtenir un système dont la dynamique propre soit suffisamment complexe pour être à la fois autodéterminée, adoptant certaines habitudes, tout en étant également sensible à des perturbations qui engendrent des changements de comportements, il semble nécessaire de choisir des modèles plus complexes que ceux présentés jusqu'ici. Les CTRNN sont de bons candidats pour cela. Il a en effet été montré par Funahashi et Nakamura (1993) que ce type de réseau pouvait approcher tout système dynamique avec une précision arbitraire moyennant un nombre suffisant d'unités. Nous décrivons leur fonctionnement plus en détail dans la section suivante.

2.2 Réseaux de neurones récurrents à temps continu

Les réseaux de neurones récurrents à temps continu (CTRNN) sont un type de réseaux de neurones artificiels caractérisés par une topologie qui autorise toutes les connexions et par le fait que chaque neurone possède sa propre dynamique.

2.2.1 Topologie

Dans un CTRNN, toutes les connexions sont autorisées, y compris les auto-connexions. La figure 2.2 représente un réseau complètement récurrent. L'organisation de ce type de réseau est circulaire, il n'est pas possible de distinguer des couches successives de traitement de l'information. On retrouve ici la propriété de clôture opérationnelle qui définit les systèmes autonomes selon Varela (1989). Si un observateur distingue des entrées provenant de l'environnement et des sorties produites par le réseau, l'organisation circulaire ne permet pas d'expliquer les sorties uniquement en fonction des entrées. En effet, l'état interne du réseau détermine tout autant son évolution que les entrées provenant de l'environnement.

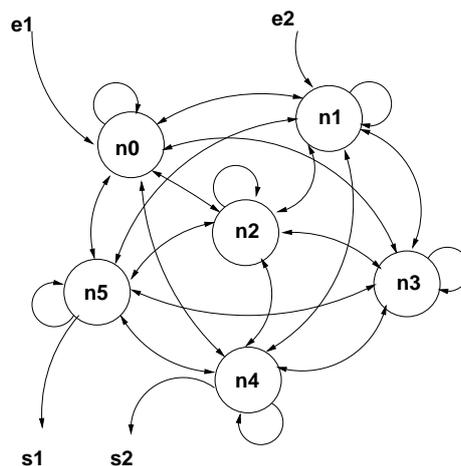


FIG. 2.2 – Exemple de réseau de neurones artificiels complètement récurrent. Certains neurones, ici n0 et n1, reçoivent une entrée extérieure au réseau. La sortie de certains neurones, ici n4 et n5, est considérée comme sortie du réseau. Les neurones sont tous inter-connectés et sont aussi auto-connectés. Les connexions entre deux neurones différents sont bi-directionnelles.

2.2.2 Dynamique des neurones

L'évolution de l'activation d'un neurone est gouvernée par l'équation différentielle suivante :

$$\tau_i \frac{\delta y_i}{\delta t} = -y_i + \sum_{j=1}^k w_{ji} \sigma(y_j + b_j) + I_i \quad (2.1)$$

dans laquelle y_i est l'activation du neurone i , b_j un biais, $\sigma(y_i + b_j)$ est la fonction d'activation de type sigmoïde, w_{ji} le poids de la connexion entre le neurone i et le neurone j , I_i est une entrée externe sur le neurone i et k est le nombre total de neurones. L'activité des neurones évolue donc de manière continue, les entrées du neurone ne déterminent pas de manière directe son activité mais participent seulement au calcul de la dérivée de cette activité. De plus, chaque neurone ayant sa propre constante de temps, ils évoluent selon une fréquence spécifique. Dans un même réseau, il peut donc y avoir des neurones qui réagissent de manière quasi réactive à leurs entrées alors que d'autres auront une grande inertie. La dynamique globale du réseau peut donc refléter des dynamiques opérant à différentes échelles de temps.

La figure 2.3 présente un exemple d'activité d'un CTRNN. Chaque courbe représente l'évolution de l'activation d'un neurone dans le temps. Nous observons que globalement l'activité de chaque neurone oscille à une fréquence qui lui est propre.

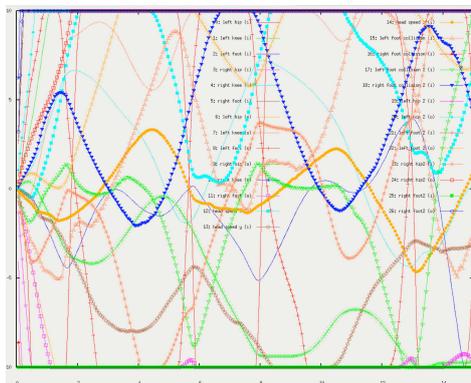


FIG. 2.3 – D'après De Loor et al. (2007). Activité d'un CTRNN. Chaque courbe représente l'évolution de l'activation d'un neurone dans le temps.

Ce type de réseaux de neurones artificiels est devenu un standard pour la modélisation de contrôleurs d'agents autonomes dans une approche dynamique. Comme nous l'avons dit précédemment ils présentent l'avantage théorique de pouvoir reproduire le comportement de n'importe quel système dynamique.

Ces réseaux ont été étudiés par Randall Beer. Il propose notamment dans (Beer, 1995) une étude systématique de la dynamique des CTRNN de 1 et 2 neurones et quelques résultats concernant la dynamique des réseaux contenant plus de neurones. Il montre ainsi par exemple que le poids des auto-connexions joue sur la stabilité des points d'équilibre du système constitué par un neurone. Si ce poids est inférieur à 4, le point d'équilibre est forcément stable,

alors qu'avec un poids supérieur à 4, le neurone peut rester actif en l'absence d'entrée externe. Il peut alors y avoir un changement qualitatif du comportement du système. Différents portraits de phase possibles pour des CTRNN de 1 et 2 neurones sont présentés en fonction de variations des paramètres, c'est-à-dire des poids et des biais. L'analyse de la dynamique des CTRNN permet également de donner des clés pour le paramétrage des réseaux en fonction du comportement global désiré. En particulier, son analyse montre que les biais peuvent être calculés de manière à éviter la saturation des neurones (Beer, 2006), c'est-à-dire à obtenir des réseaux fortement oscillants, qui vont pouvoir exhiber un comportement rythmique.

Cependant, ces analyses sont faites sur des CTRNN isolés, dans le sens où aucune entrée externe variable n'est prise en compte. Le fonctionnement de CTRNN de plus de deux neurones et qui reçoivent des entrées qui varient avec le temps échappe à ces analyses, même si certaines hypothèses peuvent être déduites de l'analyse des réseaux isolés à peu de neurones. Ceci rend impossible un paramétrage entièrement "à la main" d'un CTRNN utilisé comme contrôleur d'agent éactif. Il existe des algorithmes d'apprentissage pour les CTRNN, *cf.* Pearlmutter (1990). Ils ne sont cependant pas adaptés à nos besoins car ils supposent de connaître la séquence temporelle détaillée correspondant au comportement recherché. L'approche évolutionnaire fournit une heuristique qui permet de trouver des ensembles de paramètres qui sont des solutions pour la simulation de systèmes dynamiques particuliers. La Robotique Évolutionnaire (Pfeifer et Scheier, 1999; Harvey et al., 2005; Beer et Gallagher, 1992), combine l'approche évolutionnaire et l'utilisation de réseaux de neurones artificiels dynamiques pour la simulation d'agents ou la création de robots autonomes artificiels. Nous allons dans la section suivante rappeler les principes de cette approche et décrire les techniques usuellement utilisées pour la mettre en place.

2.3 Approche Évolutionnaire

L'approche évolutionnaire est une méthode de conception inspirée de l'évolution naturelle darwinienne. Elle permet de concevoir des systèmes complexes qui partagent certaines propriétés avec les systèmes vivants.

La théorie de la sélection naturelle considère que tout individu d'une population survit pendant un temps limité puis meurt, et que les individus les plus adaptés ont tendance à avoir plus de descendants. Ainsi, l'espèce évolue, s'adaptant aux conditions dominantes. La sélection naturelle suppose trois principes : hérédité, variation et sélection. Le principe de sélection introduit l'idée d'une direction guidant l'évolution, cependant le moteur de l'évolution n'est pas l'optimisation d'un critère, mais plutôt la conservation de solutions viables, conduisant à des populations adaptées à une niche écologique. Ainsi, la sélection naturelle n'amène pas spécifiquement des solutions optimales. Un trait qui est neutre du point de vue adaptatif a donc de grandes chances de se propager. Il peut être la trace d'un avantage adaptatif passé comme l'apparition spontanée d'un nouveau trait qui pourra se révéler positif dans des conditions futures encore inconnues, il pourra aussi n'être jamais réellement utile à l'espèce en terme d'adaptation ou même se révéler néfaste face à un changement dans l'environnement. C'est pourquoi, la phylogenèse, c'est-à-dire l'évolution d'une espèce pourrait être présentée en terme de dérive naturelle, plutôt qu'en terme d'optimisation (Varela et al., 1993).

En ce qui concerne l'évolution artificielle (Holland, 1975), une direction est explicitement donnée. Le concepteur sait ce qu'il attend de l'évolution. L'algorithme évolutionnaire consiste à appliquer une pression sélective sur une population pour obtenir des individus qui satisfont un ou plusieurs critères que le concepteur a fixé. La figure 2.4 illustre les différentes étapes d'un algorithme génétique. La première étape consiste à générer une population initiale de taille n de manière aléatoire, c'est-à-dire à fournir à n individus autant de génotypes construits à partir de séquences arbitraires de symboles identifiés auparavant. Le phénotype de chaque individu est alors évalué et obtient un score correspondant aux critères choisis. Ceci permet de sélectionner les individus les mieux "adaptés" pour constituer la génération suivante. La création du génotype de la nouvelle génération se fait en respectant les principes d'hérédité et de variation. Typiquement, le génotype d'un individu sera le résultat du croisement du génotype de deux parents qui subira quelques mutations, mais il peut être aussi une simple recopie avec possibilités de mutations. La nouvelle population obtenue est alors de nouveau évaluée. Le cycle d'évaluation-reproduction continue jusqu'à ce qu'un niveau prédéfini d'adaptation soit atteint ou bien que le nombre de générations évaluées soit trop grand.

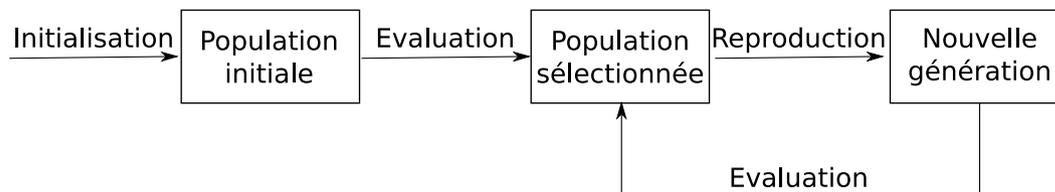


FIG. 2.4 – Principe d'un algorithme évolutionnaire.

L'évolution artificielle peut être utilisée pour trouver des solutions "originales" souvent très robustes. En effet, le concepteur pose le problème et l'algorithme propose une solution, qui n'est pas basée sur des hypothèses qu'auraient pu poser le concepteur et qui est potentiellement plus adaptée.

Doncieux et al. (2009) relèvent pour la robotique évolutionnaire, prise dans un sens large d'approche évolutionnaire appliquée à des contrôleurs de robots, les principaux problèmes suivants :

- ▷ Le passage de la simulation, dans laquelle les contrôleurs sont évalués, à l'implantation sur un robot peut entraîner des chutes de performances en raison de l'imperfection de la simulation à reproduire l'environnement réel. Il est préférable de réaliser l'évolution sur les robots directement, cependant le temps nécessaire sera plus important. Nous n'approfondissons pas ce point car les agents que nous voulons obtenir sont destinés à peupler des environnements virtuels.
- ▷ La robotique évolutionnaire peut-elle être appliquée pour un apprentissage automatisé "en-ligne", dans lequel l'algorithme évolutionnaire est "embarqué" chez un agent qui modifie en ligne ses paramètres. Ce problème n'est pas discuté ici puisqu'il s'agit d'un type d'apprentissage qui va à l'encontre des principes d'autonomie que nous cherchons à respecter dans la construction d'agents artificiels.

- ▷ Les conditions dans lesquelles les contrôleurs sont évalués peuvent influencer la généralité des solutions, la robustesse du comportement n'étant alors pas assurée dans des conditions différentes. Pour résoudre ce problème, il faut s'attacher à évaluer les contrôleurs dans des conditions représentatives de l'environnement avec lequel ils seront amenés à interagir. Ce point est particulièrement délicat dans le cas d'environnements ouverts et donc potentiellement imprévisibles. Pour améliorer la généralité des solutions, introduire du bruit sur les capteurs est une première piste (Di Paolo et Harvey, 2003). De même, utiliser des conditions aléatoires renforce la robustesse du comportement évalué.
- ▷ La robotique évolutionnaire a été appliquée avec succès à des tâches simples, le passage à des comportements plus complexes n'est toutefois pas évident. Cette question est celle de l'efficacité des algorithmes évolutionnaires pour des fonctions d'évaluation qui présentent une allure complexe. En effet, pour certaines tâches, une grande partie de l'espace de recherche correspond à des "plateaux" à partir desquels il est difficile d'obtenir des solutions plus adaptées. Ceci entraîne le "bootstrap problem" qui se pose si tous les individus de la première génération sont tous autant mal adaptés pour que le processus d'évolution puisse engendrer une meilleure solution à partir de cette population initiale. Selon la fonction d'évaluation utilisée, la recherche peut également être sujette à une convergence prématurée sur un optimum local ou sur un "plateau".

Pour permettre une exploration plus efficace de l'espace des solutions, plusieurs méthodes ont été proposées. Les principales sont les suivantes :

- ▷ L'évolution incrémentale consiste à faire évoluer des comportements graduellement de plus en plus complexes (Harvey et al., 1994). L'évolution commence alors avec une tâche simple. Une fois que des scores satisfaisants sont obtenus, la population "évoluée" est alors utilisée comme population initiale pour une tâche plus complexe. Comme nous le verrons dans la section 3.2.1, Tuci et al. (2002) par exemple commencent par faire évoluer un comportement phototaxique chez des agents. à partir de la population obtenue, ils cherchent ensuite à obtenir des agents qui apprennent une association entre la position de la lumière et un but à atteindre. Izquierdo et al. (2008) procèdent aussi de manière incrémentale en augmentant progressivement le nombre de tests effectués à la suite d'une phase d'association entre la nourriture et une température. Ils s'assurent ainsi que l'association est mémorisée de manière durable. La tâche est ensuite de nouveau complexifiée, puisque les agents sont évalués sur leur capacité à mémoriser l'association pendant plusieurs essais mais aussi à modifier la température associée à la nourriture.
- ▷ L'évolution conjointe de la topologie d'un réseau de neurones utilisé comme contrôleur et de ses paramètres (Stanley et Miikkulainen, 2002) permet d'ajuster l'espace de recherche si nécessaire en partant d'une topologie minimale. Lorsque l'on utilise un réseau de neurones artificiels comme contrôleur pour un agent, classiquement le concepteur définit un nombre de neurones et une topologie. Les paramètres qui sont recherchés par l'algorithme génétique sont alors les poids sur les connexions. Il n'y a cependant pas de méthode évidente pour déterminer le nombre de neurones nécessaires et quelles connexions utiliser. Il a été alors proposé de faire évoluer la

topologie, c'est-à-dire la structure du réseau, conjointement aux poids.

- ▷ L'approche multi-objectif (Schaffer et Grefenstette, 1985; Deb, 1999) définit des fonctions d'évaluation indépendantes correspondant à différents objectifs. L'exploration se fait alors par la recherche d'un compromis entre les objectifs.
- ▷ La co-évolution (Sims, 1994; Nolfi et Floreano, 1998) consiste à évaluer les individus en les mettant en situation de compétition, ce qui permet d'avoir une évaluation plus dynamique qui s'affine à mesure que la population est mieux adaptée.
- ▷ La préservation de la diversité (Mouret et Doncieux, 2009; Spector et Klein, 2005) permet à de nouvelles solutions de faire leurs preuves et peut permettre d'éviter la convergence prématurée.
- ▷ L'homme peut également intervenir en tirant parti de l'évolution.

Nous présentons dans la suite différents algorithmes qui mettent en œuvre certaines de ces solutions. Dans la section 2.3.1, nous présentons l'algorithme NEAT qui permet de faire évoluer à la fois la topologie du réseau et les poids des connexions. L'algorithme multi-objectif NSGA-II présenté dans la section 2.3.2 a été élaboré avec l'intention d'optimiser la recherche de solutions satisfaisant plusieurs critères. L'algorithme Microbial que nous présentons ensuite dans la section 2.3.3 est lui l'expression d'une volonté de simplicité. Pour Harvey (2001), la convergence prématurée est un faux problème. En effet, il démontre l'existence de "passages" qui permettent de sortir des optima locaux, même à partir d'une population génétiquement proche. L'évolution se fait alors par une dérive, engendrée par des mutations neutres pour l'adaptation mais qui peuvent permettre l'apparition d'un trait nouveau et utile lorsqu'une nouvelle mutation apparaît.

2.3.1 Évolution de la topologie d'un réseau de neurones : algorithme NEAT

Gruau et al. (1996) réussirent à trouver des solutions au problème d'équilibrage de pôles⁴, en faisant évoluer la structure d'un contrôleur en même temps que les poids. Cependant, la résolution de ce problème était alors plus lente qu'avec d'autres méthodes d'évolution utilisant une topologie fixe. Stanley et Miikkulainen (2002) proposent l'algorithme NEAT⁵, et arrivent à la conclusion inverse : l'évolution de la topologie conjointement aux poids permet de trouver des solutions plus rapidement pour ce même problème.

L'encodage génétique utilisé facilite le croisement. Le génome est composé d'une liste de nœuds et d'une liste de connexions. Parmi les caractéristiques d'une connexion, on trouve un

⁴ Le *pole balancing problem* est un problème de référence dans les domaines de la théorie du contrôle et des réseaux de neurones artificiels pour la conception et le test de contrôleurs de systèmes dynamiques linéaires complexes et instables. Voir <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.9205&rep=rep1&type=pdf>

⁵ Pour *NeuroEvolution of Augmenting Topologies*

identifiant d'innovation qui permet de tracer l'histoire de l'évolution du réseau. Des mutations sont possibles sur les poids des connexions mais également sur la structure. Les mutations structurelles peuvent intervenir soit par l'ajout d'une connexion entre deux nœuds existants, soit par l'ajout d'un nœud qui vient s'interposer sur une connexion, celle-ci étant détruite pour créer deux nouvelles connexions dont le poids de l'une est le même que l'ancienne et le poids de l'autre est mis à 1, ce qui réduit la perturbation engendrée par la mutation.

Grâce aux identifiants d'innovation, le croisement peut se faire de manière cohérente par rapport à l'histoire des mutations. Le système connaît en effet les gènes *correspondants*, c'est-à-dire qui sont apparus simultanément, les gènes *disjoints* (qui correspondent à des évolutions parallèles) ainsi que les gènes *en excès* (qui correspondent à des mutations intervenant après la dernière mutation de l'autre individu). Ce maintien de l'information sur l'évolution historique du génome permet de plus de préserver l'innovation en créant des niches. En effet, une nouvelle structure a peu de chance d'être efficace immédiatement et il est souhaitable de lui laisser un peu de temps pour évoluer. Une distance de compatibilité peut être calculée entre deux génomes en utilisant le nombre de gènes disjoints, de gènes en excès et la différence sur les poids des connexions. Une liste d'espèces est maintenue, puis à chaque génération, les génomes sont placés dans les espèces par comparaison avec un individu représentant de chaque espèce en utilisant un seuil de distance de compatibilité. Un dernier point important dans la description de l'algorithme NEAT est que la population initiale utilisée correspond à des solutions minimales, dans lesquelles les entrées sont directement liées aux sorties (il n'y a pas d'unités cachées).

Par une série d'expériences, Stanley et Miikkulainen (2002) montrent l'efficacité de l'algorithme dans sa globalité. En étudiant les performances de l'algorithme amputé de certains composants (mutations structurelle, espèces, population minimale et croisement), ils montrent également que chaque composant est nécessaire au bon fonctionnement de l'algorithme. L'algorithme NEAT n'a pas été utilisé à notre connaissance pour faire évoluer des CTRNN pour lesquels il faut non seulement déterminer des poids sur les connexions, mais également un biais et une constante de temps propre à chaque neurone.

2.3.2 L'approche multi-objectif : algorithme NSGA-II

Deb et al. (2002) proposent un algorithme génétique adapté à des problèmes multi-objectifs. Dans ce type de problèmes, il existe souvent non pas une mais un ensemble de solutions optimales. Elles sont appelées les solutions Pareto-optimales ou encore solutions non dominées et aucune de ces solutions n'est meilleure que les autres sur l'ensemble des objectifs. La figure 2.5 illustre pour un problème à deux objectifs la notion de solutions Pareto-optimales de premier rang et de second rang.

L'algorithme NSGA⁶ est basé sur une procédure de tri par non domination, i.e. les individus dominés sont désavantagés, et sur l'utilisation de niches qui permet de maintenir une population équilibrée par rapport aux différents objectifs (Srinivas et Deb, 1995). Le tri

⁶ Pour *Nondominated Sorting Genetic Algorithm*

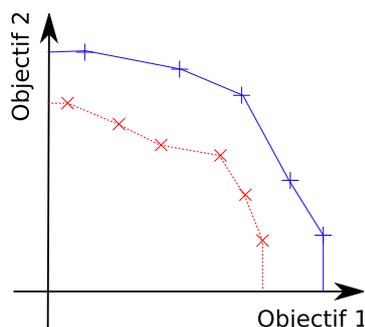


FIG. 2.5 – Solutions Pareto-optimales pour un problème à deux objectifs pour lequel on cherche à maximiser les scores. Chaque solution a un score par objectif. Les solutions marquées par un + appartiennent au front de Pareto, elle ne sont dominées par aucune autre solution. Les solutions marquées par un × appartiennent au deuxième front de Pareto, elles ne sont dominées que par des solutions du premier front de Pareto.

s'effectue en identifiant les individus non dominés dont l'ensemble constitue le premier front. Ces individus se voient attribuer un même score d'adaptation. Puis le mécanisme de niche est utilisé pour maintenir une diversité des solutions, en divisant le score d'un individu par le nombre d'individus qui lui sont proches. Pour cela, la distance entre tous les individus est calculée. Deux individus sont considérés comme proches si cette distance est inférieure à un paramètre. Les individus de ce premier front sont ensuite ignorés et la même procédure est utilisée pour construire le second front d'individus non dominés, auxquels on attribue un score inférieur à ceux du premier front. Si nécessaire, ce procédé est répété jusqu'à ce que la population entière ait été classée dans différents fronts.

Pour construire une nouvelle population, l'algorithme NSGA calcule l'ensemble des fronts. Cependant, ceci n'est pas indispensable et a été optimisé dans NSGA-II. En effet, ce calcul peut s'arrêter une fois que les fronts déjà calculés contiennent suffisamment d'individus pour créer la nouvelle population, *cf.* figure 2.6(b). De plus, pour préserver une diversité dans les solutions, NSGA-II n'utilise pas de paramètre prédéfini mais trie les individus en fonction d'une distance qui correspond à une densité autour d'eux. Les individus sont d'abord triés selon leur score pour un objectif et la densité autour de chaque individu est calculée en utilisant la distance aux deux voisins. La distance globale est la somme des densités pour chaque objectif, *cf.* figure 2.6(a). Toutefois, il n'est nécessaire de trier les individus selon cette distance seulement si l'ensemble du front ne peut pas être sélectionné. En effet, pour une population de N individus, l'ensemble des solutions du front dominant est sélectionné tant que le nombre d'individus déjà sélectionnés est inférieur à N .

Deb et al. (2002) montrent l'efficacité de l'algorithme sur des problèmes multi-objectifs en le comparant à d'autres algorithmes évolutionnaires multi-objectifs. NSGA-II est une procédure de tri de solutions non dominées rapide qui propose une stratégie élitiste et qu'il n'est pas nécessaire de paramétrer selon le problème. Nous n'avons pas connaissance de travaux utilisant l'algorithme NSGA-II pour l'évolution de CTRNN.

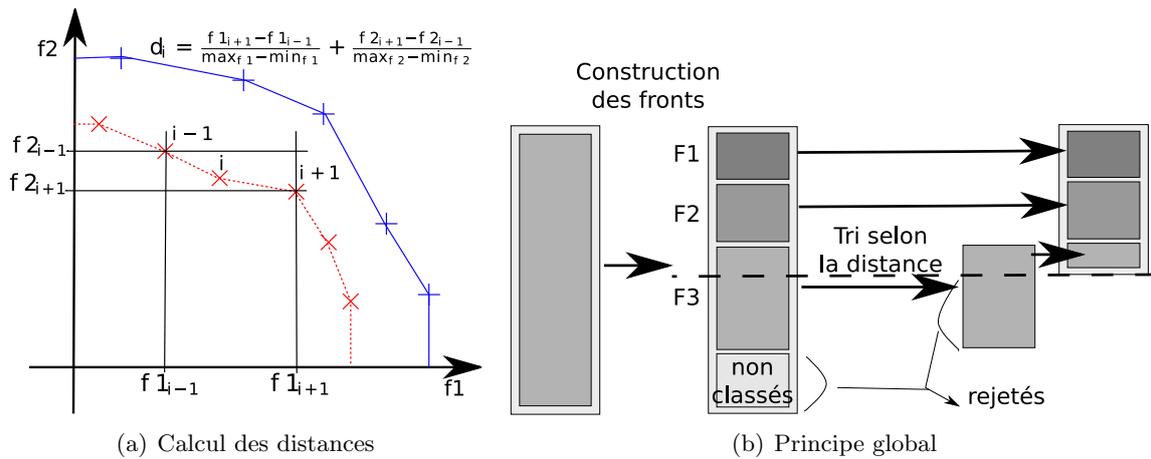


FIG. 2.6 – (a) Calcul de la distance d_i sur un problème à 2 objectifs. f_{j_i} correspond au score de l'individu i pour l'objectif j . max_{f_j} et min_{f_j} sont les valeurs maximales et minimales des scores pour l'objectif j . (b) Principe de l'algorithme NSGA-II. Les individus sont classés dans des fronts, puis triés selon leur distance si nécessaire pour construire une nouvelle population.

2.3.3 Algorithme Microbial

Dans une approche qui vise à proposer un algorithme génétique simplifié au maximum, Harvey (2009) propose Microbial qu'il encourage à utiliser aussi bien dans un but pédagogique qu'à des fins pratiques.

Les algorithmes génétiques sont traditionnellement basés sur des mécanismes générationnels, où les enfants viennent remplacer de façon synchrone les parents. Il est cependant possible d'imaginer que le remplacement d'un individu par un nouvel individu arrive de manière événementielle. La population n'évolue alors plus par paliers générationnels, mais de manière plus continue, certains individus plus "âgés" se retrouvant en concurrence avec de nouveaux individus. Harvey défend ce type d'algorithme génétique car il permet entre autres d'implémenter la sélection sous une forme minimaliste ou encore de paralléliser facilement l'évaluation des individus. Le nom de Microbial vient d'une analogie avec la reproduction bactérienne ou microbienne, dans laquelle le transfert de l'information génétique se fait de manière horizontale, par opposition à un transfert vertical de génération en génération.

La méthode de sélection utilisée traditionnellement consiste à appliquer une probabilité de sélection proportionnelle à l'adaptation, appelée sélection par roulette. Harvey critique cette méthode qui nécessite souvent une transformation pour que la sélection soit plus homogène, par exemple la probabilité de sélection d'un individu peut être proportionnelle à son rang dans la population plutôt qu'à son score. Une sélection par tournoi remplace de manière avantageuse ce type de sélection. Pour des tournois à deux individus, la probabilité de sélection de chacun est en effet la même qu'avec une sélection proportionnelle au rang, alors que la complexité est moindre puisqu'il n'y a pas besoin de trier la population.

Une autre question soulevée par les algorithmes génétiques est de choisir les individus

qui sont éliminés. L'algorithme Microbial propose d'utiliser la sélection par tournoi pour déterminer quel individu est remplacé. Les "parents" sont alors deux individus choisis au hasard et celui qui perd le tournoi est remplacé par un nouvel individu. Le nouvel individu est obtenu soit en croisant les deux parents et en appliquant des mutations, soit en reproduisant un individu et en le faisant muter.

Harvey (2001) relativise l'importance de la recombinaison dans l'évolution de la population. En effet, même en utilisant une population initiale aléatoire, dont le génotype est à la base très varié, le matériel génétique converge rapidement. Pour Harvey, ce sont donc les mutations qui sont le moteur de l'évolution. Il s'intéresse aussi d'ailleurs aux mutations "neutres", qui sans affecter directement le score d'un individu se propagent dans la population et permettent combinées à des mutations ultérieures de franchir des paliers d'adaptation. L'algorithme Microbial utilise néanmoins la recombinaison. Ainsi, le perdant du tournoi est combiné au gagnant avec une probabilité de remplacement de chacun de ses gènes, puis muté.

Les quelques lignes suivantes (Harvey, 2009), opèrent la sélection par rang, la recombinaison et la mutation, l'élitisme ainsi qu'une distribution géographique qui préserve une variabilité dans la population. Les individus qui participent à un tournoi appartiennent au même *dème*. Un dème correspond à une sous-population, déterminée par un voisinage géographique. Ainsi, les individus ne peuvent être en compétition qu'avec leurs voisins selon leur position dans la population.

```
void microbial_tournament(void) {
    int A,B,W,L,i;
    A=P*rnd(); // A est choisi au hasard dans la population P
    B=(A+1+D*rnd())%P; // B est choisi au hasard dans le deme D
    if (eval(A)>eval(B)) {W=A; L=B;} // Un tournoi donne le gagnant W et
    else {W=B; L=A;} // le perdant L
    for (i=0;i<N;i++) { // Pour chacun des N genes
        if (rnd(<REC) // le taux de recombinaison est REC
            gene[L][i]=gene[W][i]; // elle copie chez L le gene de W
        if (rnd(<MUT) // le taux de mutation est MUT
            gene[L][i]^1; // elle modifie le gene de L
    }}
```

Harvey (2001) illustre par un exemple l'efficacité de l'algorithme Microbial sur un problème pour lequel l'espace de recherche comporte de très nombreuses dimensions. Ainsi, malgré la convergence rapide du matériel génétique, des mutations neutres entraînent une évolution "en plateaux" dans laquelle l'adaptation augmente subitement parfois. Harvey (2009) ne prétend pas que Microbial soit plus efficace qu'un algorithme génétique standard, mais qu'il apporte les mêmes fonctionnalités dans une version très simple. De plus, il a été utilisé avec succès pendant des années à l'université de Sussex pour le paramétrage de réseaux de neurones dont des CTRNN.

2.4 Synthèse

Dans ce chapitre, nous nous sommes intéressés aux outils techniques candidats à la simulation d’agents éactifs. Les architectures classiques ne satisfaisant pas les contraintes liées à l’autonomie constitutive, nous nous sommes tournés vers les réseaux de neurones artificiels récurrents qui permettent de simuler une dynamique en préservant la situation de l’agent.

En particulier, les CTRNN ont des propriétés topologiques et dynamiques qui permettent d’envisager la mise en place de comportements relevant d’une dynamique propre qui peut être perturbée et évoluer. Ces contrôleurs peuvent être directement couplés à l’environnement par des capteurs et des effecteurs. Leur organisation n’est toutefois pas dépendante d’une vision entrées/sorties et la dynamique interne de tels réseaux peut amener la conservation d’une “identité” soumise à des perturbations. Ils sont capables d’approximer tout système dynamique, dont les systèmes chaotiques sont une sous-classe. Il est alors envisageable qu’un CTRNN reproduise une dynamique robuste, mais également “surprenante”, suscitant un sentiment d’engagement dans l’interaction. Ils constituent donc de bons candidats pour l’objectif que nous nous sommes donnés, qui est de nous approcher d’une autonomie constitutive.

Le paramétrage “à la main” de ce type de systèmes est cependant aujourd’hui impossible à faire. En effet, l’analyse de ce type de systèmes dynamiques est encore actuellement un défi mathématique. Pour l’instant l’analyse du comportement de CTRNN n’a pu être faite que pour un nombre très petit de neurones et sans tenir compte d’entrées variant dans le temps.

C’est pourquoi nous avons présenté l’approche évolutionnaire qui permet de rechercher des paramètres adaptés pour des contrôleurs de robots ou d’agents artificiels simulés. Différentes méthodes ont été décrites qui permettent d’améliorer cette recherche. Nous avons présenté trois algorithmes génétiques qui mettent en œuvre certaines de ces méthodes.

Faire évoluer des agents contrôlés par des réseaux de neurones récurrents dans des environnements ouverts ne se ramène pas à un problème d’optimisation mais à un problème pour lequel, dans un espace de recherche très vaste, la spécification de la tâche peut être amenée à évoluer, par exemple, lorsqu’une approche incrémentale est choisie.

Dans la pratique, il semble que l’algorithme Microbial se soit révélé adapté à l’évolution de CTRNN pour le type de problèmes qui nous intéresse, il est en effet utilisé depuis plus de 10 ans par certains chercheurs du CCNR⁷ dans le domaine de la robotique évolutionnaire. Nous n’avons pas connaissance d’application des autres techniques citées dans ce chapitre pour l’évolution de CTRNN contrôlant des agents artificiels. Ce type de réseaux présente une dynamique beaucoup plus complexe que celle des réseaux de neurones utilisés classiquement pour valider les performances d’algorithmes évolutionnaires. Leur efficacité dans le cadre qui nous intéresse n’a donc pas été démontrée.

En conclusion de ce chapitre, nous choisissons d’utiliser pour la simulation d’agents

⁷ *Center for Computational Neuroscience and Robotics, Université du Sussex*

éactifs virtuels des CTRNN dont le paramétrage se fait par des méthodes évolutionnaires du type de l'algorithme Microbial. Dans le chapitre suivant, nous argumentons ce choix en confrontant le modèle du CTRNN aux propriétés externes que nous donnions en synthèse du chapitre 1 comme nécessaires chez un agent éactif virtuel.

Chapitre 3

Argumentation pour la génération d'agents éactifs avec des CTRNN

EN synthèse du chapitre 1, nous considérons comme nécessaires chez un agent éactif virtuel les propriétés externes suivantes : robustesse, adaptativité, indéterminisme des comportements et capacité d'évolution. Nous présentons dans la section 3.1 des travaux qui montrent que des agents contrôlés par des CTRNN peuvent posséder les propriétés de robustesse, d'adaptativité et d'indéterminisme des comportements. Dans la section 3.2, nous discutons des travaux s'intéressant à l'apprentissage et argumentons sur le fait que la capacité d'évolution des comportements constitue un point dur pour la synthèse d'agents éactifs virtuels contrôlés par des CTRNN.

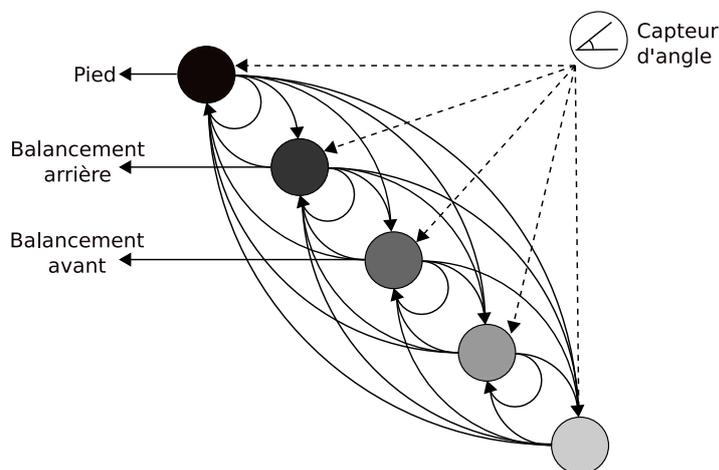
3.1 Exemples de travaux pour l'étude de l'autonomie

Les travaux que nous présentons dans cette section adoptent la démarche qui consiste à proposer un mécanisme pour la réalisation de certaines propriétés et à le tester à travers des simulations qui permettent d'interpréter le comportement produit sur la base de la connaissance des mécanismes en jeu mais aussi sur la base de ce qui est observé.

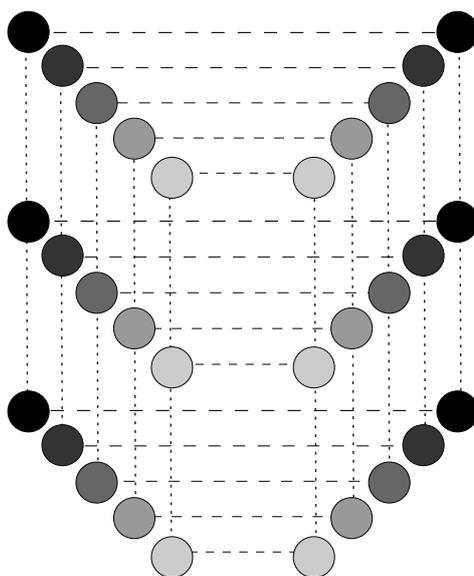
3.1.1 Robustesse des comportements

Il est important de s'assurer qu'un agent contrôlé par un CTRNN puisse adopter des comportements robustes car comme nous le présentions au chapitre 1, les systèmes dynamiques présentent des caractéristiques chaotiques, certes que nous recherchons mais pas au point que le comportement soit complètement aléatoire. L'expérience de Beer et Gallagher (1992) met en avant la capacité d'un agent contrôlé par un CTRNN à produire

un comportement robuste dans un environnement bruité. Il s'agit de concevoir un animat hexapode capable de locomotion en utilisant une approche dynamique. Les auteurs modélisent le contrôleur d'une patte comme un système dynamique et, pour simplifier le problème, font l'hypothèse qu'il y a des symétries avant-arrière et droite-gauche. La figure 3.1 décrit



(a) Contrôleur d'une patte



(b) Couplage entre les contrôleurs de pattes

FIG. 3.1 – D'après Beer et Gallagher (1992). (a) Architecture du contrôleur d'une patte d'animat. Chaque neurone reçoit une entrée provenant d'un capteur d'angle. Les sorties de trois neurones sont utilisées comme commandes de moteurs permettant le déplacement de la patte. (b) Couplage entre les contrôleurs des différentes pattes. Les neurones de même niveau de gris ont les mêmes seuils et constantes de temps. Les lignes interrompues représentent les connexions latérales. Les lignes pointillées représentent les connexions avant-arrière. Les connexions de même style entre neurones de même niveau de gris ont des poids identiques. Le couplage se fait donc en respectant des symétries avant-arrière et droite-gauche.

l'architecture d'un réseau de neurones contrôlant une patte ainsi que la manière dont des contrôleurs de pattes sont couplés pour obtenir le contrôleur d'un hexapode. Les contrôleurs de chacune des pattes ont donc les mêmes paramètres. Ils sont modélisés par un CTRNN qui contrôle une force verticale sur la patte, qui permet de la lever et de l'abaisser et deux forces de rotation opposées, qui simulent deux muscles antagonistes. L'angle de la patte est utilisé comme entrée sensorielle par le contrôleur. Les paramètres sont déterminés par un algorithme génétique.

Dans un premier temps, les auteurs recherchent et obtiennent des contrôleurs de pattes indépendants qui produisent un mouvement rythmique. Ceux-ci sont très dépendants de l'information sensorielle sur l'angle de la jambe. Cependant, des contrôleurs qui n'ont aucune information sensorielle ont de moins bonnes performances. Pour obtenir des contrôleurs à la fois performants et robustes à d'éventuelles défaillances sensorielles, Beer et Gallagher (1992) les font évoluer en tenant compte de leurs performances avec un retour sensoriel mais aussi sans.

De la même manière, la robustesse du comportement d'un animat complet dépend des conditions dans lesquelles il a été *évolué*. Un autre résultat intéressant de cette expérience est que l'évolution incrémentale n'est pas forcément plus rapide. L'évolution incrémentale consiste à faire d'abord évoluer des contrôleurs de pattes indépendants, puis à faire évoluer le couplage entre ces contrôleurs. Les auteurs remarquent qu'en couplant ces contrôleurs de pattes obtenus indépendamment, les solutions trouvées n'étaient pas aussi bonnes qu'en faisant évoluer directement l'ensemble du contrôleur de l'hexapode. Réciproquement, en isolant un contrôleur de patte qui a été initialement évolué en couplage avec d'autres, on ne retrouve pas un mouvement rythmique aussi bon que dans la première version.

Ceci met en avant ce qui est à la fois un avantage et un inconvénient de l'approche dynamique. La mise en place d'une approche incrémentale pour la complexification des comportements est difficile, car le travail déjà effectué n'est pas capitalisable. Le comportement est le résultat de la dynamique globale du système couplée à celle de l'environnement. Nous retrouvons ici les problématiques liées à la question de l'identité du système. Identifier des sous-systèmes semble de peu d'utilité. Cependant, ce caractère émergent du comportement le rend plus robuste.

En effet, le système maintient par sa dynamique un état relativement indépendant du contexte immédiat, ce qui n'est pas le cas des approches purement fonctionnelles dans lesquelles l'histoire de l'activité du système n'est pas prise en compte. Dans l'approche dynamique, le système n'est pas entièrement soumis à l'environnement; il montre une certaine inertie dans son activité qui lui permet d'exhiber un comportement robuste dans un environnement complexe et bruyant.

3.1.2 Adaptativité

Une expérience de Di Paolo (2000) illustre comment doter des agents artificiels de la caractéristique d'adaptativité face à des changements brutaux de l'environnement. Un

comportement est adapté s’il participe à la viabilité continue du système. Un comportement est adaptatif s’il restaure la viabilité menacée d’un système. Di Paolo reprend l’analyse de Taylor (1962) qui rend compte des propriétés d’adaptation du système visuel en se basant sur la théorie de l’ultra-stabilité développée par Ashby (1960). Les expériences de Taylor portent sur l’adaptation à une inversion horizontale ou verticale du champ visuel. Les sujets s’adaptent à la perturbation au point par exemple de pouvoir skier en étant soumis à cette inversion au moyen de lunettes.

Avec la théorie de l’ultra-stabilité, Ashby (1960) propose une modélisation de l’adaptativité. Ainsi, un système ultra-stable possède ce que Ashby nomme des variables essentielles. Une boucle sensorimotrice permet au système d’interagir avec son environnement. à travers ces interactions, les variables essentielles sont susceptibles d’être déséquilibrées. Un système ultra-stable est capable dans ces conditions de se réorganiser jusqu’à retrouver un équilibre. Il modifie ses paramètres comportementaux pour assurer sa viabilité en restabilisant les variables essentielles. Pour cela, une boucle sensorimotrice fermée n’est pas suffisante, il est nécessaire d’introduire un autre type de rétroaction.

L’expérience consiste alors à simuler un robot contrôlé par un CTRNN. Le contrôleur est couplé à deux capteurs permettant de détecter la présence d’une cible, ce qui perturbe la dynamique du système. La figure 3.2 illustre un tel agent. Certains neurones contrôlent la vitesse de déplacement de deux roues symétriques permettant ainsi d’appliquer un mouvement au robot.

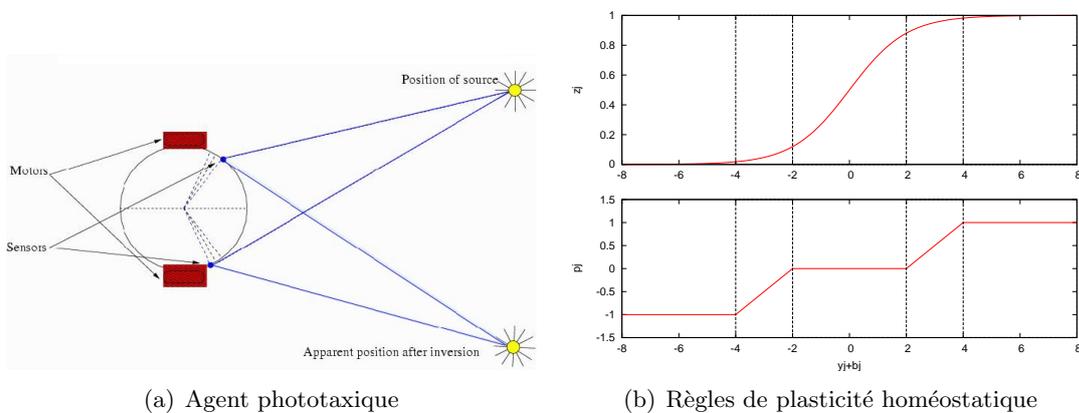


FIG. 3.2 – D’après Di Paolo (2000). (a) Un agent phototaxique a deux capteurs de lumière dont l’orientation peut varier et deux moteurs qui lui permettent de se déplacer. La figure indique la position réelle d’une source lumineuse et sa position apparente après inversion des capteurs de l’agent. (b) La plasticité s’applique (probabilité $p_j \neq 0$) lorsque l’activité du neurone sort d’un certain intervalle ($-2 < y_j + b_j < 2$). Cet intervalle est considéré comme délimitant la zone de viabilité du neurone. Si l’activité du neurone s’éloigne de cette zone de viabilité, le système se réorganise, en modifiant les poids des connexions, pour retrouver sa stabilité.

Les variables essentielles pour le robot sont liées à la stabilité de l’activité de chaque neurone. Di Paolo définit des intervalles homéostatiques pour ces activités. Si l’activité d’un neurone s’écarte de sa zone homéostatique, on considère que la variable essentielle qui lui est

associée est déstabilisée. Cela déclenche une règle de plasticité sur les connexions.

La plasticité homéostatique consiste à faire varier les poids des connexions quand cette stabilité est menacée, c'est-à-dire si les sorties de ces neurones dépassent certaines bornes inférieures ou supérieures. Les paramètres des règles de plasticité sont obtenus en utilisant une approche phylogénétique : On sélectionne et on améliore progressivement les robots qui atteignent des cibles posées au hasard grâce à un algorithme génétique. Cette méthode permet d'obtenir des robots qui exhibent un comportement phototaxique par le biais d'une dynamique sensorimotrice, mais surtout, et c'est remarquable, qui vont être capables à l'échelle ontogénétique de s'adapter à l'inversion de leurs capteurs en adoptant un comportement de compensation rappelant les expériences de Taylor.

3.1.3 Indéterminisme

Toujours en utilisant une plasticité homéostatique, Iizuka et Di Paolo (2007) proposent une expérience dans laquelle des changements de buts vont être co-déterminés par l'environnement et la dynamique interne d'un agent. L'agent artificiel doit choisir entre deux comportements qui consistent à approcher soit une lumière de type A (comportement A), soit une lumière de type B (comportement B). Les agents sont du même type que ceux décrits dans la section 3.1.2. Ils ont deux capteurs pour chacun des types de lampe et deux moteurs qui leur permettent de se déplacer dans l'environnement. Ils sont contrôlés par un réseau de neurones récurrent.

Deux régions homéostatiques A et B sont définies pour les variables internes, qui correspondent à deux intervalles d'activation des neurones, *cf.* figure 3.3. Si les dynamiques neuronales restent dans ces régions, la topologie du réseau ne bouge pas. Mais si les variables internes prennent des valeurs en dehors de ces intervalles, une règle de plasticité synaptique est utilisée localement de telle sorte que la dynamique du réseau retourne dans une des régions homéostatiques. Un processus d'évolution est utilisé pour obtenir des agents capables de montrer les deux types de comportement tout en maintenant leurs variables internes dans les régions homéostatiques correspondantes. Ainsi, quand un agent approche la lumière de type A, respectivement B, chacun des neurones du réseau maintient son activation dans la région homéostatique A, respectivement B.

Il existe donc deux régions bien distinctes dans lesquelles la dynamique neuronale est stable et correspond à deux comportements différents. Une préférence comportementale peut alors apparaître dans la mesure où, si la dynamique se stabilise dans une région, l'agent adopte le comportement correspondant de manière durable. Cependant, si des perturbations déstabilisent le système, la dynamique s'écarte temporairement de la zone stable et va évoluer de manière à retrouver une stabilité. Deux situations sont alors envisageables : soit la dynamique se restabilise sur le comportement en place avant la perturbation, soit la plasticité amène la dynamique neuronale dans la région de stabilité correspondant à l'autre comportement.

Les agents évolués sont mis en présence simultanément des deux types de lampes, posi-

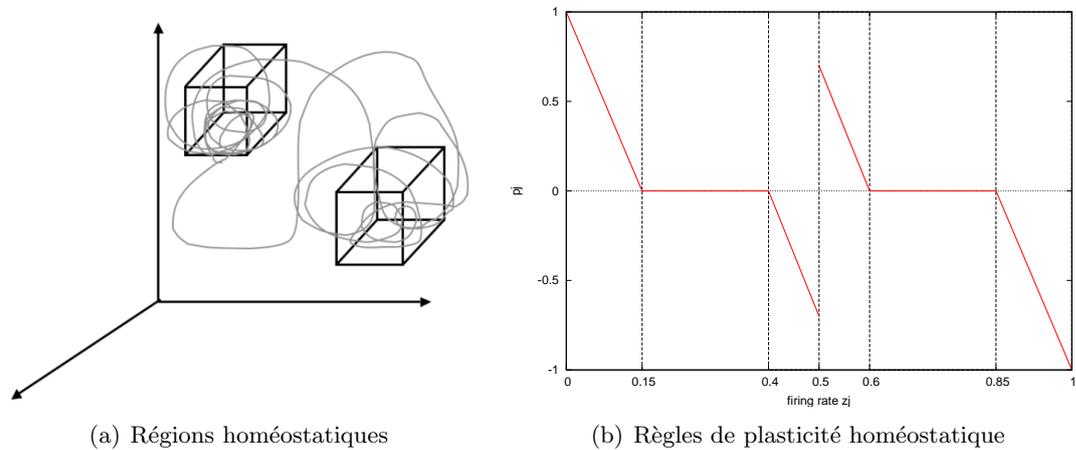


FIG. 3.3 – D’après Di Paolo et Iizuka (2008). (a) Représentation schématique des deux régions homéostatiques dans l’espace des activités des neurones. (b) Facilitation de la plasticité pour les neurones internes en fonction de l’activité du neurone. Deux zones stables, pour lesquelles la probabilité de plasticité est nulle, correspondent aux comportements A et B.

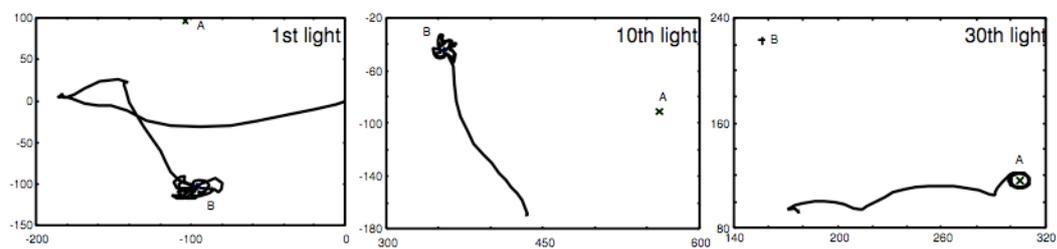


FIG. 3.4 – D’après Iizuka et Di Paolo (2007). Trajectoires spatiales d’un agent à la première présentation, à la dixième et à la trentième.

tionnées aléatoirement. Dans cette situation, l’agent adopte toujours un des comportements, c’est-à-dire que l’agent approche l’une des deux lampes mais il ne reste pas au milieu et ne s’éloigne pas des lampes non plus, *cf.* figure 3.4. Les auteurs observent également une préférence comportementale face à des perturbations externes. Ainsi, si les deux lampes sont déplacées, l’agent maintient le plus souvent une préférence pour la lampe qu’il approchait précédemment. Cependant, il arrive également que le déplacement des lampes entraîne un changement de comportement chez l’agent.

Nous venons d’illustrer la capacité des CTRNN à adopter trois propriétés externes d’agents éactifs virtuels, qui sont la robustesse des comportements, l’adaptativité et l’indéterminisme. Le principal problème que nous identifions dans les travaux que nous venons de présenter réside dans le fait qu’ils restent cantonnés à des tâches dites de bas niveau. Pour imaginer pouvoir traiter des problèmes dits de haut niveau, le système doit avoir la capacité de construire progressivement ses connaissances, d’élargir progressivement son répertoire de comportements. Ceci correspond à une capacité d’évolution du comportement des agents qui est la quatrième propriété externe que nous avons identifié pour des agents éactifs virtuels. Dans la section 3.2, nous discutons des travaux portant sur cette propriété chez des agents contrôlés par des CTRNN et argumentons sur le fait qu’elle constitue un point dur.

3.2 Faire évoluer un système autonome

La capacité d’évolution d’un agent éactif virtuel dépend de ses capacités d’adaptation mais également d’apprentissage. Un apprentissage à l’échelle phylogénétique, c’est-à-dire à l’échelle de l’espèce, peut entraîner une capacité d’évolution adaptative chez un individu, que nous nommerons apprentissage “réflexe”. L’apprentissage à l’échelle ontogénétique, celle de l’individu, est étroitement lié à l’histoire particulière de cet individu. Nous le désignons par le terme apprentissage “construit”. Dans la section 3.2.1, nous proposons une distinction entre ces deux notions à partir de travaux sur l’évolution d’agents contrôlés par des CTRNN. Cela nous amènera à constater qu’il n’existe pas de travaux démontrant que les CTRNN permettent de simuler des agents capables d’un apprentissage “construit”. Nous formulerons alors, dans la section 3.2.2, un objectif à long terme qui consiste à doter des entités virtuelles de la capacité à co-évoluer avec leur environnement.

3.2.1 Adaptation vs apprentissage, phylogénèse vs ontogénèse

Pendant un certain temps, les mécanismes de l’apprentissage ont été envisagés comme étant clairement séparés de ceux qui produisent le comportement. L’apprentissage animal a été en effet expliqué en terme de plasticité modifiant la topologie du système nerveux. Dans cette vision, la production du comportement était attribuée à l’activité neuronale et l’apprentissage était compris comme modulant cette activité. Ceci a été repris dans l’artificiel et a donné lieu à différentes techniques pour entraîner des systèmes artificiels établies comme des métaphores de la plasticité synaptique. Citons, par exemple pour des méthodes d’apprentissage supervisé, l’apprentissage par rétro-propagation du gradient de

l’erreur (Rumelhart et al., 1986) ou encore l’algorithme Cascade-correlation (Fahlman et Lebiere, 1991). Cependant, il est désormais acquis que la plasticité des connexions n’est pas le seul mécanisme de l’apprentissage, qui est désormais compris comme un processus dynamique intervenant à différentes échelles de temps et complètement intégré aux processus qui produisent le comportement (Izquierdo et al., 2008). L’approche dynamique se propose donc également de simuler l’apprentissage comme un processus pleinement intégré à la dynamique interne du système, c’est-à-dire sans utiliser de mécanismes dédiés à cet apprentissage, tels que des modules spécialisés dans l’apprentissage ou encore une plasticité supposée permettre l’apprentissage¹. La majorité de ces travaux se consacrent à la conception d’agents qui se comportent différemment dans un nombre discret d’environnements. Nous commençons par en donner un exemple.

3.2.1.1 Modèles écologiques de l’apprentissage

Dans le modèle initialement proposé par Yamauchi et Beer (1994a), les mécanismes à l’origine de l’apprentissage ne sont pas donnés par le concepteur. Ils sont le résultat de l’évolution des agents sous la pression d’une sélection artificielle, qui favoriserait les individus qui apprennent. Pour cela, les agents sont modélisés selon une approche dynamique en utilisant des CTRNN qui sont paramétrés grâce à un algorithme génétique. Le point important ici est que les concepteurs ne veulent faire appel à aucun mécanisme particulier dédié à l’apprentissage. Leur hypothèse est que l’évolution artificielle du réseau peut permettre d’obtenir un contrôleur qui intègre des comportements réactifs et des comportements adaptatifs qui vont engendrer un apprentissage.

Dans leur expérimentation, les agents peuvent se déplacer dans un environnement en une dimension qui contient un but et un repère. Ils sont dotés de capteurs de faible portée et ne peuvent donc capter qu’un élément à la fois. Les éléments qu’ils peuvent capter sont un repère et un but. Un moteur leur permet de se déplacer dans l’environnement. Les agents peuvent être confrontés à deux types d’environnements : le but est soit derrière le repère, soit de l’autre côté et comme il peut également être à gauche ou à droite, il y a quatre combinaisons potentielles de situation (voir figure 3.5). Ils sont évalués sur plusieurs essais. à chaque essai, ils sont placés au centre de l’environnement et le but est placé aléatoirement d’un côté ou de l’autre. La relation entre le but et le repère change de manière occasionnelle. La tâche des agents est donc d’apprendre cette relation, qui reste stable au cours d’une série d’essais pour atteindre le but en utilisant le repère comme indice.

Dans cette version, l’expérimentation n’aboutit pas à des résultats satisfaisants en raison, selon les auteurs, du temps nécessaire aux algorithmes génétiques pour trouver une solution. Ils proposent alors une version modulaire, dans laquelle trois réseaux dynamiques, qu’ils font évoluer séparément, simulent chacun une partie de la tâche. Un module permet d’atteindre le but quand il est proche du repère, un second module permet d’atteindre le but quand il est éloigné du repère. Un dernier module détermine le type d’environnement, c’est-à-dire la relation entre le repère et le but. La sortie de ce module active l’un des réseaux spécialisés

¹ Ceci est en accord avec le caractère émergent du comportement d’un système autonome de manière constitutive que nous relevons comme une des contraintes pour le développement d’un agent énaactif dans le chapitre 2.

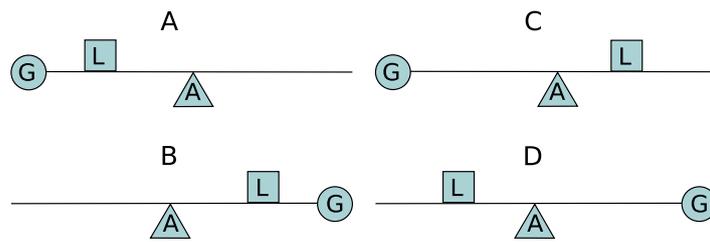


FIG. 3.5 – D’après Yamauchi et Beer (1994a). Le triangle représente l’agent A, le cercle représente le but G et le carré représente le repère L. Les configurations de l’environnement possibles peuvent être regroupées en deux catégories : soit le repère est proche du but (A et B), soit le repère est éloigné du but (C et D).

en fonction du type d’environnement. Avec cette approche modulaire, les auteurs obtiennent des agents qui ont le comportement désiré. Cependant, cette seconde version ne correspond plus aux objectifs fixés puisque le concepteur explicite les mécanismes à utiliser pour réaliser la tâche.

Tuci et al. (2002) reprennent cette expérimentation pour tenter d’analyser les raisons de l’échec de la première version. Ils identifient deux raisons possibles. Soit l’apprentissage associatif est un comportement trop complexe pour être obtenu en utilisant une modélisation d’aussi bas-niveau ; soit l’expérience contient des contraintes qui empêchent l’évolution d’aboutir au résultat recherché. Ils procèdent à une analyse écologique pour déterminer quel type de pression sélective permet, ou du moins favorise, l’apparition des mécanismes d’apprentissage. Dans leur expérience, les agents peuvent se déplacer dans un environnement fermé en deux dimensions. Ils sont équipés de capteurs infrarouges qui leur permettent de détecter des obstacles, de capteurs de lumière ambiante, d’un capteur de lumière orienté vers le sol qui leur permet de détecter le but et de deux moteurs pour se déplacer. En simulant des lésions sur les capteurs des agents évolués, ils montrent que si les capteurs d’obstacles ont bien un rôle dans le comportement moteur produit, il n’y a pas de différence comportementale significative entre un agent “intact” et un agent dont les capteurs de lumière sont altérés. Ceci suggère que l’information venant de la source lumineuse n’a aucun impact sur le comportement de l’agent. En effet, “faire attention” à la lumière n’a pas de valeur adaptative pour les agents tant qu’ils ne sont pas capables de faire l’association entre la lumière et le but. Cependant, les auteurs s’attendaient à ce que, par dérive génétique, des agents “sensibles” à la lumière finissent par faire l’association. Ce n’est pas le cas, toutes les lignées évoluées adoptent le comportement qui consiste à aller toujours du même côté. Pour une même valeur adaptative, i.e. un score semblable, ce comportement, étant moins complexe, est en effet plus simple à évoluer.

Les auteurs déduisent de cette analyse écologique, qu’il est nécessaire de tenir compte du rapport entre la complexité des mécanismes et leur valeur adaptative pour exercer les pressions sélectives adéquates. Dans le cas d’une association, il faut donc que le stimulus que l’on veut associer à une réponse ait une utilité adaptative même sans l’apprentissage de l’association. Par exemple, ici, si le but est plus souvent du côté de la lumière, les agents qui sont attirés par la lumière auront un avantage adaptatif sur ceux qui vont toujours tout droit, même si aucun d’entre eux n’est capable de faire l’association entre cette lumière et le but. Pour Tuci et al. (2002), la mise en place du comportement phototaxique est nécessaire

avant d’essayer de mettre en place l’apprentissage. Une analyse écologique doit permettre à l’évolution de mettre en place progressivement les mécanismes nécessaires à la réalisation d’une tâche.

En modifiant la valeur adaptative du comportement des agents face à la lumière (augmenter le score d’un comportement photosensible), il est alors possible de faire évoluer une population d’agents qui apprennent la relation entre lumière et but. Pour cela, un type d’environnement est privilégié. Précisément, ici, le but est plus souvent du même côté que la lumière. Ceci permet à l’évolution de procéder par étapes. En effet, sans apprentissage, des agents qui vont toujours vers la lumière sont plus adaptés au type d’environnement prédominant que des agents dont le comportement n’est pas du tout influencé par la lumière, puisqu’il n’y a plus une chance sur deux d’arriver au but sans tenir compte de la lumière. Une fois ce comportement stabilisé dans la population, l’évolution va amener l’apprentissage de la relation entre lumière et but. Les nouveaux individus sont alors adaptés aux deux types d’environnement. Notons ici que cette méthode d’évolution par étapes peut sembler en opposition avec les conclusions de Beer et Gallagher (1992) présentées dans la section 3.1.1, pour qui l’évolution incrémentale est moins efficace que la recherche d’une solution directe au problème global.

Le travail de Tuci et al. (2002) montre la possibilité de concevoir des agents capables d’apprendre des relations dans leur environnement sans utiliser de mécanisme dédié et alors que la production de leur comportement respecte les principes d’autonomie qui constituent le socle de l’approche dynamique. Ils peuvent être soumis à deux types d’environnement, dans lesquels leur comportement doit différer. L’apprentissage consiste donc à déterminer parmi un répertoire de deux comportements, lequel est adapté à l’environnement dans lequel ils se situent. Il s’agit donc de basculer d’un comportement à un autre. Izquierdo et al. (2008) s’intéressent à une autre forme d’apprentissage associatif, qui consiste à s’adapter à des environnements quelconques. Leurs travaux font l’objet de la section suivante.

3.2.1.2 Apprentissage conditionné

L’expérience de Izquierdo et al. (2008) porte sur l’apprentissage associatif sur un intervalle continu. Il s’agit de reproduire le comportement de vers qui apprennent à associer une température à la présence de nourriture. Ces vers sont en effet capables d’associer nourriture et température selon les environnements qu’ils ont expérimentés. S’ils sont affamés, ils éviteront la température qui n’est pas associée à la nourriture. Au contraire, si dans un environnement à une température particulière, ils ont trouvé de la nourriture, ils rechercheront cette même température.

Izquierdo et al. (2008) cherchent à reproduire ce comportement dans une version discrète et une version continue. Dans la version discrète, il n’y a que deux types d’environnements possibles, qui correspondent à deux températures différentes. Dans la version continue, la gamme d’environnements est plus étendue, puisque la température peut prendre n’importe quelle valeur dans un intervalle déterminé. Pour les auteurs, bien que les deux comportements soient proches, il ne s’agit pas du même type d’apprentissage puisque la version continue implique une réelle association et non plus un simple basculement entre différents modes

d'interaction. Tout comme dans l'expérience de Tuci et al. (2002) décrite à la section précédente, les agents sont modélisés selon une approche dynamique : un CTRNN reproduit un contrôleur couplé à des capteurs et à un moteur. Les capteurs permettent aux agents de sentir la température et la nourriture. Le moteur leur permet d'ouvrir ou de fermer leur bouche. Les paramètres sont évolués par algorithme génétique.

Dans cette expérience, un essai s'organise autour de deux phases qui peuvent se répéter : une phase d'association et une phase de test. Pendant la phase d'association, l'agent est placé dans un environnement à une température donnée en présence de nourriture. Après un délai, durant lequel il n'y a plus ni nourriture, ni température, la phase de test commence. L'agent est stimulé avec une température, qui peut être identique à la température initiale ou non. L'agent a réussi s'il *ouvre la bouche* quand la stimulation correspond à la température associée à la nourriture en début d'essai ou s'il *ferme la bouche* quand la température est différente. Il reçoit une récompense positive ou négative en fonction de son comportement. Plusieurs tests peuvent se succéder, séparés par des délais aléatoires. Enfin, l'agent doit également être capable d'apprendre une nouvelle association si l'environnement change. Un même essai peut donc comporter plusieurs séries de phases d'association et phases de tests. La figure 3.6 décrit la structure d'un essai individuel réussi qui ne comporte qu'une phase d'association et durant lequel un seul test est effectué pour l'association.

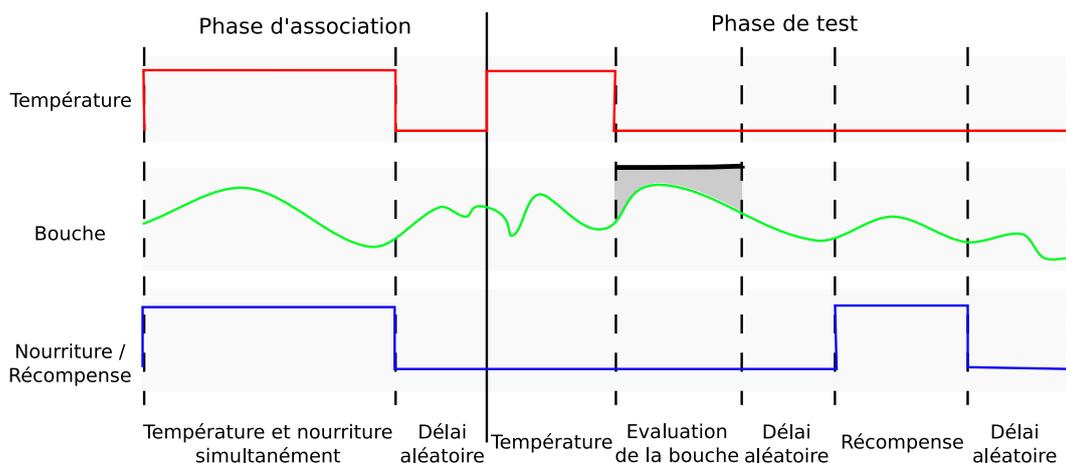


FIG. 3.6 – D'après Izquierdo et al. (2008). Structure d'un essai. Un essai est constitué de deux phases qui peuvent éventuellement être répétées. Il y a une phase d'association et une phase de test (délimitées par des traits noirs verticaux). Chaque phase est composée d'évènements (délimités par des traits pointillés verticaux). Pendant la phase d'association, la température (en rouge) et la nourriture (en bleu) sont présentées simultanément. Un délai de longueur aléatoire suit. Pendant la phase de test, une température est d'abord présentée. Ensuite, l'agent est évalué sur son comportement face à cette température. Pour cela, l'état de sa bouche (en vert) est comparé à l'action correcte à effectuer (ligne noire en gras). L'erreur est représentée par la zone grisée. Après un nouveau délai, un signal de récompense est produit en fonction de l'action effectuée auparavant. Enfin, un nouveau délai aléatoire se produit avant le début de l'essai suivant. Plusieurs phases de test peuvent suivre une phase d'association. Durant un même essai, plusieurs phases d'association peuvent se produire. L'état interne des agents n'est jamais réinitialisé pendant un essai, mais seulement entre différents essais.

Izquierdo et al. (2008) montrent que l’apprentissage d’associations liées à des valeurs continues est possible avec des agents auxquels aucun mécanisme particulier n’est fourni pour l’apprentissage. Ils montrent que le contrôleur se comporte comme une machine à état ce qui lui permet de s’adapter à l’environnement. Nous avons vu deux exemples qui illustrent chacun un type d’apprentissage différent. Cependant, dans ces deux cas, l’apprentissage est en quelque sorte inné. Il n’y a pas de construction du comportement mais une adaptation immédiate à un changement dans l’environnement. La section suivante propose une réflexion sur ce troisième type d’apprentissage que constitue l’évolution progressive de l’individu dûe à l’histoire de ses interactions avec l’environnement.

3.2.1.3 S’adapter à la nouveauté

L’évolution d’un individu est la capitalisation d’apprentissages faits à l’occasion d’adaptations successives à l’environnement. Dans les travaux sur l’apprentissage précédemment cités, le comportement à apprendre est déjà en place. L’apprentissage consiste à réadapter une réponse à un changement dans l’environnement. Cependant, il existe un autre type d’apprentissage qui conduit l’individu à construire de nouveaux types d’interaction avec l’environnement, de nouveaux comportements. Ces apprentissages sont le résultat de l’histoire unique qu’un individu a vécu. Pour Izquierdo, la niche écologique détermine fortement ce qui peut être appris : en un sens rien n’est nouveau pour l’espèce car c’est l’évolution qui détermine ce avec quoi elle interagit, ce à quoi elle “fait attention”. Néanmoins, nous faisons l’expérience de la nouveauté. Ainsi, tous les jours, nous sommes susceptibles d’apprendre quelque chose. Par nos expériences, nous construisons des connaissances. Elles peuvent également nous être transmises par autrui. Dans ce type d’apprentissage, qui nous intéresse plus particulièrement, le comportement appris est le reflet d’une série d’évènements auxquels est confronté l’individu. Ce type d’apprentissage implique une interaction plus complexe avec l’environnement.

Nous distinguons l’apprentissage “réflexe” qui apparaît à l’échelle phylogénétique, c’est-à-dire celle de l’espèce, de l’apprentissage “construit” qui, lui, apparaît à l’échelle de l’individu, l’échelle ontogénétique. Nous nous intéressons à ce dernier, c’est-à-dire à l’acquisition d’une connaissance par un individu à travers les interactions qu’il expérimente.

3.2.2 Objectif co-évolution

Dans la mesure où l’on considère que l’interaction avec son environnement va modifier la manière dont l’agent le perçoit, il y a co-construction du monde perçu et de l’agent (Varela, 1989). L’adaptation à la nouveauté, en entraînant un nouveau type d’interaction entre l’agent et son environnement, modifie donc l’environnement. Cet environnement modifié va venir en retour influencer le comportement de l’agent. Cela donne lieu à un cercle vertueux dans lequel l’agent évolue en fonction de l’histoire des interactions qu’il a eu avec son environnement.

Ainsi, des agents réellement autonomes seraient capables d’adapter par eux-mêmes leur comportement à celui des autres agents avec lesquels ils interagiraient en fonction des expériences précédentes qu’ils auraient vécues avec chacun d’entre eux. Dans le domaine de

la formation, nous pourrions imaginer un tuteur virtuel qui fournit une aide spécifique à un couple apprenant/formateur ; dans le domaine du jeu vidéo, un adversaire virtuel qui adapte sa stratégie à celle d'un joueur au fur et à mesure des parties ; ou encore, dans le domaine du théâtre virtuel, un acteur virtuel qui joue des circonstances particulières à une représentation en fonction des précédentes. Ces domaines font d'ores et déjà l'objet de travaux de recherches au sein même du LISyC (voir par exemple Tencé et al. (2010), Buche et al. (2009), De Loor et al. (2010)). Cependant, ces travaux ne s'inscrivent pas dans une approche éactionniste de l'autonomie des agents et présentent les limites mentionnées précédemment en terme d'adaptation et d'évolution. Ces exemples nécessitent des capacités de raisonnement complexes et il est bien sûr actuellement inenvisageable d'espérer obtenir de tels comportements dans le cadre éactionniste où nous nous plaçons.

Imaginons un exemple plus minimaliste faisant intervenir des agents dotés chacun d'un comportement sensorimoteur qui leur est propre. Si ces agents perçoivent les mouvements de leurs congénères et sont capables d'y répondre en modifiant leur propre comportement, nous pouvons alors imaginer que leur réunion donne lieu à une "chorégraphie" co-construite par l'ensemble des agents en interaction. Si les comportements se stabilisent, il y a eu émergence de règles d'interaction entre ces agents, que nous pouvons interpréter, pour rester dans l'image évoquée auparavant, comme les pas d'une danse. Cette danse est le résultat de la co-évolution des comportements des agents en interaction.

Le travail de Prepin et Revel (2007) s'inspire de la dynamique de l'interaction entre l'humain et la machine pour aller vers une interaction entre agents artificiels permettant la mise en place de comportements sociaux, tels que l'échange de rôle ou l'imitation. La co-construction d'un comportement commun peut être obtenu en utilisant l'imitation (Andry et al., 2001). L'imitateur apprend en adaptant ses actions à ses sensations. De plus, l'alternance des rôles imitateur-imité entre les partenaires engendre une communication qui amène à une synchronisation, prototype d'un comportement co-construit. Cependant, pour un comportement plus complexe dans lequel les rôles ne sont pas symétriques, l'imitation ne suffit pas.

Un objectif à long terme des recherches actuelles en sciences cognitives et en intelligence artificielle est d'aboutir à cette co-évolution entre un agent autonome et son environnement, l'agent construisant des connaissances à partir des expériences qu'il fait de son environnement.

De Loor et al. (2009) énoncent trois propositions pour bâtir une intelligence artificielle basée sur l'éaction qui permettrait d'atteindre une réelle co-évolution dans le contexte de la réalité virtuelle :

- 1. Irréversibilité** : un agent artificiel doit pouvoir modifier activement son environnement et celui-ci doit également évoluer. Certaines modifications doivent être irréversibles.
- 2. Ontogenèse** : pour atteindre des comportements complexes, il faut imaginer l'évolution d'un individu en terme d'auto-organisation plutôt qu'en terme d'auto-adaptation. L'agent artificiel doit pouvoir adapter ses mécanismes d'adaptation. Pour permettre cette ontogenèse, l'utilisation d'un guidage est proposée.
- 3. Signification** : il ne peut y avoir de sens que dans l'interaction entre l'agent et un utilisateur.

La conception d’un agent éactif virtuel doit être orientée par ces propositions. Ce travail de thèse s’attache essentiellement à la question de l’ontogenèse artificielle. Cependant, bien qu’elles ne soient pas l’objet direct de notre attention, les questions de l’irréversibilité et de la signification doivent rester présentes pour construire une proposition de guidage compatible avec l’objectif à long terme de co-évolution.

3.3 Synthèse

Dans ce chapitre, nous nous sommes intéressés à la pertinence de l’approche dynamique évolutionnaire pour la synthèse d’agents constitutivement autonomes. La première section a montré la possibilité d’obtenir des capacités comportementales caractéristiques de l’autonomie constitutive en utilisant cette approche. Nous avons ainsi relaté des travaux utilisant des CTRNN combinés à une approche évolutionnaire pour synthétiser des agents artificiels dont la dynamique interne est compatible avec la notion d’autonomie constitutive. Les capacités comportementales obtenues sont bien évidemment atteignables avec des méthodes de conception plus classiques. Cependant, nous rappelons que notre objectif se situe dans le cadre de l’autonomie forte et de l’éaction qui impose des contraintes bien spécifiques sur nos modèles (voir chapitre 1). Cette étude confirme l’intérêt de l’utilisation de l’approche dynamique évolutionnaire dans cet objectif.

Néanmoins, nous considérons que le point dur dans la synthèse d’agents constitutivement autonomes *suffisamment intéressants* pour susciter un sentiment d’engagement dans l’interaction réside dans leur capacité à produire un comportement variable mais également à développer de nouveaux comportements. C’est pourquoi, nous nous sommes penchés dans la seconde section sur l’évolution du comportement à l’échelle de l’individu. Nous distinguons l’apprentissage “réflexe” qui relève de l’adaptation de comportements établis à l’échelle phylogénétique et l’apprentissage “construit” qui relève de la production d’un nouveau comportement à l’échelle ontogénétique. Nous avons alors formulé un objectif à long terme, qui est de pouvoir obtenir des systèmes capables de co-évoluer avec leur environnement, ce qui en réalité virtuelle, pourra se concrétiser à travers une interaction plus fine entre l’entité artificielle et l’utilisateur.

Pour synthétiser nos réflexions sur ce qui caractérise un agent éactif virtuel, nous finalisons notre proposition d’une définition :



AEV-3 : Un agent éactif virtuel est un agent situé simulé dans un environnement virtuel, dont le comportement sensorimoteur résulte du couplage entre sa dynamique interne et la dynamique de l’environnement. La dynamique opérationnellement close de l’agent éactif pourra être obtenue en utilisant un CTRNN comme contrôleur paramétré selon une approche évolutionnaire. Son comportement correspond à une réorganisation du fonctionnement interne qui compense les perturbations de l’environnement. En particulier, un comportement nouveau pourra se construire à l’échelle ontogénétique dans l’interaction de l’agent avec son environnement.

L'objectif à long terme étant une co-évolution entre un humain et un système de réalité virtuelle, évaluer la possibilité d'un apprentissage comportemental construit suite à des interactions à l'échelle ontogénétique est un premier pas qui devient notre objectif à court terme. Par conséquent, après avoir proposé des propriétés internes et externes relatives à ce que serait un agent éactif virtuel, après avoir porté notre choix sur un modèle prometteur, le CTRNN, après avoir finalisé une proposition de définition d'un agent éactif virtuel, la seconde partie de cette thèse s'attache désormais à la question de l'ontogenèse artificielle comportementale. Ceci nous amène à proposer une étude de la possibilité d'évolution de la dynamique d'un système de manière à ce que celui-ci apprenne de nouveaux comportements en utilisant le guidage comme moyen d'interaction avec le système sans perte d'autonomie de celui-ci. Dans le chapitre suivant, nous développons notre proposition en construisant une méthode pour l'apprentissage à l'échelle individuelle de comportements sensorimoteurs chez un agent éactif.

Chapitre 4

Proposition d'une démarche pour l'apprentissage ontogénétique

RESPECTER les principes d'autonomie dans la construction d'agents artificiels tout en cherchant à reproduire des comportements interactifs et co-évolutifs est un problème très difficile pour l'intelligence artificielle. Guidés par les considérations énaactives, nous nous intéressons à la "brique de base" que sont les comportements sensorimoteurs qui, dans ce paradigme, sont considérés comme le socle pour les raisonnements de haut niveau. Dans la section 4.1, nous définissons les invariants sensorimoteurs à partir de travaux issus des sciences humaines et nous explicitons comment nous exploitons cette notion dans le domaine de l'artificiel. La section 4.2 développe ensuite l'apprentissage "construit" à travers l'acquisition d'invariants sensorimoteurs. Nous justifions notre proposition et la démarche que nous utilisons pour la mettre en œuvre.

4.1 Définition d'un invariant sensorimoteur

4.1.1 Du côté des sciences humaines

D'après Charles Lenay, spécialiste de l'étude phénoménologique de la perception, celle-ci n'est pas uniquement sensorielle mais sensorimotrice :



“Le rôle essentiel joué par l’action dans l’émergence progressive de représentations structurées montre que ce qui est perçu, reconnu, ne sont pas à proprement parler les invariants de la sensation, mais ce sont les invariants de cercles sensorimoteurs inséparables de l’activité du sujet (GIBSON 1966, 1986; PAILLARD 1971, VARELA 1979; TURVEY 1995). [...] c’est par son action que le sujet cherche et construit des règles de liaisons constantes entre action et sensation. Ces règles, qui permettent d’anticiper les effets des actions sur les sensations correspondent à la perception des formes et de leurs positions spatiales. La perception d’un objet consiste en la découverte de régularités dans la liaison entre variation d’action (mobilité de l’organe de perception) et variation de sensation (produite par ces actions).”

Lenay (1997)

La citation précédente postule donc que la perception des objets se fait grâce à la mise en place de comportements qui vérifient une régularité dans la boucle sensorimotrice. Ces comportements permettent alors une mise en évidence d’*invariants sensorimoteurs* qui peuvent être identifiés comme des signatures du comportement de perception d’un objet. Plus généralement, nous considérons que tout comportement sensorimoteur correspond à des règles liant action et sensation. Nous utiliserons cette notion d’invariants sensorimoteurs pour caractériser des comportements sensorimoteurs et nous essayons ici de la préciser.

Pour cela, nous pouvons nous appuyer sur les expériences de suppléance perceptive qui consistent à doter des sujets de dispositifs particuliers leur permettant de percevoir leur environnement sans utiliser la vision. Le dispositif utilisé par Collins et Bach-Y-Rita (1973) apparaît dans la figure 4.1(a). En simplifiant au maximum ce dispositif, *cf.* figure 4.1(b), Lenay (1997) montre qu’un sujet est capable de reconstituer l’espace autour de lui, et donc de percevoir son environnement, s’il a des sensations sur lesquelles ses actions ont des conséquences. Ainsi, même avec des données sensorielles extrêmement basiques, la capacité à diriger l’organe de sensation rend possible la perception. Son dispositif minimal est constitué d’une cellule photoélectrique connectée à un stimulateur tactile qui se déclenche de manière binaire selon un seuil d’activation de la cellule photoélectrique. Celle-ci est fixée à la main du sujet, qui peut donc explorer son environnement en déplaçant son bras et sa main. Dans ces conditions, le sujet est capable de localiser une cible lumineuse dans son environnement. Cette expérience montre la nécessité de l’exploration active dans la constitution de la perception. Le couplage du sujet à son environnement permet la mise en évidence d’invariances dans les liaisons entre action et sensation, qui autorisent la localisation relativement au sujet des sources de sensation.

Un autre type d’expérience issue de la psychologie expérimentale peut également nous aider à mieux expliciter la notion d’invariant sensorimoteur. Il s’agit d’une expérimentation reproduisant la capacité d’un sujet à s’adapter à des perturbations sensorielles fortes. Dans le chapitre 1, nous mentionnons les travaux de Di Paolo (2000) issus de ceux de Taylor (1962) qui portent sur ce sujet. Cette capacité a été étudiée par de nombreux chercheurs. Stratton a été le premier à utiliser un dispositif expérimental modifiant la vision de sujets dont il étudiait ensuite l’adaptation. Ce type de dispositif est représenté sur la figure 4.2. Les sujets s’adaptent rapidement à l’inversion du champ visuel, qu’elle soit verticale ou horizontale et sont alors en mesure d’effectuer des tâches nécessitant une coordination sensorimotrice. Les expériences de

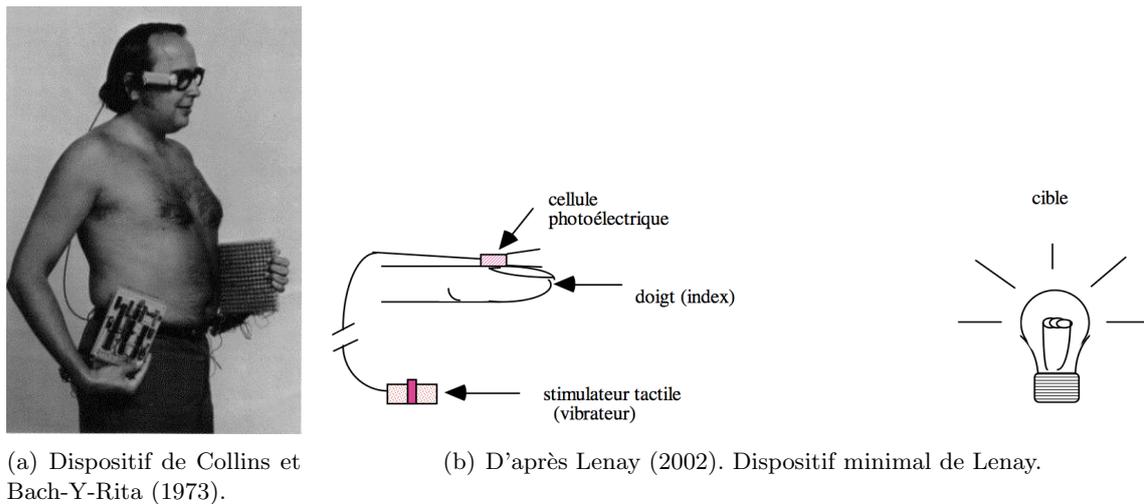


FIG. 4.1 – Dispositifs de suppléance perceptive de Bach-Y-Rita (a) et de Lenay (b)

Stratton (1897) l'amènent à considérer que la vision qui nous semble "normale", en tant que nous ne nous sentons pas désorientés, réside dans la cohérence entre le regard et au-moins un autre sens. Ainsi, un système d'objets visuels ne peut pas être "normal" ou inversé par lui-même. La notion d'inversion suppose que les impressions visuelles et les impressions motrices ou tactiles entrent en désaccord.

Dans cette formulation, nous retrouvons la notion d'invariant sensorimoteur. En effet, ici, la vision "normale", qui permet au sujet de skier par exemple est liée aux régularités qui existent entre actions et sensations. La mise en place d'un comportement est donc l'assimilation des règles de variation des sensations en fonction des actions produites. Dans le cas d'une transformation du couplage entre un sujet et son environnement, que ce soit par une modification de son corps ou l'utilisation d'un dispositif particulier, les invariants sensorimoteurs liés à certaines tâches peuvent être modifiés. Il est alors nécessaire de rechercher une nouvelle relation sensorimotrice qui permette d'effectuer ces tâches.

4.1.2 Dans le domaine de l'artificiel

Avant d'aborder l'exploitation de la notion d'invariant sensorimoteur dans le domaine de l'artificiel, nous tenons ici à préciser qu'en cherchant à mettre en évidence des invariants sensorimoteurs chez des agents énaactifs, nous nous donnons un outil pour aborder la production du comportement et sa transformation qui peut être considérée comme un apprentissage. Pour Declerck (2010), les théories dites de la perception active réduisent la perception à un invariant sensorimoteur. Or, ceci fait débat au sein des sciences cognitives car pour Declerck par exemple, la perception est possible par la recherche et la maîtrise d'invariants sensorimoteurs mais elle ne s'y réduit pas. L'objet qui est perçu n'est pas un invariant sensorimoteur, le sens est donné par une activité non motrice.

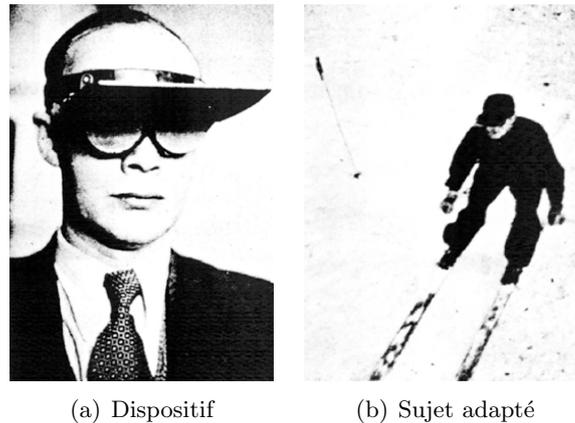


FIG. 4.2 – (a) Dispositif expérimental d’inversion visuelle. (b) Un sujet adapté est capable de skier en portant le dispositif d’inversion visuelle.



“L’activité sensorimotrice [...] permet d’incarner de la “signification” dans l’actualité d’un champ phénoménal. Elle fournit un vêtement qualitatif à des creusets de signification, mais cette signification ne se réduit pour sa part en aucune manière à cette activité.[...] Cette distinction nous semble d’autant plus nécessaire que la plupart des théories dites de la perception active ont parfois tendance à vouloir réduire l’expérience perceptive à sa couche sensorimotrice.”

Declerck (2010)

Dans notre domaine qui est celui du virtuel, la question du sens, celle de la possibilité d’une expérience perceptive chez un agent artificiel nécessiterait que les sciences cognitives avancent dans ce débat. Nous ne pourrions de toute façon pas accéder à cette expérience si tant est qu’elle puisse avoir lieu. Les comportements que nous étudions font sens pour un observateur humain et nous cherchons à avancer dans la compréhension de l’apprentissage de ces comportements par un agent éactif sans qu’il soit nécessaire de déterminer l’existence potentielle d’une signification pour l’agent.

Dans un espace virtuel, le couplage entre un agent éactif et l’environnement est réalisé par des capteurs et des effecteurs, comme c’est le cas dans un environnement réel pour un robot. La simulation nous permet d’avoir accès aux données captées et à l’activité des effecteurs. L’analyse de leur relation peut permettre de mettre en évidence des invariants sensorimoteurs nécessaires à la réalisation de certains comportements. Ceci constitue une piste dans l’étude du comportement d’agents autonomes.

Nous avons simulé deux agents artificiels dont nous contrôlons les déplacements de manière à ce que l’un effectue un carré et l’autre un sablier dans un espace en deux dimensions, comme illustré sur les figures 4.3 et 4.4. Il ne s’agit pas ici d’agents éactifs, le comportement est programmé pour suivre les trajectoires précises. Nous enregistrons les valeurs sur ses capteurs et les commandes motrices pendant tout l’épisode. Il y a deux repères dans l’environnement et un capteur par repère. La valeur captée dépend de la distance de

l'agent au repère comme ceci : $I_i = gain * (1 - d_i/D)$. Nous remarquons que des motifs particuliers apparaissent dans le lien entre les actions et les sensations de l'agent. Ces motifs peuvent être mis en relation avec le comportement de l'agent. Ainsi, à chaque étape correspond un type d'invariant sensorimoteur. Ces simulations ont pour objectif de donner un aperçu global de ce que nous entendons par mise en évidence d'invariant sensorimoteur et non pas de réaliser une analyse détaillée de ceux-ci. Les détails sur cette illustration sont donc fournis en annexe A.

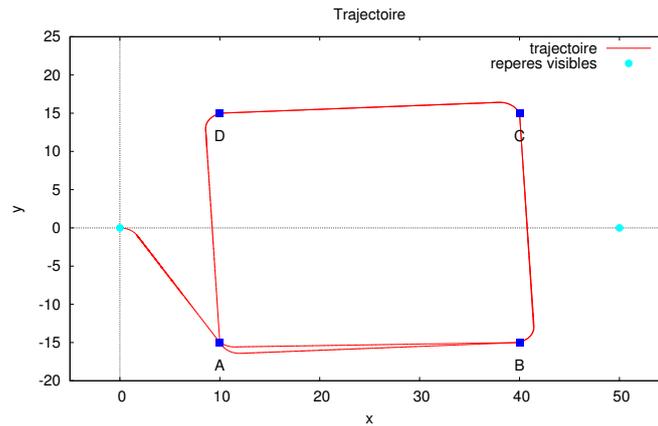


FIG. 4.3 – Trajectoire d'un agent dans un environnement en 2D contrôlé pour réaliser un carré

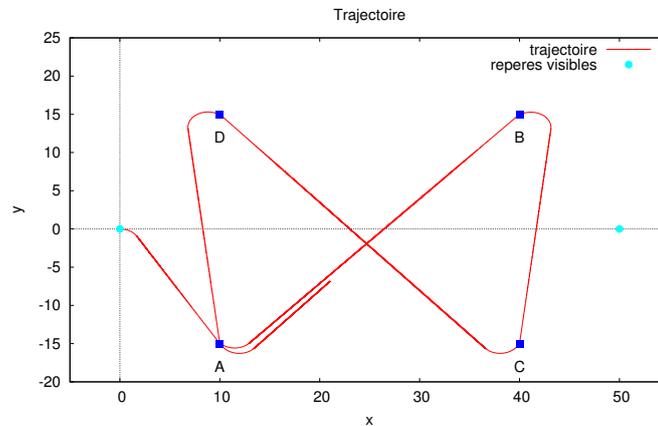


FIG. 4.4 – Trajectoire d'un agent dans un environnement en 2D contrôlé pour dessiner un sablier

4.2 Acquisition d'invariants sensorimoteurs

L'apprentissage à l'échelle ontogénétique implique une assimilation d'invariants sensorimoteurs, déterminée au-moins en partie par l'influence de l'environnement. Par nécessité de simplification, nous mettons de côté ce que nous appelons l'indéterminisme des compor-

tements dans le chapitre 3¹, qui correspond à la capacité d’un agent autonome à maintenir différents buts en même temps et à déterminer quel comportement il adopte à un instant donné en fonction à la fois de facteurs externes et de facteurs internes. Pour mieux nous concentrer sur l’apprentissage, nous considérerons des agents dotés d’un comportement propre unique, correspondant à certains invariants sensorimoteurs. L’acquisition d’un nouveau comportement sensorimoteur consiste alors en la transformation de ces invariants sensorimoteurs de manière irréversible.

4.2.1 D’un invariant sensorimoteur à un autre

Cette transformation pourrait être, par exemple, de faire passer un agent du comportement correspondant à la figure 4.3 à celui correspondant à la figure 4.4. Autrement dit, un agent contrôlé par un CTRNN, et donc dont on ne programme pas la trajectoire, a un comportement “inné” qui consiste à se déplacer pour dessiner un carré entre deux balises. Au niveau sensorimoteur, des régularités correspondantes à des invariants sensorimoteurs se retrouvent, *cf.* figure A.1 en annexe A. L’apprentissage en modifiant le comportement modifie également ces invariants sensorimoteurs jusqu’à obtenir le comportement de déplacement en sablier et donc des invariants sensorimoteurs sensiblement différents, *cf.* figure A.2 en annexe A.

Cependant pour que ce soit un apprentissage au sens où nous l’entendons, il faut que cette transformation du comportement soit déclenchée par une interaction avec l’environnement. En effet, si la transformation du comportement se fait par elle-même, comme le résultat d’un développement “programmé” de l’individu, il s’agit d’une aptitude “innée” et non d’un apprentissage réalisé par l’individu. Pour aborder réellement l’ontogenèse, nous devons atteindre un développement propre à un individu et à son histoire. Ainsi, obtenir des agents qui passeraient de la figure 4.3 à la figure 4.4 sans élément déclencheur ne nous satisferait pas. L’idéal serait de pouvoir déclencher un changement chez ces agents les amenant potentiellement à différents comportements, que nous pourrions alors qualifier “d’acquis”. Suivant l’interaction menée avec le même agent éactif, celui-ci adopterait alors un comportement particulier reflétant son histoire. Pour un même agent, les scenarii 4.5(a) et 4.5(b) représentés ci-après seraient par exemple possibles.

Pour illustrer cette transformation que nous appelons apprentissage ontogénétique, nous prenons un exemple plus simple². Dans un environnement en une dimension, schématiquement une droite, nous plaçons une balise qui est mobile. Un agent est muni d’un capteur sensible à cette balise et d’un moteur qui lui permet de se déplacer dans cet environnement. Le comportement “inné” de cet agent, déterminé à sa conception, consiste à avancer dans la direction signalée par la balise (comportement A). L’invariant sensorimoteur pour l’agent est la relation entre une commande motrice et une différence sur les capteurs correspondant à une

¹ Nous voulons dire ici que nous ne chercherons pas à obtenir des agents qui choisissent un comportement parmi ceux qu’il maîtrise. Cela ne signifie pas que nous voulons pour autant obtenir un comportement “figé”, précis, puisque l’autonomie constitutive engendre une possibilité de variation dans le comportement et que nous appelons cette variabilité pour susciter un engagement.

² Cet exemple sera repris dans le chapitre suivant dans une expérience utilisant des CTRNN.

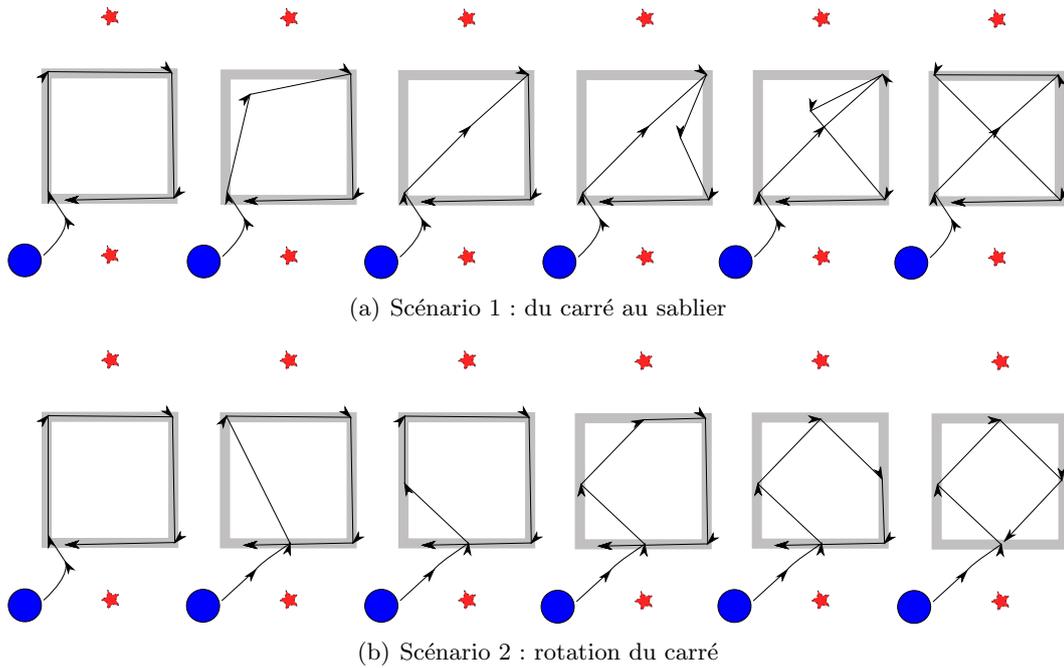


FIG. 4.5 – Scénarii possibles de transformation du comportement sensorimoteur d’un agent.

diminution de la distance à la balise. Un agent parti a priori dans la mauvaise direction doit modifier son activité motrice pour retrouver cette relation (comportement A1). S’il va dans la bonne direction, il adopte le comportement A2 consistant à poursuivre la balise, l’invariant sensorimoteur sera de maintenir cette relation. Si l’on considère un autre comportement (B) visant à rester à une distance fixe de la balise, l’invariant sensorimoteur sera la relation entre une commande motrice à ajuster pour obtenir une variation nulle ou du moins la plus faible possible sur les capteurs. La variation sur les capteurs liée à une variation de la commande motrice obéit en effet à une règle déterminée par la position de la balise par rapport à l’agent. Un comportement sensorimoteur particulier consistera donc à ajuster la commande motrice en fonction de la règle en vigueur et d’un objectif.

4.2.2 Ontogenèse : le résultat de l’histoire d’un individu...

Nous verrons dans le chapitre 5 qu’il est possible d’obtenir, en utilisant des techniques d’évolution artificielle, des agents artificiels qui adoptent les comportements A et B. Ceci est le résultat d’une complexification du comportement au fil des générations. C’est ce que nous appelons une évolution à l’échelle phylogénétique, puisque chaque génération hérite des progrès des générations précédentes. Les événements déclenchant la modification du comportement sont des mutations qui interviennent à l’échelle de “l’espèce”. Dans le cas des travaux que nous avons présentés dans la section 3.2.1, il s’agissait également d’apprentissage à l’échelle phylogénétique. Le comportement se modifie de manière “innée”. Par exemple, l’“espèce” pourrait “s’adapter” de manière à ce qu’un individu “naisse” avec le comportement

A et bascule sur le comportement B suite à un évènement quelconque ou bien après un délai déterminé. Nous avons bien un individu qui met en œuvre les différents comportements, mais ce développement est en quelque sorte prédéfini avant même qu’il n’interagisse avec l’environnement.

L’apprentissage à l’échelle ontogénétique que nous voulons obtenir doit faire la part belle à la singularité de l’histoire des interactions entre chaque agent et son environnement. Nous voulons ici que les évènements à l’origine d’un changement dans le comportement de l’agent adviennent à l’échelle de l’individu. Prenons un autre exemple plus concret. Un agent est sur une piste de ski, balisée par des drapeaux. Il sait descendre la piste en longeant les drapeaux toujours sur sa droite. Nous voulons le faire slalomer entre les drapeaux. Pour interagir avec lui, nous pouvons lui crier “gauche” ou “droite”. L’agent commence la descente et passe à droite du premier drapeau après quoi il se dirige vers le deuxième drapeau toujours vers la droite. Nous lui crions alors “gauche”. L’agent tourne vers la gauche. Si le simple fait d’avoir crié permet à l’agent de finir la descente en slalomant, nous n’avons pas réellement affaire à un comportement construit. Il s’agit d’un comportement “inné”, potentiellement déjà présent chez l’agent. à la conception, l’agent “sait” déjà slalomer et l’évènement déclencheur qui est le cri ne fait que le faire basculer d’un mode de comportement à un autre. Pour obtenir un comportement nouveau, l’environnement doit avoir une plus grande incidence sur celui-ci. Ainsi, par exemple, si nous crions à l’agent de tourner à chaque drapeau, il pratique un slalom régulier, mais si nous ne lui crions nos consignes que tous les deux drapeaux, il s’agira d’une autre forme de slalom, d’un autre comportement légèrement différent.

Pour obtenir un comportement nouveau, il est donc nécessaire d’avoir un environnement plus ouvert, qui perturbe le comportement de l’agent de manière à l’amener à un changement dans son activité motrice. Cet environnement doit pouvoir produire des évènements pour lesquels la réaction de l’agent n’a pas été prévue à sa conception. Bien sûr, dans notre exemple, la réaction de l’agent aux cris est préprogrammée. Cependant, le comportement auquel nous pouvons attribuer un sens dépend lui de la séquence d’évènements associant cris et drapeaux selon une certaine chronologie. Cette association n’est pas donnée *a priori* à l’agent. Pour arriver à une expérience singulière entre l’agent et son environnement, il est nécessaire par conséquent que le comportement de l’environnement lui-même ne soit pas préprogrammé et qu’il soit perturbé, influencé, par celui de l’agent. Ainsi, le comportement d’un agent pourra être profondément lié à l’histoire de ses interactions avec l’environnement. Cette construction du comportement correspond à la notion de co-évolution de l’agent et de son environnement.

Cette situation peut être initiée de deux manières. D’une part, l’environnement de l’agent peut abriter d’autres agents artificiels. Dans ce cas, nous avons une société d’agents artificiels qui s’entre-influencent et développent des comportements coordonnés au sein de leur système. D’autre part, l’environnement de l’agent artificiel peut être modifié par un utilisateur, qui tient compte du comportement de l’agent pour l’orienter. Cela permet à l’utilisateur humain de garder la main sur l’interaction de l’agent avec son environnement et de lui donner du sens (De Loor et al., 2009). Notre démarche est ancrée dans des problématiques de réalité virtuelle et nous voulons également que l’humain puisse donner du sens à l’interaction de l’agent éactif avec son environnement, c’est pourquoi ce deuxième cas nous intéresse.

4.2.3 ... dans un environnement

L'intervention d'un utilisateur dans l'environnement virtuel doit donc être perceptible par l'agent artificiel. Elle se traduit sous la forme d'une modification des sensations de l'agent, d'une perturbation des régularités sensorimotrices. Ce peut être par exemple, le déplacement d'un pot de fleurs qui jusqu'ici était immobile. Dans ce cas, l'invariant sensorimoteur de l'agent par rapport à ce pot de fleurs sera modifié. D'une situation dans laquelle sans variation d'action, il n'y a pas de variation dans les sensations, l'agent passe à une situation dans laquelle sans variation d'action, il y a une variation dans les sensations. Cependant, pour que le comportement évolue, il faut également que l'agent réagisse à la perturbation initiée par l'utilisateur. Par conséquent, le déplacement du pot de fleurs doit entraîner un mouvement de l'agent. Pour cela, nous devons raisonnablement supposer qu'il préexiste une relation sensorimotrice par rapport à l'objet chez l'agent. Il y aura donc bien un ancrage "inné" chez l'agent, mais qui sera à un niveau comportemental générique. C'est le cas de la relation entre les cris et les virages chez l'agent qui slalome dans notre exemple précédent.

Imaginons cette fois une porte à ouverture automatique. Un agent est, de manière préétablie, capable d'ouvrir cette porte car son comportement tient compte des relations sensorimotrices qui existent entre ce qu'il fait et ce qu'il perçoit de la porte. Comment, dans cette situation, un utilisateur peut-il intervenir pour modifier l'environnement de manière à transformer le comportement, par exemple pour que l'agent traverse la porte une fois celle-ci ouverte ? Agir directement sur le comportement de l'agent en forçant ses commandes motrices pour qu'il passe la porte ne semble pas pertinent, puisque cela risque de rendre les règles intégrées par l'agent inutiles. De la même manière, agir sur le comportement de la porte en la déplaçant pour que l'agent se retrouve de l'autre côté aura pour conséquence de transformer le couplage sensorimoteur de l'agent le rendant inexploitable. Dans un cas, nous pouvons parler d'agent désafférenté, qui génère une activité motrice mais qui ne se réalise pas. Dans le second cas, ce sont les afférences sensorielles qui sont perturbées. Pourtant, dans ces deux situations, le couplage "normal" entre l'agent et la porte est perturbé, *cf.* figure 4.6, et il n'y a plus de lien de causalité entre les actions et les sensations. Ces deux modes d'intervention semblent alors inadaptés³.

Nous pensons que l'utilisateur doit agir de manière indirecte pour influencer le comportement de l'agent par rapport à cet objet avec lequel il a déjà une relation sensorimotrice. Il s'agit donc bien de transformer un invariant sensorimoteur non pas dans le sens où les règles sont complètement changées, mais dans le sens où les actions de l'agent sont contraintes différemment par ces règles, ces dernières étant étendues pour permettre un comportement plus complexe. Nous illustrons dans la figure 4.7 ce que nous entendons par modification de l'invariant sensorimoteur à travers une action indirecte d'un utilisateur. Pour plus de lisibilité, nous traçons la position de l'agent et l'ouverture de la porte, plutôt que les données internes de l'agent. Des événements externes sont déclenchés quand l'agent s'approche de la porte et que celle-ci s'ouvre. Ces événements viennent perturber le comportement de l'agent qui s'approche plus de la porte. Le comportement se transforme progressivement et, en réagissant

³ En fait, ces deux types d'interaction ne sont pas distinctes du point de vue de l'agent sauf si celui-ci peut "ressentir" la modification de l'activité motrice. En l'absence de ce retour, le lien sensorimoteur est rompu et finalement cet utilisateur "fait à la place de" l'agent.

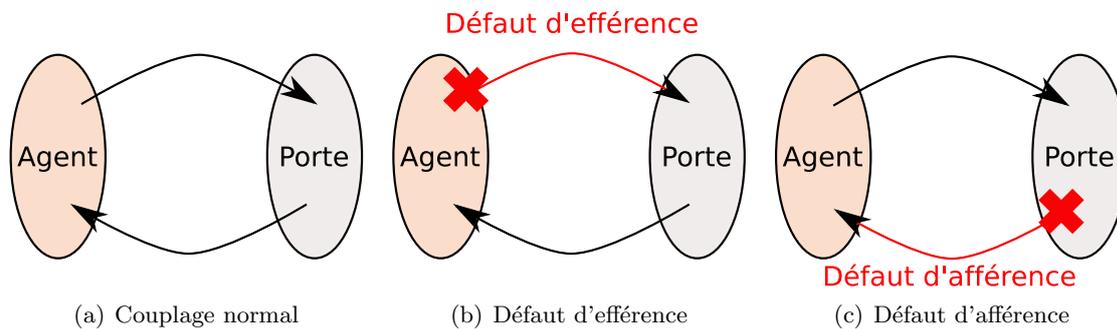


FIG. 4.6 – Perturbations du couplage sensorimoteur agent/porte. (a) Mode de couplage normal. (b) En forçant les commandes motrices de l’agent, nous agissons sur le couplage en introduisant un défaut d’efférence. (c) En agissant sur le comportement de la porte, un défaut d’afférence vient perturber le couplage.

à ces évènements, l’agent traverse la porte.

Notre proposition est d’utiliser un guidage en association avec des stimuli auxquels l’espèce est par nature sensible sans avoir de réponse particulière. Le guidage va constituer une interaction privilégiée entre l’environnement et l’agent dans la mesure où il va permettre la mise en place d’une influence de l’environnement sur l’agent dirigée vers la construction d’un comportement nouveau, comme l’illustre la figure 4.8. Cette proposition constitue donc une piste qui peut mener vers une co-évolution. Pour cela, il faudra cependant prendre également en compte l’influence de l’agent sur son environnement.

Le guidage que nous proposons ne doit cependant pas être compris comme un signal associé à des mécanismes dédiés à l’apprentissage, tel qu’un renforcement. Il s’agit simplement d’un élément de la dynamique de l’environnement que nous contrôlons pour initier une interaction particulière avec l’agent. L’apprentissage de la bicyclette peut être une métaphore du type d’apprentissage que nous voulons mettre en place. En effet, tout comme un parent accompagne l’enfant qui donne ses premiers coups de pédales, le guidage accompagne l’agent dans son interaction avec un environnement. Il ne s’agit ni de renforcement, ni d’imitation, mais plutôt d’une orientation du comportement au fur et à mesure qu’il se construit. Chez l’enfant, il existe déjà des éléments qui le prédisposent à apprendre, il est motivé et sait écouter ses parents. Ainsi, l’apprentissage que nous visons doit s’ancrer dans un comportement préexistant.

Nous imaginons ce guidage comme un encadrement de l’agent qui viendrait corriger un comportement qui s’écarte de l’objectif, plutôt qu’imposer un modèle comme cela pourrait être le cas avec de l’imitation ou lors d’un conditionnement. Cette solution conserve l’autonomie de l’agent qui construit ses propres règles à partir des limites fournies par le guidage.

De plus, le guidage offre selon nous des possibilités pour aborder l’apprentissage de comportements relativement complexes car il doit permettre un apprentissage par étapes. En effet, le guidage n’est pas lié à un but immuable, mais au comportement de l’agent à un instant donné pour un observateur par rapport à un objectif à ce même instant. Ceci n’exclue

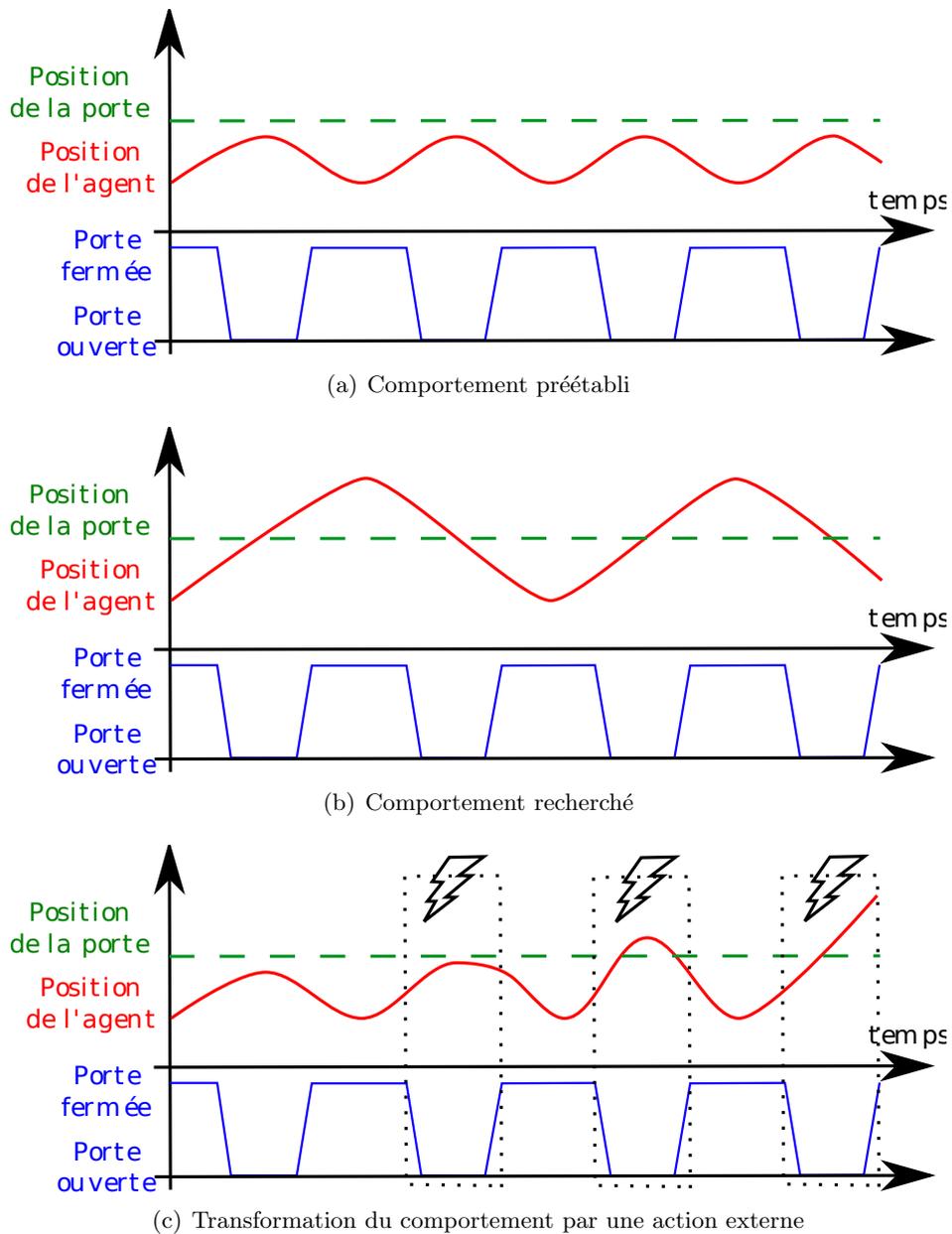


FIG. 4.7 – Transformation progressive d'un comportement sensorimoteur. (a) Comportement prédéterminé : l'agent s'approche de la porte qui s'ouvre, puis recule, la porte se referme, etc... (b) Comportement recherché : l'agent s'approche de la porte, celle-ci s'ouvre, il traverse, la porte se referme, etc... (c) Transition : un utilisateur déclenche des évènements qui vont perturber l'agent et l'amener à modifier son comportement sensorimoteur aux moments jugés clés par cet utilisateur, de manière à ce que l'agent modifie durablement son comportement par rapport à la porte.

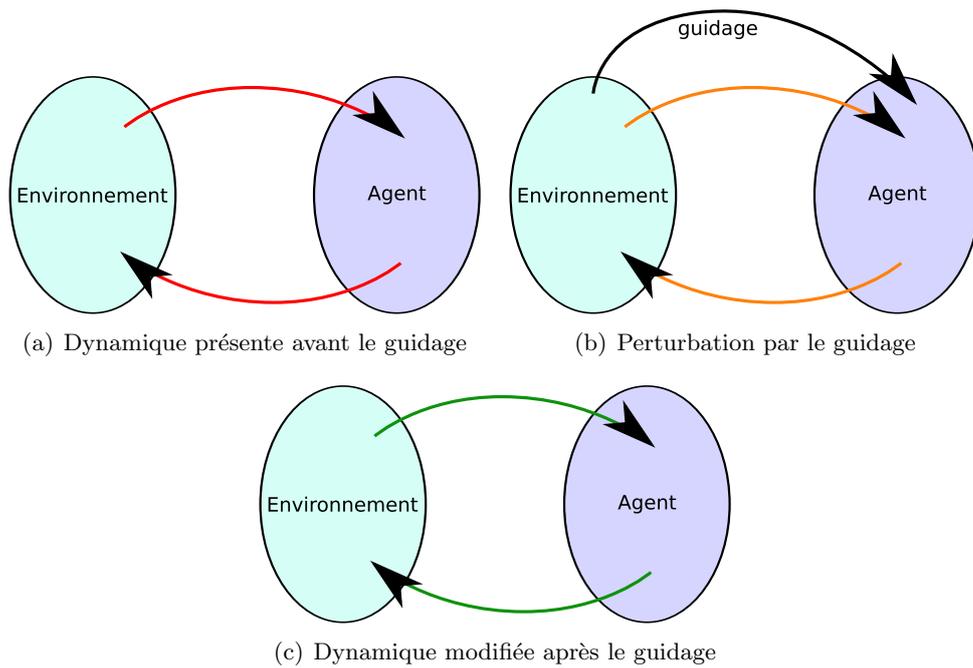


FIG. 4.8 – (a) Avant l’apprentissage, l’agent maintient une certaine dynamique avec son environnement. (b) Le guidage introduit un changement dans la dynamique de l’environnement, ce qui se répercute sur la dynamique de l’agent. (c) Après l’apprentissage, l’agent maintient une nouvelle dynamique avec son environnement, même en l’absence du guidage.

pas que si l'agent a un comportement satisfaisant au regard de l'observateur, celui-ci puisse décider de modifier l'objectif pour complexifier ce comportement et enrichir le répertoire de l'agent.

Enfin, le guidage nous intéresse car il s'agit d'une interaction sensorimotrice qui peut se transposer directement dans des applications de réalité virtuelle. Ainsi, l'utilisateur d'une application de réalité virtuelle pourrait jouer le rôle de guideur face à des agents artificiels et ainsi influencer leur comportement sans pour autant leur imposer une loi de fonctionnement. Ceci contribuerait à un sentiment d'autonomie des agents et à la créativité dans les interactions réel/virtuel.

Pour valider cette proposition, notre ambition est de faire apprendre par guidage de nouveaux invariants sensorimoteurs à des agents énaactifs virtuels. Cette tâche est cependant complexe et la question du développement incrémental se pose. Nous avons vu dans le chapitre 3 que certains auteurs (Beer et Gallagher, 1992) montrent qu'il est plus efficace de poser le problème entier directement, alors que d'autres (Tuci et al., 2002) prônent la résolution de la tâche par étapes de plus en plus complexes. Cette contradiction n'est pourtant qu'apparente puisque dans le premier cas, c'est la construction de l'agent qui est incrémentale, il s'agissait en effet d'assembler des "morceaux d'agents" évolués séparément, alors que dans l'autre cas c'est le comportement qui est progressivement complexifié. Ces deux travaux vont dans le sens d'une démarche écologique, qui considère un agent "entier" dans "sa niche". L'évolution incrémentale a pour but de faciliter la production d'un comportement complexe en s'assurant que la recherche part d'une population qui possède déjà certains traits comportementaux jugés comme nécessaires pour passer à ce comportement complexe. Cependant, à chaque étape, l'analyse de l'influence d'un trait comportemental sur l'autre conditionne la réussite de l'évolution. Toute la difficulté de cette approche réside donc dans la définition d'une fonction d'évaluation du comportement d'un agent par rapport à une tâche ou une sous-tâche. à chaque étape, la sous-tâche doit permettre d'atteindre des conditions qui rendent l'étape suivante possible mais aussi être suffisamment "facile" pour qu'une solution soit trouvée.

Notre démarche sera donc progressive. Tout d'abord, nous nous assurerons qu'un même type d'agents est capable d'adopter des invariants sensorimoteurs différents. Ensuite, nous nous assurerons qu'il est possible de faire apprendre par guidage un comportement simple de type associatif à ces agents. Enfin, nous réunirons ces deux exigences pour faire passer des agents artificiels d'un comportement sensorimoteur à un autre en les guidant.

En choisissant les moments où il déclenche des événements de guidage, l'utilisateur oriente l'apprentissage car il peut tenir compte de la réaction de l'agent. Le guidage doit donc être un processus d'accompagnement progressif. Le guideur suit, durant tout l'épisode d'apprentissage, le comportement effectif de l'agent et doit idéalement adapter les événements qu'il produit aux progrès réalisés.

Cependant, il peut être souhaitable de pouvoir dans un premier temps automatiser le guidage pour tester nos propositions. Ceci suppose que le guideur soit programmé pour déclencher certains événements de guidage en fonction du comportement de l'agent. Il paraît ambitieux de tenter de programmer un guideur capable de mesurer les progrès de l'agent au cours de l'apprentissage et de s'adapter en conséquence au niveau de maîtrise de l'agent à mesure que celui-ci évolue. Surtout, programmer un guideur qui tienne compte de la

réaction de l’agent à différents évènements et construisent ses propres règles de guidage est ici inenvisageable. Pour pouvoir utiliser des guideurs automatiques, nous devons par conséquent avoir des agents qui réagissent de manière prévisible à une stratégie de guidage. Nous parlons ici de la façon dont le guidage perturbe l’activité motrice de l’agent et non pas d’une éventuelle modification de la dynamique sensorimotrice globale liée au comportement que nous voulons lui apprendre. En effet, si l’apprentissage n’est le résultat que du seul guidage sans incidence des autres éléments de l’environnement, nous retombons dans un cas correspondant à une adaptation “innée” du comportement.

Il est pourtant possible de produire de manière automatique un guidage plus ou moins fin. Nous proposons ici différents profils de guideurs automatiques :

- ▷ Le guideur “tout ou rien”. Il se manifeste dès qu’il considère que le comportement de l’agent ne correspond pas à celui que nous voulons lui apprendre, sans tenir compte d’éventuels progrès. Ce type de guideur est facilement implémentable puisqu’il suffit qu’il connaisse les actions de l’agent et celles qui sont attendues dans sa situation. Si elles diffèrent, un évènement est produit.
- ▷ Le guideur “informatif”. Il produit des signaux de guidage différents selon l’erreur de l’agent. Cela peut se réaliser par des évènements de guidage de nature variée (“gauche” ou “droite” dans l’exemple du slalom) ou par un signal dont l’intensité dépend de l’écart entre le comportement de l’agent et celui attendu. En ce sens, le guidage contient plus d’information. Ceci nécessite de définir des règles de guidage plus fines.
- ▷ Le guideur “progressif”. Il est possible d’assouplir un peu le guidage tout en gardant un processus automatique simple. Un guideur que nous qualifierons de “progressif” fonctionnera toujours selon des règles strictes, mais modifiera l’évènement lié au guidage en fonction du comportement passé de l’agent. Par exemple, face à un agent qui s’écarte du comportement désiré, le guideur “progressif” déclenche tout d’abord un évènement sous la forme d’un signal faible. Puis, si l’agent ne modifie pas son comportement, le guidage augmente progressivement. Ce guideur doit être doté d’une mémoire.

Des profils combinant les caractéristiques de ces guideurs types peuvent également être envisagés. Ainsi, nous pouvons imaginer un guideur informatif progressif. Cependant, l’information liée à l’attente au niveau comportemental peut interférer avec la progression du guidage. Si nous voulons éviter d’avoir à mettre en place un couplage complexe entre l’agent et le guideur, il sera alors vraisemblablement nécessaire de limiter le caractère informatif pour pouvoir mettre en place la progression. De manière simple, nous pouvons avoir un guidage dont le signe forme le caractère informatif et l’intensité l’aspect progressif. Ceci permet de n’avoir qu’un capteur continu au niveau de l’agent pour permettre le couplage.

4.2.4 Positionnement par rapport aux techniques d’apprentissage artificiel classiques

L’apprentissage ontogénétique que nous cherchons à obtenir est une forme d’apprentissage artificiel. Le paramétrage du contrôleur constitue un autre apprentissage artificiel, qui a lieu, lui, à l’échelle phylogénétique. Nous ne développerons pas ici les techniques existantes

en terme d'apprentissage artificiel, pour lesquelles nous renvoyons à la lecture des ouvrages de Cornuéjols et Miclet (2002) ou de Haykin (2009).

Si les algorithmes évolutionnaires utilisés pour paramétrer les contrôleurs peuvent être considérés comme des exemples d'apprentissage supervisés, l'apprentissage ontogénétique par guidage se rapproche plus d'un apprentissage par renforcement. Dans un cas, nous avons un apprentissage qui se fait hors ligne dans lequel la sélection des solutions se fait par comparaison avec un résultat recherché. Dans le second cas, l'apprentissage se fait en ligne et l'agent obtient une information de l'environnement sur la qualité des actions qu'il effectue. Cependant, cet apprentissage par renforcement n'est pas caractérisé par une récompense (ou punition) explicite de telle ou telle action du système et les algorithmes classiques de renforcement (Sutton et Barto, 1998) ne sont pas applicables. En effet, il n'existe pas de représentation explicite dans le contrôleur des états possibles et des actions à y associer. C'est une des propriétés nécessaires chez des agents éactifs virtuels tels que nous les avons défini, pour laquelle notamment nous avons choisi les CTRNN. Comme nous le verrons, ceci rend les choses bien plus difficiles à mettre en œuvre qu'avec d'autres modèles compatibles avec les approches classiques de renforcement.

4.3 Synthèse

Nous avons présenté la notion d'invariants sensorimoteurs, qui sont les régularités qui existent entre des actions et des sensations, permettant d'accéder à des relations entre ce qui est fait et ce qui est senti. Cette notion issue des sciences cognitives peut être utilisée dans le domaine de l'artificiel en laissant de côté la question du sens.

Nous l'utilisons comme un outil qui nous permet d'aborder les comportements sensorimoteurs dans une approche dynamique. Nous avons ainsi illustré comment la notion d'invariant sensorimoteur peut être exploitée dans la caractérisation du comportement d'un agent artificiel. Nous avons ensuite développé notre objectif qui est de faire acquérir de nouveaux invariants sensorimoteurs à un agent éactif.

Pour cela, nous avons proposé de transformer des invariants sensorimoteurs. La notion de co-évolution a été explicitée en tant que transformation des invariants sensorimoteurs reflétant une évolution de l'agent à l'échelle ontogénétique mais reflétant aussi une évolution de l'environnement. Ceci nous a permis de préciser la notion d'apprentissage ontogénétique qui est le résultat de la co-évolution d'un agent éactif avec un environnement.

Afin d'être exploitable dans un contexte de réalité virtuelle, cette co-évolution doit passer par l'intervention d'un utilisateur qui a un objectif comportemental pour l'agent. Des critères ont été définis pour la mise en place d'une interaction qui permette d'aboutir à un apprentissage ontogénétique.

Comme nous l'avions annoncé au chapitre 3, nous nous appuyerons sur la notion de guidage qui est désormais justifiée comme moyen d'action indirecte de l'utilisateur sur le comportement de l'agent par rapport à un élément de l'environnement. Ce guidage constitue

une influence qui peut s’exercer sur l’agent en perturbant la boucle sensorimotrice sans pour autant mettre à mal de manière brutale les règles de liaison sensorimotrice de l’agent pour l’objet en question. Le guidage advient sous la forme d’évènements déclenchés par l’utilisateur aux moments qu’il juge pertinent. Ces évènements sont captés par l’agent et vont provoquer une réaction de celui-ci qui réorganise sa dynamique pour en tenir compte. Le guidage peut alors avoir des conséquences sur le comportement sensorimoteur de l’agent. Par la réorganisation qu’il entraîne, le guidage doit aider l’agent à assimiler les évènements sensorimoteurs concomitants. Ainsi, il n’est pas spécifique à l’apprentissage d’un comportement précis, mais doit permettre de faire assimiler à un agent artificiel une variété de comportements. De cette manière, l’apprentissage se réalise à l’échelle ontogénétique et reflète l’histoire unique d’un agent en couplage avec un environnement particulier pendant une période précise.

Dans la suite, nous présentons trois expériences visant à évaluer si le guidage peut effectivement permettre un apprentissage ontogénétique chez des agents contrôlés par des CTRNN. La première expérience consiste à complexifier un comportement sensorimoteur en utilisant une approche phylogénétique. Puis nous nous intéressons à l’apprentissage d’une association par guidage et enfin nous tentons de réaliser un apprentissage de comportement sensorimoteur à l’échelle ontogénétique. Nous prévenons le lecteur que ces expérimentations ne se sont pas révélées très fructueuses, mais qu’elles permettent de mieux cerner les difficultés et les moyens futurs à déployer dans la poursuite de cet objectif.

Chapitre 5

Évolution phylogénétique incrémentale de comportements sensorimoteurs

La première étape pour atteindre un apprentissage sensorimoteur à l'échelle ontogénétique consiste à évaluer si des agents contrôlés par des CTRNN peuvent respecter des invariants sensorimoteurs différents. Notre objectif étant de faire passer les agents de l'un à l'autre en les guidant, ces deux invariants doivent être différents mais suffisamment proches. Nous devons donc proposer deux tâches sensorimotrices différentes nécessitant les mêmes capacités sensibles et motrices. La section 5.1 propose une expérience répondant à cet objectif dans un environnement en 2 dimensions. Les difficultés rencontrées nous amène à nous engager dans une simplification du problème qui préserve les caractéristiques nécessaires à notre argumentation. La section 5.2 expose une expérience minimaliste dans laquelle nous caractérisons les relations sensorimotrices chez un agent. Enfin, dans la section 5.2.1, nous utilisons une approche phylogénétique pour complexifier le comportement sensorimoteur obtenu.

5.1 Comportement sensorimoteur dans un espace en 2 dimensions

L'expérience que nous voulons construire veut vérifier qu'il est possible de faire évoluer des agents qui avec la même architecture expriment des invariants sensorimoteurs différents, mais suffisamment proches pour qu'il soit possible de faire adopter à l'un des agents l'autre invariant sensorimoteur. Notre première idée a été de faire réaliser des figures à nos agents dans un environnement en deux dimensions. Il est possible de distinguer deux positions (non symétriques) dans un espace en 2-D en ayant accès à la distance de ces positions à deux points. Nous plaçons dans l'environnement deux balises qui servent de repères.

5.1.1 Tâche

Les agents sont situés dans un environnement contenant deux balises qui servent de référents spatiaux. La tâche d'un agent est de se déplacer dans cet environnement selon une trajectoire correspondant à une figure particulière. En l'occurrence, un agent de la population A doit réaliser un carré comme indiqué sur la figure 5.1(a). La tâche d'un agent de la population B est de dessiner un sablier (figure 5.1(b)).

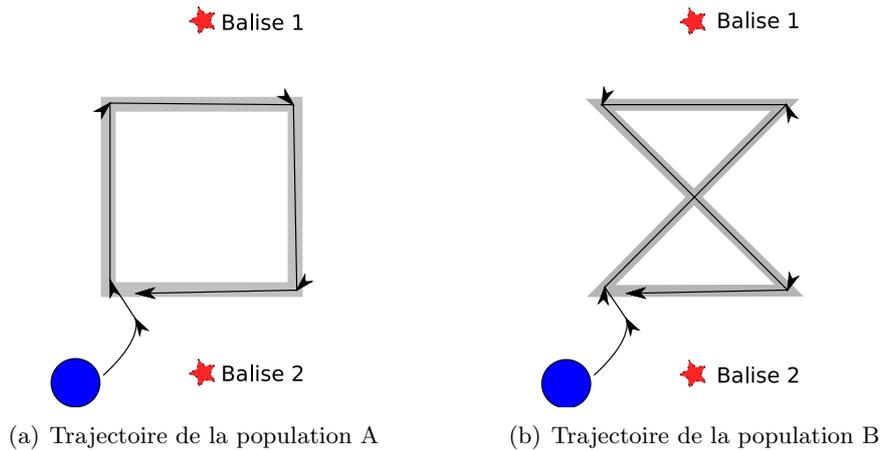


FIG. 5.1 – Trajectoires correspondant aux deux invariants sensorimoteurs recherchés. (a) : les agents de la population A ont pour tâche de dessiner un carré. (b) : les agents de la population B doivent réaliser une figure en forme de sablier.

5.1.2 Agent

Un agent est doté de deux capteurs sensibles chacun à une balise. Deux moteurs lui permettent de se déplacer dans l'environnement, la différence de vitesse entre les deux moteurs permettant une rotation. Le contrôleur d'un agent est un CTRNN. La figure 5.2 représente un tel agent muni d'un contrôleur à 4 neurones.

Chacun des neurones artificiels évolue selon l'équation suivante :

$$\tau_i \frac{\delta y_i}{\delta t} = -y_i + \sum_j w_{j,i} z_j + I_i \quad (5.1)$$

où y_i représente l'activation du neurone i , τ_i est une constante de temps dans l'intervalle $[\exp 0, \exp 5]$, $w_{i,j}$ est le poids de la connexion entre le neurone i et le neurone j (intervalle $[-10, 10]$), z_j est un signal provenant du neurone j , et I_i est une entrée venant d'un capteur pour les neurones sensitifs. La simulation informatique nécessite de discrétiser l'équation différentielle 5.1. Pour cela, nous utilisons une méthode d'Euler dans laquelle le changement

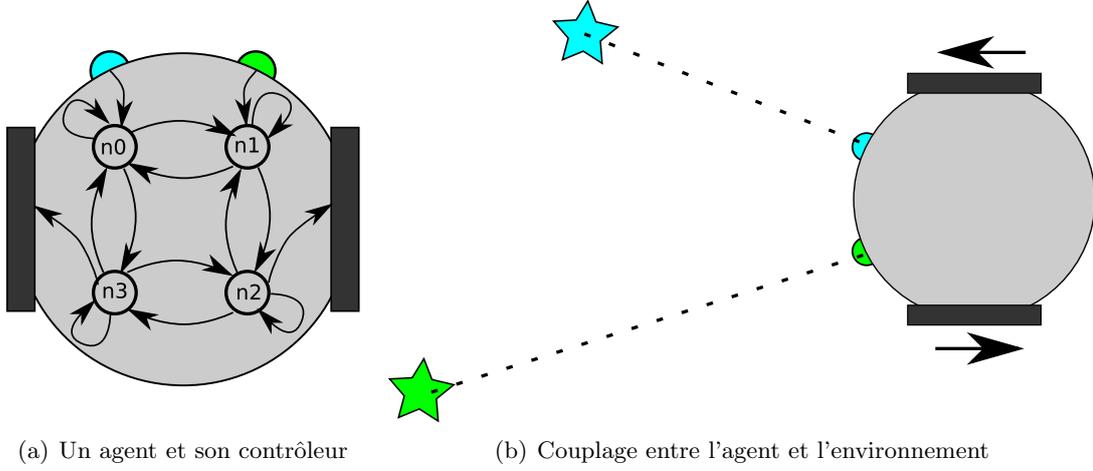


FIG. 5.2 – (a) : contrôleur réalisé par un CTRNN de 4 neurones. Les capteurs fournissent chacun une entrée à un neurone. La sortie de deux autres neurones est utilisée pour la commande des moteurs. (b) : couplage entre l'agent et l'environnement. Chaque capteur est spécifique à une balise, la valeur captée dépend inversement de la distance. La combinaison des commandes des deux moteurs produit le déplacement

d'état de y_i est approximé de la manière suivante :

$$y_i(t + d) = y_i(t) + \frac{d}{\tau_i} (-y_i(t) + \sum_j w_{j,i} z_j(t) + I_i) \quad (5.2)$$

Le même pas de temps est utilisé pour l'évolution de tous les neurones, dont l'évaluation se fait de manière synchrone.

Le signal z_i est calculé en appliquant une fonction de transfert sigmoïde à l'activation du neurone i de la manière suivante, où b_i est un biais pour le nœud i (intervalle $[-10,10]$) :

$$z_i = \frac{1}{1 + e^{-(y_i + b_i)}} \quad (5.3)$$

La valeur sur les capteurs est calculée par la formule 5.4, où $gain$ est un gain, identique pour les deux capteurs, dans l'intervalle $[0,40]$, d_i est la distance entre l'agent et la balise cible, et D est la portée du capteur :

$$I_i = gain * (1 - d_i/D) \quad (5.4)$$

L'activité d'un moteur est calculée en normalisant le signal provenant du neurone moteur dans l'intervalle $[-1,1]$ et en multipliant le résultat par un gain (intervalle $[0,40]$). Cette activité correspond à la vitesse tangentielle au point où est placé le moteur. Le déplacement de l'agent s'effectue en appliquant un mouvement de rotation ω donné par la différence entre les vitesses tangentielles des deux moteurs v_1 et v_2 , divisée par le diamètre de l'agent \varnothing :

$$\omega = \frac{v_1 - v_2}{\varnothing} \quad (5.5)$$

Un mouvement de translation est ensuite appliqué. Il correspond à la moyenne vectorielle des vitesses des moteurs. Pour introduire du bruit, les capteurs et les moteurs sont activés de manière asynchrone et aléatoire. Le bruit et la variabilité entre les essais sont des éléments importants pour obtenir des dynamiques intéressantes plus robustes (*cf.* chapitres 1 et 2).

5.1.3 Algorithme évolutionnaire

Les paramètres des réseaux de neurones artificiels sont recherchés en utilisant l'algorithme Microbial présenté dans la section 2.3.3. Ils sont directement encodés dans le génotype par des nombre réels. Un réseau de N neurones est défini par $2 + 2N + N^2$ paramètres :

- ▷ Un gain utilisé pour les deux capteurs
- ▷ Un gain utilisé pour les deux moteurs
- ▷ Un poids différent pour chaque connexion (N^2 poids)
- ▷ Un biais différent pour chaque neurone (N biais)
- ▷ Une constante de temps différente pour chaque neurone (N constante de temps)

La population utilisée est de 50 individus. Elle est initialisée aléatoirement. Microbial effectue une sélection par tournoi, 25 tournois sont effectués par génération entre deux agents tirés aléatoirement. Les individus sélectionnés sont conservés pour la génération suivante et pour produire après mutation l'autre moitié de la nouvelle population. Le taux de croisement est de $1/2$ comme le recommande Harvey (2009). Le taux de mutation est de $\frac{1}{4N+2+2N+N^2}$ pour chacun des paramètres¹. La mutation consiste à faire varier le paramètre de ± 0.05 . Ces paramètres sont toutefois maintenus dans les intervalles donnés ci-dessus.

5.1.4 Évolution incrémentale

L'évolution des individus se fait de manière incrémentale. Nous décrivons ici les différentes phases. Dans le processus d'évolution des agents, chaque phase s'appuie sur les populations évoluées dans les phases précédentes.

Phase 1 : Approche de la balise 1 à partir d'une position initiale fixe avec une orientation initiale aléatoire

Dans un premier temps, nous cherchons à obtenir des individus qui approchent de la balise 1. Au début d'un essai, les agents sont positionnés entre les deux balises. La balise 1 est en $(0,0)$, la balise 2 en $(50,0)$, les agents sont placés en $(25,0)$. Leur orientation par rapport aux balises est aléatoire. Le score d'un agent pour un essai est donné par la distance

¹ Le taux de mutation optimal selon Harvey (2001) pour permettre une exploration autour d'un optimum est le taux de mutation minimal entraînant un changement dans le phénotype.

minimale à la balise atteinte durant l'essai.

$$s_i = \frac{D - d}{D} \quad (5.6)$$

où D est la distance initiale à la balise 1 et d est la distance finale à cette même balise.

Les agents effectuent 10 essais par génération. Leur score final est une moyenne pondérée des scores obtenus sur les 10 essais :

$$S = \frac{2}{n(n+1)} \sum_{j=1}^n j * s_j \quad (5.7)$$

où j est le rang du score s_j (ils sont triés du plus fort au plus faible) et n le nombre total d'essais.

Nous avons fait évoluer des populations de 50 individus, munis de contrôleurs à 4, 5, 6 et 7 neurones. Les populations initiales sont constituées d'agents paramétrés de manière aléatoire. L'évolution s'arrête si le meilleur agent a un score supérieur à 0.9 et qu'il existe depuis au moins 5 générations, ou bien si la génération 10000 est atteinte. Pour chaque taille de CTRNN, nous avons alors simulé la population qui obtient le meilleur score en fin d'évolution. Le tableau 5.1 récapitule les résultats pour la simulation de ces populations pendant 10 essais.

		4 neurones	5 neurones	6 neurones	7 neurones
Première phase	Meilleur score	0.85	0.96	0.90	0.96
	Score moyen	0.6	0.85	0.59	0.70
	Nombre d'évolutions nécessaires	50	7	20	2

TAB. 5.1 – Évolution incrémentale : scores selon le nombre de neurones pour la première phase.

Phase 2 : Approche de la balise 1 à partir d'une position quelconque

La tâche des agents est légèrement complexifiée. Leur orientation par rapport aux balises est toujours aléatoire. La balise 1 est en (0,0), la balise 2 en (50,0), les agents sont placés en (25,y) avec $-25 < y < 25$. Leur score est calculé de la même manière qu'à l'étape précédente.

Nous avons fait évoluer des populations de 50 individus, munis de contrôleurs à 4, 5, 6 et 7 neurones. La population de départ est constituée à partir des agents obtenus à l'étape précédente. L'évolution s'arrête si le meilleur agent a un score supérieur à 0.95 et qu'il existe depuis au moins 5 générations, ou bien si la génération 10000 est atteinte. Pour chaque taille de CTRNN, nous avons alors simulé la population qui obtient le meilleur score en fin d'évolution. Le tableau 5.2 récapitule les résultats pour la simulation de ces populations pendant 10 essais.

		4 neurones	5 neurones	6 neurones	7 neurones
Seconde phase	Meilleur score	0.88	0.94	0.90	0.92
	Score moyen	0.57	0.61	0.71	0.63

TAB. 5.2 – Évolution incrémentale : scores selon le nombre de neurones pour réaliser la deuxième phase.

Phase 3 : Approche de la balise 1 puis de la balise 2 à partir d'une position quelconque

La tâche des agents est désormais d'atteindre la balise 2 après avoir atteint la balise 1. Les conditions initiales sont les mêmes que pendant la phase 2. La fonction d'évaluation est modifiée en conséquence et la durée d'un essai doublée. Le score d'un agent devient :

$$s_i = \begin{cases} 2 & : A \text{ puis } B \text{ atteintes} \\ 1 + \frac{D_2 - d_2}{D_2} & : A \text{ atteinte} \\ \frac{D_1 - d_1}{D_1} & : \text{sinon} \end{cases} \quad (5.8)$$

où D_1 est la distance initiale à la balise 1, D_2 est la distance à la balise 2 au moment où la balise 1 est atteinte, d_1 et d_2 sont les distances finales aux balises respectives.

Nous avons fait évoluer 10 populations de 50 individus, munis de contrôleurs à 4, 5, 6 et 7 neurones. La population de départ est constituée à partir des agents obtenus à l'étape précédente. L'évolution s'arrête si le meilleur agent a un score supérieur à 1.95 et qu'il existe depuis au moins 5 générations, ou bien si la génération 10000 est atteinte. Le tableau 5.3 récapitule les résultats.

		4 neurones	5 neurones	6 neurones	7 neurones
Troisième phase	Meilleur score	1.02	1.07	1.07	1.16
	Score moyen	0.78	0.72	0.86	0.8

TAB. 5.3 – Évolution incrémentale : scores par phase selon le nombre de neurones pour la troisième phase. Le score optimal théorique est 2.

La figure 5.3 représente les trajectoires du meilleur agent obtenu, lors de 10 essais. Ces trajectoires ont été lissées pour la lisibilité de la figure. Les trajectoires brutes sont présentées en annexe B ainsi que celles correspondant à des simulations dans des conditions différentes. Au début d'un essai, l'agent est placé entre les deux balises sur un segment matérialisé par des pointillés sur la figure en $(25, y)$ avec $-25 < y < 25$. Son orientation de départ est aléatoire. Les croix bleues représentent la position des balises. Globalement, l'agent approche de la balise 1, il la contourne et change brutalement de direction.

Pour vérifier que les agents sont bien sensibles aux deux balises, nous simulons les individus sélectionnés par l'algorithme génétique en modifiant les conditions initiales des

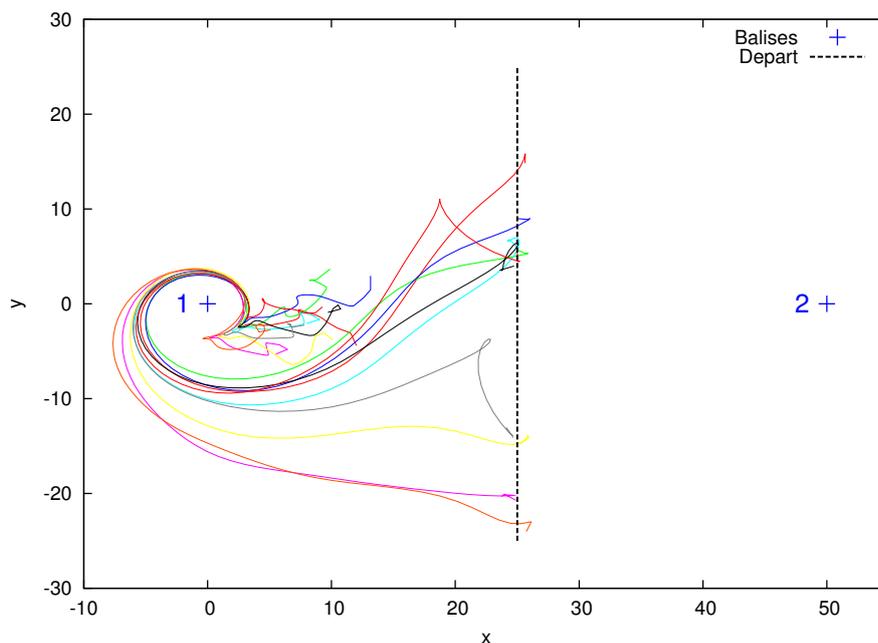


FIG. 5.3 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3.

essais. Au début d’un essai, les agents sont positionnés en $(-25, y)$ avec $-25 < y < 25$. Les conditions initiales par rapport à la balise 1 sont donc les mêmes que dans la simulation précédente (il y a une symétrie des conditions initiales par rapport à la balise 1, les distances de départ sont donc équivalentes), seules celles concernant la balise 2 sont modifiées. Leur orientation par rapport aux balises est aléatoire. Les trajectoires lissées correspondant à ces essais sont représentées sur la figure 5.4.

La différence dans le comportement de l’agent que ces modifications entraînent montre que l’agent est sensible à la balise 2 au moins au début d’un essai, il semble alors qu’il se “repère” dans un environnement en deux dimensions caractérisé par deux points. En effet, quand seules les conditions par rapport à la balise 2 sont modifiées, l’agent n’approche plus de la balise 1. Son comportement dépend donc également de ce qu’il capte en provenance de la balise 2. Cependant, même quand il est placé entre les deux balises, l’agent ne semble pas vraiment “se diriger” vers la balise 2 après avoir atteint la balise 1 (cf. annexe B). Ces résultats sont tout de même intéressants, puisqu’ils nous laissent envisager la possibilité d’obtenir des comportements reflétant une capacité à se localiser dans l’espace.

Néanmoins, notre objectif ici est avant tout de creuser la question de l’évolution ontogénétique, ces expériences ne constituant que de simples préliminaires à l’apprentissage. Les difficultés que nous avons rencontrées pour obtenir ces résultats nous amène aux réflexions suivantes :

1. Le contrôleur. Les CTRNN ne peuvent peut-être pas reproduire toutes les dynamiques. Nous rappelons toutefois qu’il a été démontré la capacité des CTRNN à approximer tout système dynamique (Funahashi et Nakamura, 1993). Ceci écarte théoriquement la

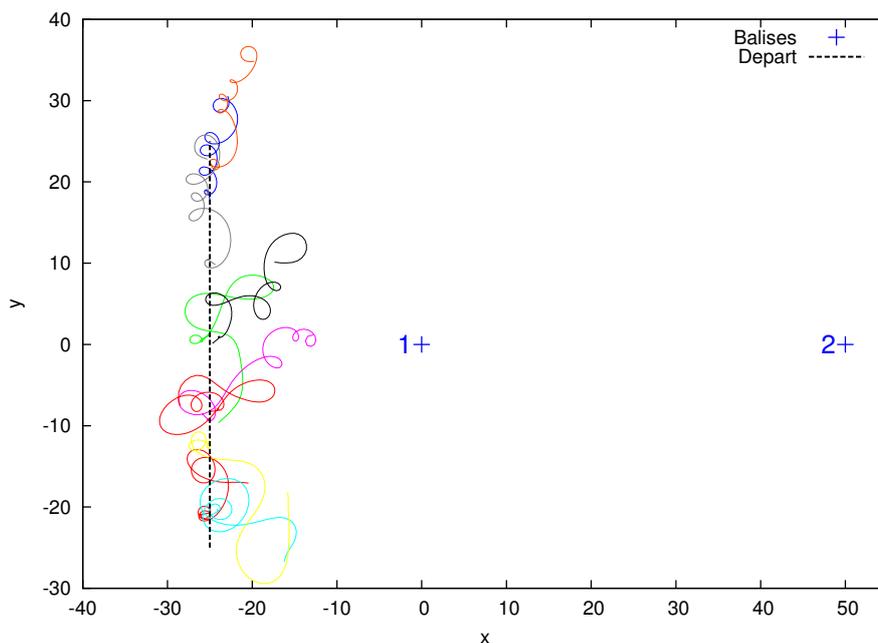


FIG. 5.4 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3 en modifiant uniquement les conditions initiales liées à la balise 2.

remise en cause du contrôleur.

2. L'algorithme évolutionnaire Il y a $2 + 2n + n^2$ paramètres à déterminer pour un réseau de n neurones. Ces paramètres sont des valeurs réelles pour lesquelles nous ne pouvons que choisir un intervalle de manière empirique. L'espace des phases est donc vaste et le hasard joue un rôle important dans la survenue d'une solution. L'évaluation des agents nécessite de simuler leur comportement, ce qui implique un besoin de ressources de calcul. à titre d'exemple, pour cette expérience, il faut environ 1 seconde pour simuler 3 générations. Nous avons fixé à 10000 générations la limite d'une tentative d'évolution et effectué jusqu'à 70 tentatives par taille de réseau. Nous n'avons pas ici décrit toutes les expériences qui nous ont permis de fixer nos paramètres. Ainsi, nous avons modifié les intervalles dans lesquels les biais, poids et gains sont recherchés. à notre connaissance, dans les travaux concernant l'évolution de CTRNN, ces intervalles sont déterminés de manière empirique à la suite d'une grande expérience dans le domaine, que nous ne possédions pas. Nous avons aussi fait varier la durée des essais et augmenté le nombre de générations autorisées avant l'abandon d'une tentative d'évolution. Le nombre d'essais a également été augmenté. Il est possible que pour certaines des expériences que nous avons considérées comme des échecs, d'autres tentatives d'évolution dans des conditions identiques mais avec une "graine" différente auraient abouties.
3. L'architecture de l'agent. L'évolution de comportements complexes est peut-être trop ambitieuse par rapport aux capacités des agents. Des neurones supplémentaires sont peut-être nécessaires. Les capacités sensibles et motrices de l'agent sont peut-être trop limitées. Ceci correspond au principe d'équilibre écologique² défendu par Pfeifer et

² Principle of ecological balance

Scheier (1999), selon lesquels pour réaliser un agent capable d'effectuer une certaine tâche dans un certain environnement, il faut que les systèmes sensitifs, moteurs et "neuraux" soient de complexité équilibrée et que de manière plus générale, la complexité de l'agent corresponde à la complexité de la tâche.

La tâche que nous avons définie ici est plus complexe qu'elle ne peut paraître de prime abord. En effet, il ne s'agit pas ici de se diriger vers une position directement dépendante de ce qui est capté, comme c'est le cas pour des comportements phototaxiques. Ici, la tâche fait appel à une capacité de localisation dans un espace en deux dimensions accessible à travers deux points de repère. L'agent n'a donc accès seulement qu'à l'évolution conjointe de ce qu'il capte (deux réels) et de ce qu'il fait (deux réels). Nos résultats semblent montrer que cette capacité de localisation est possible avec des CTRNN. Nous sommes cependant encore loin d'agents se déplaçant dans un environnement en deux dimensions selon des figures telles que nous l'envisagions en début de chapitre.

En tenant compte de ces problèmes, nous avons donc essayé de modifier l'expérience à maintes reprises et retenté des évolutions par algorithme génétique pour aller plus loin. Nous avons essayé de nombreuses fonctions d'évaluation et différentes manières de réaliser l'évolution incrémentale. Par exemple, en faisant évoluer un comportement d'approche de la balise 1, puis un comportement d'approche de la balise 2 en l'absence cette fois de la balise 1, avant de passer au comportement d'aller-retour entre les balises. Ces différentes tentatives se sont soldées par des échecs et ont nécessité beaucoup de temps. Devant ces difficultés à obtenir un simple aller-retour entre deux repères, il était alors inenvisageable de poursuivre vers la mise en place des comportements plus complexes que nous nous fixions comme objectif en début de chapitre : à savoir le déplacement selon des trajectoires correspondant à des figures en deux dimensions (carré ou sablier).

Ceci nous amène au constat suivant : En l'état actuel des techniques pour faire évoluer des populations d'agents modélisés selon une approche dynamique, deux choix s'offrent à nous, soit nous dirigeons plus fortement la conception du contrôleur en introduisant des hypothèses sur les solutions, soit nous adaptons la complexité de la tâche et la complexité de l'agent.

La première solution peut être envisagée en mettant au point des contraintes sur la topologie, c'est ce qui est utilisé dans l'expérience de Beer et Gallagher (1992) décrite dans la section 3.1.1. Elle pourrait également consister à définir des modules réalisant des sous-tâches, comme l'ont fait Yamauchi et Beer (1994a) pour l'expérience que nous décrivions dans la section 3.2.1. De manière plus radicale, nous pourrions renoncer au CTRNN pour choisir un contrôleur dont la dynamique est moins complexe et qui serait donc plus facile à paramétrer. Néanmoins, ceci ne nous convient pas car nous voulons préserver l'aspect dynamique autonome perturbé du contrôleur, capable de produire un comportement à la fois robuste et "créatif".

La deuxième solution peut consister soit à augmenter la complexité de l'agent, c'est-à-dire à le doter de plus de capteurs ou d'effecteurs ou à augmenter la taille du contrôleur, soit à réduire la complexité de la tâche. Notre intérêt se porte sur la possibilité de développer de nouveaux invariants sensorimoteurs sur la base d'anciens. Nous choisissons donc de proposer des tâches minimalistes permettant cela. La capacité de localisation dans un espace nous semble cependant exemplaire en matière d'invariant sensorimoteur. Nous conserverons donc

l'idée que le comportement cible doit refléter cette capacité.

Cette difficulté inattendue nous encourage finalement dans notre approche, qui vise à minimiser la part phylogénétique des capacités comportementales d'un agent pour favoriser un développement à l'échelle ontogénétique de son répertoire comportemental. En effet, puisque pour respecter une certaine autonomie dans la production du comportement, nous sommes contraints de nous intéresser à la mise en place de comportements minimalistes, il faut en contrepartie étudier la possibilité d'une complexification du comportement non pas à l'échelle de la conception des agents, mais à l'échelle ontogénétique, celle de l'exécution.

5.2 Comportement sensorimoteur à travers un exemple minimaliste

Nous exposons ici l'étude d'un comportement sensorimoteur très simple. Dans un environnement en une dimension, un agent doit approcher une cible. Cet agent est contrôlé par un CTRNN. Il a un capteur sensible à la distance de l'agent à la cible. Le signal perçu s est donné par :

$$s = g_s * \left(1 - \frac{d}{D}\right) \quad (5.9)$$

où g_s est le gain associé au capteur, d la distance à la cible et D la portée du capteur. Le signal est donc d'autant plus fort que l'agent est proche de la cible.

La vitesse v de l'agent est donnée par l'activité d'un moteur selon la formule :

$$v = g_m * ((2 * z_m) - 1) \quad (5.10)$$

où g_m est le gain associé au moteur et z_m est l'activité du neurone moteur.

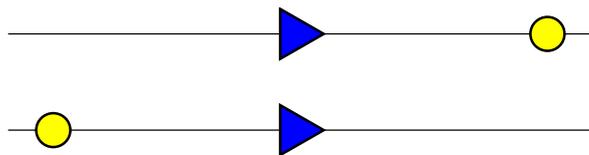


FIG. 5.5 – Configurations possibles de l'environnement. L'agent est représenté par un triangle. Il est orienté vers la droite. La cible, représentée par un disque, est soit placée devant lui, soit derrière lui.

Une cible fixe est placée dans l'environnement, soit devant l'agent, soit en arrière comme le montre la figure 5.5. L'algorithme génétique Microbial (Harvey, 2009) est utilisé pour faire évoluer une population d'agents qui ont pour tâche d'approcher la cible. Le score d'un agent pour un essai est donné par la distance finale à la cible qu'il atteint.

$$score = \max\left\{0, 1 - \frac{d_e}{d_i}\right\} \quad (5.11)$$

où d_i est la distance initiale entre l'agent et la cible et d_e est la distance finale qui les sépare. Si la distance finale est plus grande que celle de départ, le score de l'agent est nul. Un agent

est évalué sur 10 essais pendant lesquels la cible est positionnée initialement soit en $x_A + 10$ soit en $x_A - 10$, x_A étant la position de l'agent. Le score final d'un agent est le score minimal sur les 10 essais.

L'algorithme génétique a été appliqué à $70 * 5$ populations initiales de 50 agents contrôlés par des CTRNN de 1 à 5 neurones. Le tableau 5.4 présente les scores obtenus pour ces évolutions. Nous donnons ces chiffres à titre indicatif, car un même agent peut obtenir des scores différents pour 2 essais de même type. Ceci est dû au bruit présent sur les capteurs et à une grande sensibilité des contrôleurs à un changement, même léger, des conditions. Par exemple pour les contrôleurs à 3 neurones, 17 évolutions aboutissent à des scores positifs. Nous testons les meilleurs agents de ces 17 lignées, et pourtant, pour le test de 10 essais effectués, seulement 11 agents ont un score final non nul, le score final étant le score minimal sur les 10 essais. Toutefois, le score maximal sur les 10 essais est toujours très bon (5ème colonne du tableau 5.4) et les scores moyens sur les 10 essais sont corrects (colonne 6) sachant que les scores nuls sont inclus dans ces moyennes.

	Nombre d'agents testés	Nombre de scores non nuls	Meilleur score final	Score moyen final	Meilleur score pour un essai	Score moyen par essai
1 neurone	1	1	0.81	0.81	0.99	0.92
2 neurones	4	2	0.81	0.43	0.99	0.66
3 neurones	17	11	0.95	0.80	0.99	0.83
4 neurones	25	16	0.93	0.62	0.99	0.77
5 neurones	28	19	0.97	0.77	0.99	0.83

TAB. 5.4 – Statistiques pour l'évolution d'un comportement sensorimoteur minimal avec des CTRNN de 1 à 5 neurones (une ligne par type de contrôleur). La première colonne correspond au nombre d'évolutions aboutissant à un score positif au bout de 20000 générations sur les 70 évolutions lancées à nombre de neurones fixes. Les données suivantes correspondent aux résultats sur 10 essais pour les meilleurs agents de ces évolutions. La seconde colonne indique le nombre d'agents qui ont un score final non nul. La troisième colonne donne le meilleur score obtenu. Nous avons ensuite le score moyen final, calculé sans les scores moyens nuls. Le meilleur score pour un essai est ensuite présenté. Enfin, la dernière colonne correspond au score moyen par essai.

La difficulté de la tâche des agents est d'adopter un comportement moteur qui dépend de ce qui est capté. Une fois qu'une relation apparaît entre les variations de sensations et les variations d'actions, c'est-à-dire que le comportement relève de l'existence d'un invariant sensorimoteur, le score de la population augmente rapidement au fil des générations. Les résultats obtenus semblent montrer qu'avec seulement 1 ou 2 neurones les solutions soient très peu nombreuses. En augmentant le nombre de neurones, les scores obtenus sont meilleurs et le nombre d'évolutions aboutissant à un score positif est plus élevé.

La figure 5.6 montre la trajectoire du meilleur individu que nous avons obtenu, ainsi que la relation entre ses sensations et ses actions pendant deux essais types. Lorsque la cible est devant l'agent, celui-ci suit une trajectoire quasi linéaire (figure 5.6(a)). Par contre, dans

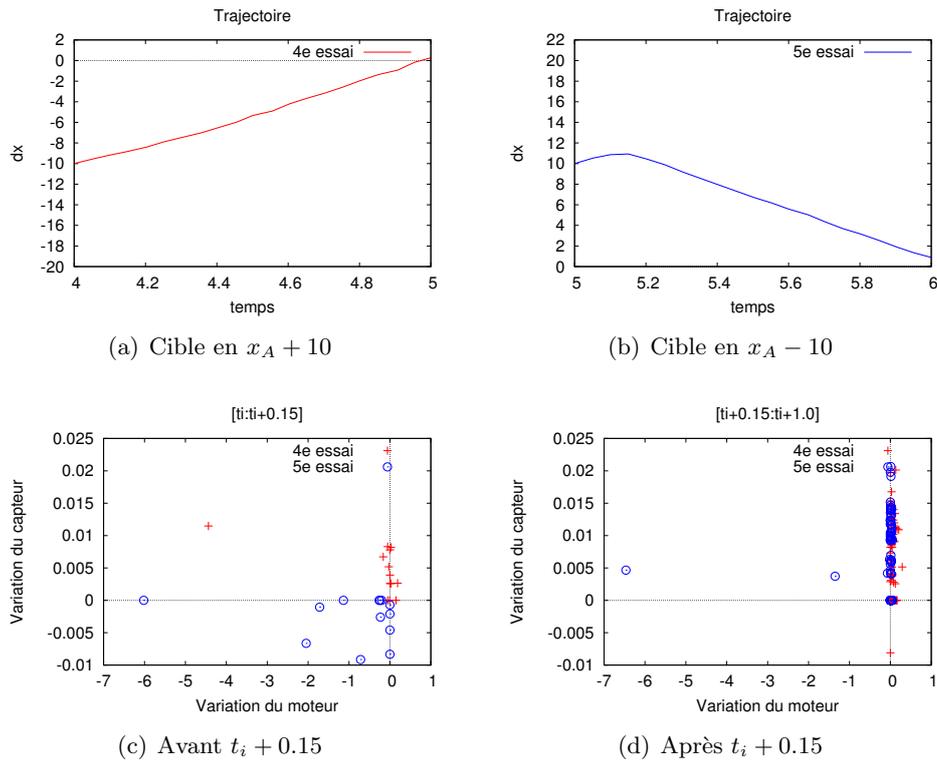


FIG. 5.6 – Analyse d’un individu. Les deux graphes du haut tracent la position de l’agent dans le repère de la cible en fonction du temps pendant un essai durant lequel la cible est en (a) $x_A + 10$ et pendant un essai durant lequel elle est en (b) $x_A - 10$. Les deux graphes du bas tracent la variation sur le capteur en fonction de la variation sur le moteur pour chaque pas de temps dans l’intervalle (c) $t_i; t_i + 0.15$ d’une part et (d) $[t_i + 0.15; t_i + 1]$ d’autre part.

le cas où la cible est placée derrière l’agent, deux phases apparaissent (figure 5.6(b)). Du temps $t=5$ à $t=5.15$, l’agent avance, puis il commence à reculer. Dans cet exemple, l’invariant sensorimoteur correspondant au comportement d’approche de la cible est la relation entre une faible variation de l’activité du moteur et une variation positive du signal capté, ce qui correspond à l’invariant identifié pour le comportement A2 dans la section 4.2. La figure 5.6(c) montre que cet invariant est respecté dès le début de l’essai lorsque la cible se trouve devant l’agent, alors que ce n’est pas le cas lorsqu’elle est placée derrière lui. En effet, nous pouvons ici identifier la relation sensorimotrice correspondant à une mauvaise direction, un passage rapide par le comportement A1, avant l’adoption du comportement A2. Cependant, dès le quinzième pas de temps, nous retrouvons cet invariant dans les deux cas, *cf.* figure 5.6(d). Il y a un changement d’activité du moteur guidé par la relation sensorimotrice.

Dans cette illustration, la découverte de l’invariant sensorimoteur a été faite à l’échelle phylogénétique pour les agents. En effet, c’est un algorithme génétique qui sélectionne les individus dont le comportement “inné” en quelque sorte respecte une certaine règle de correspondance entre actions et sensations. Notre objectif est, nous le rappelons, de faire acquérir de nouveaux invariants sensorimoteurs à des agents à une échelle ontogénétique, c’est-

à-dire à travers une adaptation “en ligne” résultant de leur confrontation à un environnement. Avant cela, nous proposons dans la section suivante de montrer comment cet invariant sensorimoteur peut être complexifié en conservant une approche phylogénétique.

5.2.1 Complexification de l’invariant sensorimoteur par une approche phylogénétique

L’expérience que nous allons maintenant présenter a pour objectif de montrer qu’il est possible de faire évoluer la population obtenue précédemment pour obtenir des agents qui adoptent un invariant sensorimoteur différent qui leur permet d’exhiber un comportement plus complexe. Cela correspond donc à l’approche incrémentale. L’architecture des agents est la même que celle décrite dans la section précédente.

5.2.1.1 Cible en mouvement

Dans un premier temps, les configurations possibles de l’environnement sont les mêmes que celles présentées dans la figure 5.5. La tâche est légèrement complexifiée en imposant un mouvement propre à la cible différent à chaque essai. De plus, les essais sont plus longs. La dimension temporelle est importante lorsque l’on utilise des CTRNN, car chaque neurone a une constante de temps spécifique. Il peut y avoir une périodicité dans le comportement de l’agent, qui a pour conséquence que le score de l’agent est dépendant de la durée de l’essai. Le système atteignant un point fixe, le comportement peut aussi basculer après un certain délai. Modifier la durée des essais peut permettre d’augmenter la robustesse du comportement.

Le score de l’agent pour un essai est donné par l’équation 5.11. Nous avons choisi d’utiliser pour créer nos populations initiales les meilleurs individus obtenus contrôlés par un CTRNN de 3 neurones artificiels. à ces individus s’ajoutent leurs copies mutées pour obtenir une population de 50 individus. L’algorithme Microbial a été appliqué 70 fois à cette population initiale. L’évolution s’arrête si un agent obtient un score supérieur à 0.95 alors qu’il a été sélectionné au moins 10 fois ou si la génération 20000 est atteinte.

Nombre d’agents testés	Nombre de scores non nuls	Meilleur score final	Score moyen final	Meilleur score intermédiaire	Score moyen intermédiaire
70	67	0.94	0.89	0.99	0.95

TAB. 5.5 – Résultats des évolutions obtenus en simulant 10 essais pour le meilleur agent obtenu à partir de chacune des 70 évolutions.

Les résultats de ces évolutions sont présentés dans le tableau 5.5. L’individu que nous avons retenu obtient un score de 0.94. Nous avons enregistré son activité sur 10 essais pendant lesquels la cible est soit en $x_A + 10$, soit en $x_A - 10$. Dans le repère que nous utilisons pour analyser le comportement de l’agent, la cible est fixe. Cependant, dans le repère de l’agent, la

cible est mobile. La vitesse de la cible est constante durant un essai, mais cette vitesse varie d'un essai à l'autre.

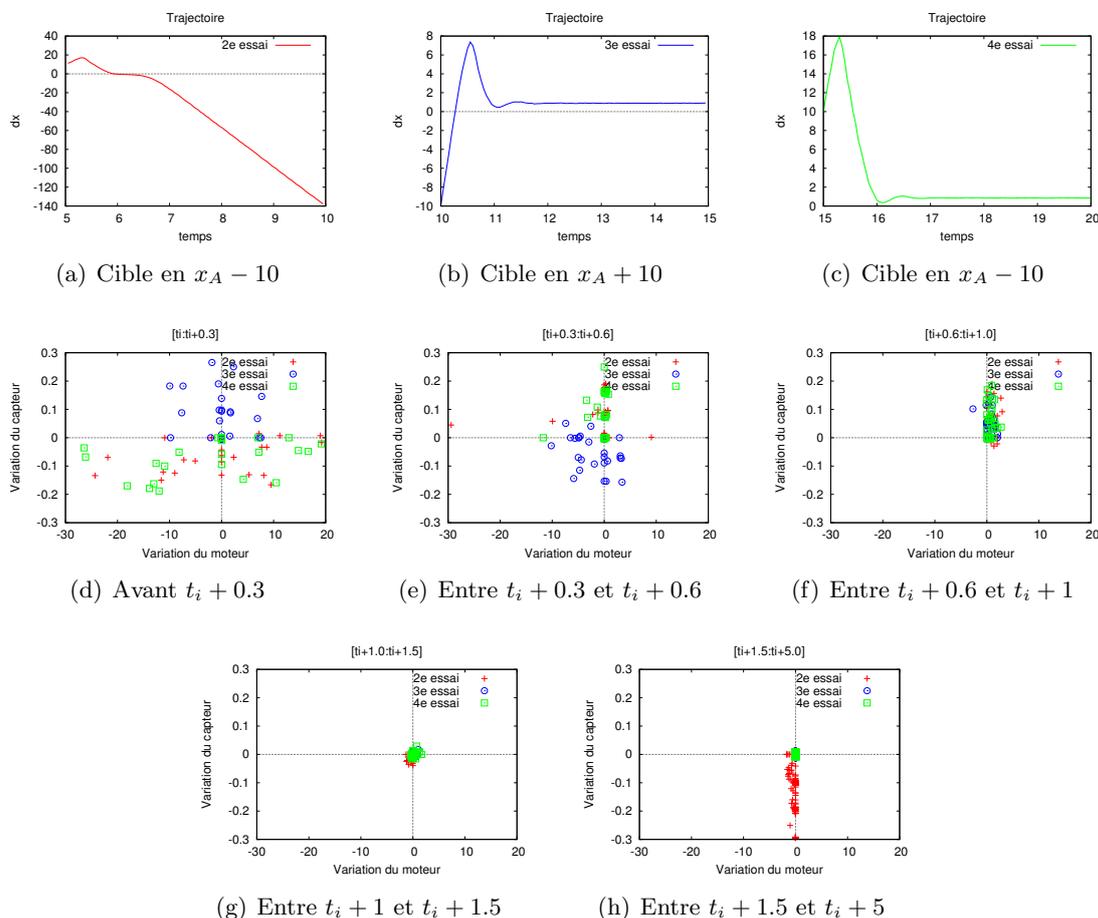


FIG. 5.7 – Analyse d'un individu. Les trois graphes du haut tracent la position de l'agent dans le repère de la cible en fonction du temps pendant un essai durant lequel la cible est soit en $x_A - 10$ (a) et (c), soit en $x_A + 10$ (b). Les graphes (d) à (h) tracent les variations sur le capteur en fonction des variations sur le moteur tous les 10 pas de temps à différentes étapes des trajectoires.

La figure 5.7 présente cette activité lors de trois essais que nous avons choisis parmi les 10 essais effectués :

- ▷ Essai 2 : La cible est en $x_A - 10$. L'agent commence par avancer, puis change de direction. Lorsqu'il atteint la cible, il reste à son niveau, mais s'en éloigne ensuite *cf.* figure 5.7(a).
- ▷ Essai 3 : La cible est en $x_A + 10$. L'agent avance dans la bonne direction, dépasse la cible, rectifie sa direction et reste jusqu'à la fin de l'essai au niveau de la cible *cf.* figure 5.7(b).
- ▷ Essai 4 : La cible est en $x_A - 10$. L'agent a le même comportement que lors de l'essai 2, mais une fois la cible atteinte, il reste à son niveau *cf.* figure 5.7(c).

La figure 5.7 présente aussi la correspondance entre variation de sensation et variation

d'action à différents moments des trois essais. Les intervalles de temps considérés correspondent aux différentes phases que nous avons décrites dans les trajectoires de l'agent :

- ▷ Avant $t_i + 0.3$, *cf.* figure 5.7(d) : Pendant les essais 2 et 4, la direction est mauvaise au contraire de l'essai 3. Ceci se retrouve dans les relations sensorimotrices avec des variations de sensations négatives et une activité motrice qui varie plus pour les essais 2 et 4.
- ▷ Entre $t_i + 0.3$ et $t_i + 0.6$, *cf.* figure 5.7(e) : La situation s'est inversée. Pour les essais 2 et 4, l'agent a adopté la bonne direction ce qui se traduit par des variations de sensation positive et que son activité motrice se stabilise. Au contraire, pour l'essai 3, l'agent s'éloigne de la cible et maintient une certaine variabilité dans ses actions.
- ▷ Entre $t_i + 0.6$ et $t_i + 1$, *cf.* figure 5.7(f) : Cette période correspond à une phase d'approche fine de la cible pour les 3 essais. L'agent se rapproche au maximum, les variations de sensation sont essentiellement positives. L'activité motrice est très stable.
- ▷ Entre $t_i + 1$ et $t_i + 1.5$, *cf.* figure 5.7(g) : L'agent adopte le mouvement de la cible, sensations et actions sont stables.
- ▷ Entre $t_i + 1.5$ et $t_i + 5$, *cf.* figure 5.7(h) : Pour les essais 3 et 4, il n'y a pas de changement. Par contre, lors de l'essai 2, l'agent "décroche" de la cible et n'ajuste plus ses actions, s'éloignant alors de la cible.

5.2.1.2 Distance initiale arbitraire

Dans un second temps, nous cherchons à raffiner le comportement des agents en variant les configurations possibles de l'environnement initial. En effet, elles sont légèrement modifiées puisque la cible est positionnée à une distance aléatoire de l'agent dans l'intervalle $[x_A - 10; x_A + 10]$, x_A étant la position de l'agent au début de l'essai. Nous prenons les meilleurs agents obtenus pour chaque tentative d'évolution pour construire nos populations initiales. Nous obtenons donc une population de 70 agents, contrôlés par un CTRNN de 3 neurones et munis d'un capteur et d'un moteur. Microbial est utilisé pour faire évoluer 70 fois cette population initiale. Les conditions d'arrêt de l'évolution sont toujours les mêmes : soit la génération 20000 est atteinte, soit un agent sélectionné au moins pour la dixième fois obtient un score supérieur à 0.95.

Nombre d'agents testés	Nombre de scores non nuls	Meilleur score final	Score moyen final	Meilleur score intermédiaire	Score moyen intermédiaire
70	70	0.90	0.80	0.99	0.91

TAB. 5.6 – Résultats des évolutions obtenus en simulant 10 essais pour le meilleur agent obtenu à partir de chacune des 70 évolutions.

Le tableau 5.6 présente les résultats de ces évolutions. L'individu que nous avons retenu obtient un score moyen de 0.90 sur 10 essais. Son comportement a été enregistré lors d'essais durant lesquels la cible est placée à différentes positions dans l'environnement. La figure 5.8 montre ainsi son activité pendant trois essais que nous avons retenu parmi les 10 essais :

- ▷ Essai 2 : La cible est en $x_A + 4.1$. L'agent part vers l'avant, dépasse largement la cible, puis change de direction et se stabilise autour de la cible, *cf.* figure 5.8(a).

- ▷ Essai 3 : La cible est en $x_A + 8.1$. L'agent a un comportement similaire à celui de l'essai 2, cf. figure 5.8(b).
- ▷ Essai 6 : La cible est en $x_A - 5.9$. L'agent part vers l'avant dans la mauvaise direction, change de sens et stabilise son mouvement dès qu'il atteint la cible, cf. figure 5.8(c).

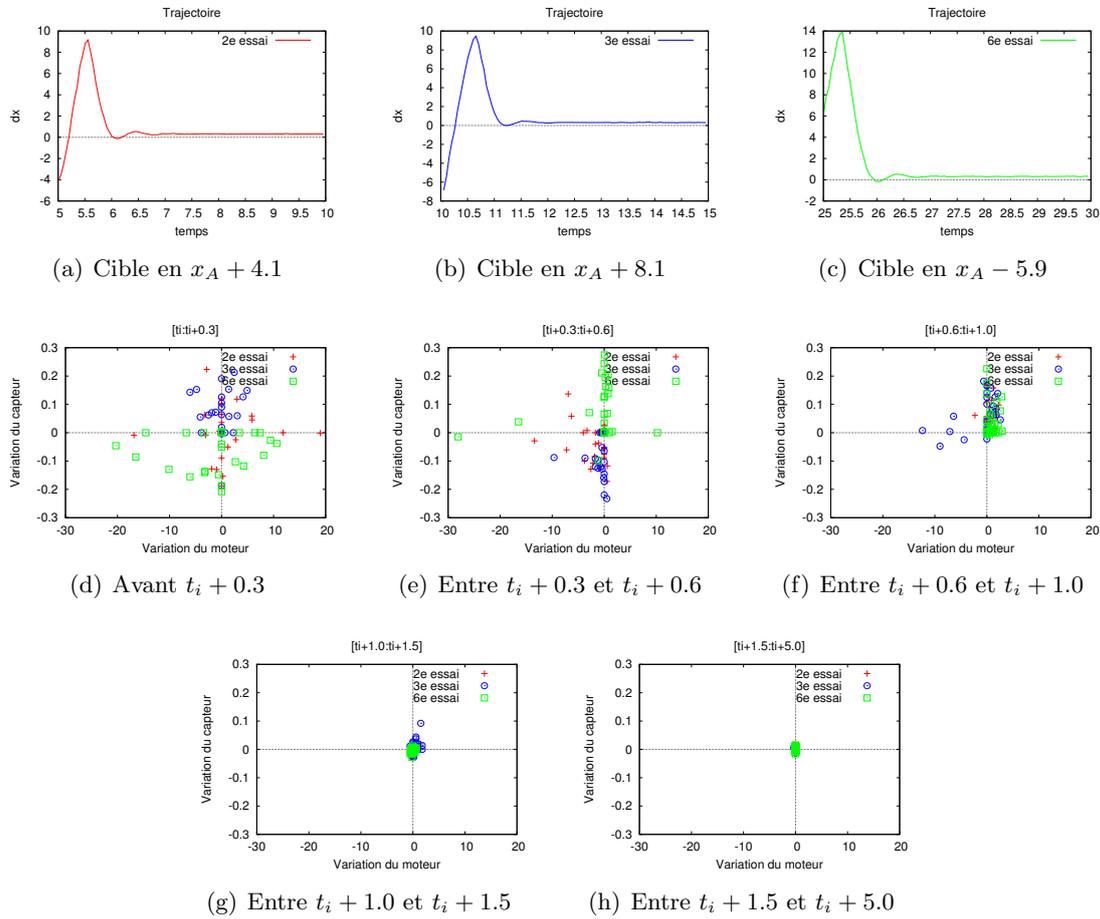


FIG. 5.8 – Analyse d'un individu. Les trois graphes du haut tracent la position de l'agent dans le repère de la cible en fonction du temps pendant un essai durant lequel la cible est soit en $x_A + 4.1$ (a), soit en $x_A + 8.1$ (b), soit enfin en $x_A - 5.9$ (c). Les graphes suivant tracent la variation sur le capteur en fonction de la variation sur le moteur pour chaque pas de temps dans divers intervalles.

La figure 5.8 présente aussi la correspondance entre variation de sensation et variation d'action à différents moments des trois essais. Comme précédemment, les intervalles de temps considérés correspondent aux différentes phases que nous avons décrites dans les trajectoires de l'agent :

- ▷ Avant $t_i + 0.3$ (figure 5.8(d)) : Pendant les essais 2 et 3, lors desquels l'agent part dans la bonne direction, il y a moins de variation dans l'activité motrice que pour l'essai 6.
- ▷ Entre $t_i + 0.3$ et $t_i + 0.6$ (figure 5.8(e)) : Pour l'essai 3, la direction est mauvaise (variations sensibles négatives), la correction se fait par un changement progressif du moteur. En effet, les variations sur le moteur sont uniquement négatives et d'intensité relativement faibles. Pour l'essai 2, il y a un changement de direction plus franc, l'agent adopte le comportement A1.

- ▷ Entre $t_i + 0.6$ et $t_i + 1$ (figure 5.8(f)) : Globalement sur les trois essais, l'activité est stabilisée, l'agent approche la cible.
- ▷ Entre $t_i + 1$ et $t_i + 1.5$ (figure 5.8(g)) : L'agent atteint la cible, sensations et actions sont stables.
- ▷ Entre $t_i + 1.5$ et $t_i + 5$ (figure 5.8(h)) : L'agent suit le mouvement de la cible de manière très réactive, les sensations et actions sont encore plus stables.

5.2.1.3 Maintien d'une distance fixe à la cible

L'étape suivante dans la complexification du comportement est de ne plus obtenir des agents qui suivent la cible simplement, ce qui correspond au comportement A dont nous parlions au chapitre précédent, mais des agents qui maintiennent une distance à cette cible, c'est-à-dire qui adoptent le comportement B. Ce comportement semble a priori plus complexe, mais repose sur un type d'invariant sensorimoteur assez proche.

Nous constituons une population initiale d'agents à partir des meilleurs agents des évolutions faites à l'étape précédente. Ce sont donc des agents qui possèdent un capteur sensible à la distance à la balise et un moteur qui leur permet de se déplacer en une dimension. Leur contrôleur est un CTRNN de 3 neurones. L'environnement est toujours le même : un environnement en une dimension dans lequel une balise mobile est placée à une distance initiale arbitraire de l'agent. L'évaluation des agents est légèrement modifiée pour correspondre à notre objectif. Ainsi, le score d'un agent pour un essai est donné par l'équation suivante :

$$score = \max\left\{0, 1 - \frac{|d_e - \Delta|}{|d_i - \Delta|}\right\} \quad (5.12)$$

où d_i est la distance initiale entre l'agent et la cible, d_e est la distance finale qui les sépare et $\Delta = 5$ est la distance à maintenir. Si la distance finale est plus grande que celle de départ, le score de l'agent est nul.

L'évolution a été faite en utilisant l'algorithme Microbial, lancé 70 fois avec la même population initiale composée de 70 agents obtenus à l'étape précédente pour l'approche incrémentale. Pour l'approche directe, l'algorithme Microbial est lancé 70 fois avec des populations initiales différentes de 70 agents paramétrés de manière aléatoire. L'évolution s'arrête si un agent de la population a un score supérieur à 0.95 et a survécu à 10 sélections ou bien si la génération 20000 est atteinte.

Le tableau 5.7 présente les scores obtenus. L'approche incrémentale est nettement meilleure. En particulier, les solutions sont beaucoup plus robustes. Un seul agent parmi 70 obtient un score final nul, alors qu'avec une évolution directe, 62 agents sur 70 ont un score final nul.

La figure 5.9 confirme l'intérêt de l'approche incrémentale en comparant les vitesses d'évolution. Les courbes représentent les évolutions les plus rapides pour chaque étape de l'approche incrémentale. Pour les étapes 1, 2 et 3, l'équation 5.11 est utilisée, pour l'étape 4 il s'agit de l'équation 5.12. Aucun des 70 essais effectués dans le cas de l'évolution directe n'a

	Nombre d'agents testés	Nombre de scores non nuls	Meilleur score final	Score moyen final	Meilleur score pour un essai	Score moyen par essai
Incrémental	70	69	0.96	0.89	0.99	0.96
Direct	70	8	0.88	0.67	0.99	0.85

TAB. 5.7 – Statistiques pour l'évolution du maintien d'une distance avec des CTRNN de 3 neurones. L'approche incrémentale (1ère ligne) est comparée avec une approche directe (2ème ligne). La première colonne correspond au nombre d'évolutions aboutissant à un score positif au bout de 20000 générations sur les 70 évolutions lancées. Les données suivantes correspondent aux résultats sur 10 essais pour les meilleurs agents de ces évolutions. La seconde colonne indique le nombre d'agents qui ont un score final non nul. La troisième colonne donne le meilleur score obtenu. Nous avons ensuite le score moyen final, calculé sans les scores moyens nuls. Le meilleur score pour un essai est ensuite présenté. Enfin, la dernière colonne correspond au score moyen par essai.

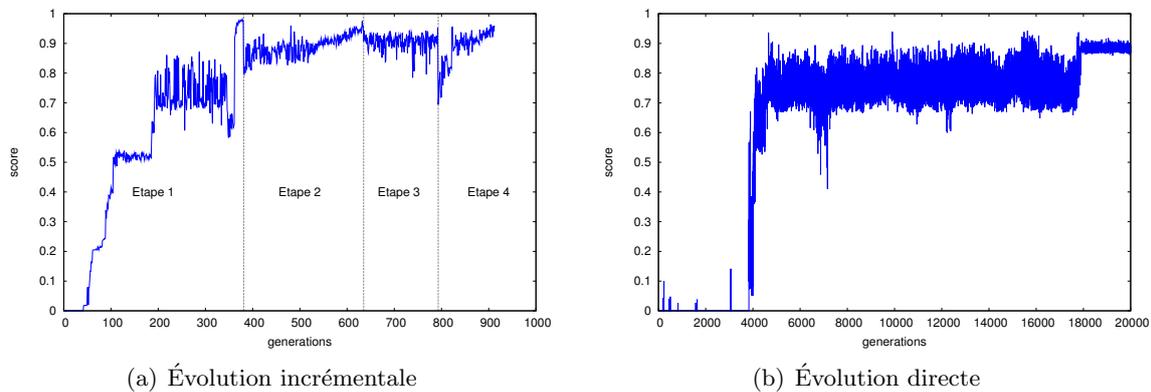


FIG. 5.9 – Évolution des meilleurs scores de chaque génération.

terminé avant la génération 20000, nous donnons la courbe d'évolution des scores du meilleur essai.

5.3 Discussion

Les expérimentations précédentes montrent que nous sommes proches d'une limite qui n'est pas évidente à caractériser pour obtenir des caractéristiques comportementales que nous avons identifiées comme préliminaires à l'obtention d'agents éactifs virtuels. Passer de deux à une dimension n'est pas gênant d'un point de vue théorique pour poursuivre nos recherches, mais attire notre attention sur les difficultés qui se présentent et nous ont amenés à effectuer quelques tests.

Ainsi, pour rendre compte de la sensibilité des CTRNN et de leur comportement

chaotique, nous étudions ici la dynamique de deux agents contrôlés par des CTRNN. Les agents simulés ici ont été obtenus à la suite des expériences présentées dans les section 5.2.1.1 et 5.2.1.2. La figure 5.7 nous montre que le comportement du premier agent n'est pas très robuste puisque lors d'un des essais il s'éloigne de la cible après l'avoir suivie pendant un moment. Les simulations du deuxième agent dont certaines correspondent aux essais présentés sur la figure 5.8 semblent indiquer un comportement cette fois plus robuste.

Nous procédons ici à quelques simulations de ces agents afin de montrer la complexité de la dynamique des CTRNN qui explique les difficultés à faire évoluer des agents réalisant les tâches que nous avons fixées.

Dans un environnement en une dimension, une cible est placée en $x_A - 10$. Pour l'agent, cette cible se déplace à une vitesse constante durant chaque essai. Nous cherchons à mettre en évidence l'influence de la variation de cette vitesse sur le comportement des agents. Pour cela, nous simulons les agents deux fois de suite avec la même vitesse. La seule différence entre ces deux essais consécutifs est dûe au bruit obtenu par le traitement asynchrone et aléatoire du capteur et du moteur, alors que l'évolution de l'activité des neurones est calculée de manière synchrone³. Nous faisons ensuite varier légèrement la vitesse pour les deux prochains essais et ainsi de suite.

La figure 5.10(a) présente les trajectoires réalisées par les agents pour des essais lors desquels la vitesse varie entre $v_x = -2.25$ et $v_x = -2.70$ par pas de 0.05. à partir d'une certaine vitesse, il semble que l'individu A ne réalise plus correctement la tâche, alors que c'est toujours le cas pour l'individu B. Pour identifier le moment où cette différence apparaît afin d'essayer de comparer les dynamiques des agents, nous réduisons alors le pas de variation de la vitesse entre les essais. Les figures 5.10(b) à 5.10(d) représentent les trajectoires obtenues pour des essais avec des pas de variation de la vitesse de plus en plus réduits.

Nous remarquons que vers $v_x = -2.5$, la vitesse de la cible semble être trop élevée pour que l'agent B réussisse la tâche de manière systématique. Avec des faibles pas de variation de la vitesse, nous pouvons constater que le comportement de l'agent B ne bascule cependant pas de manière radicale à partir d'une certaine vitesse. Des essais consécutifs avec la même vitesse ne produisent d'ailleurs pas forcément le même comportement, toutes conditions étant égales excepté l'ordre de traitement du capteur et du moteur. Ceci est particulièrement intéressant car de tels agents évalués dans ces conditions vont avoir des scores irréguliers. En retenant le score minimum sur plusieurs essais, il suffit alors d'un essai "raté" pour que l'algorithme génétique ne les élimine au profit d'agents plus réguliers mais peut-être moins performants.

La figure 5.11 représente l'activité du réseau selon la valeur du paramètre vitesse v_x . Les variables considérées sont l'activité z de chacun des neurones. Pour plus de lisibilité, le début des essais n'est pas pris en compte.

L'influence du paramètre v_x sur l'évolution de l'activité du réseau se fait de manière assez progressive, les attracteurs se situant sur une "courbe" dans l'espace des phases. L'activité est sensiblement similaire chez les deux agents. Il existe cependant une sorte de "cassure" quand v_x devient inférieur à -3 pour l'agent A et à -2 pour l'agent B, qui est illustrée par la figure

³ Sans ce bruit et dans des conditions identiques, tous les essais sont évidemment similaires.

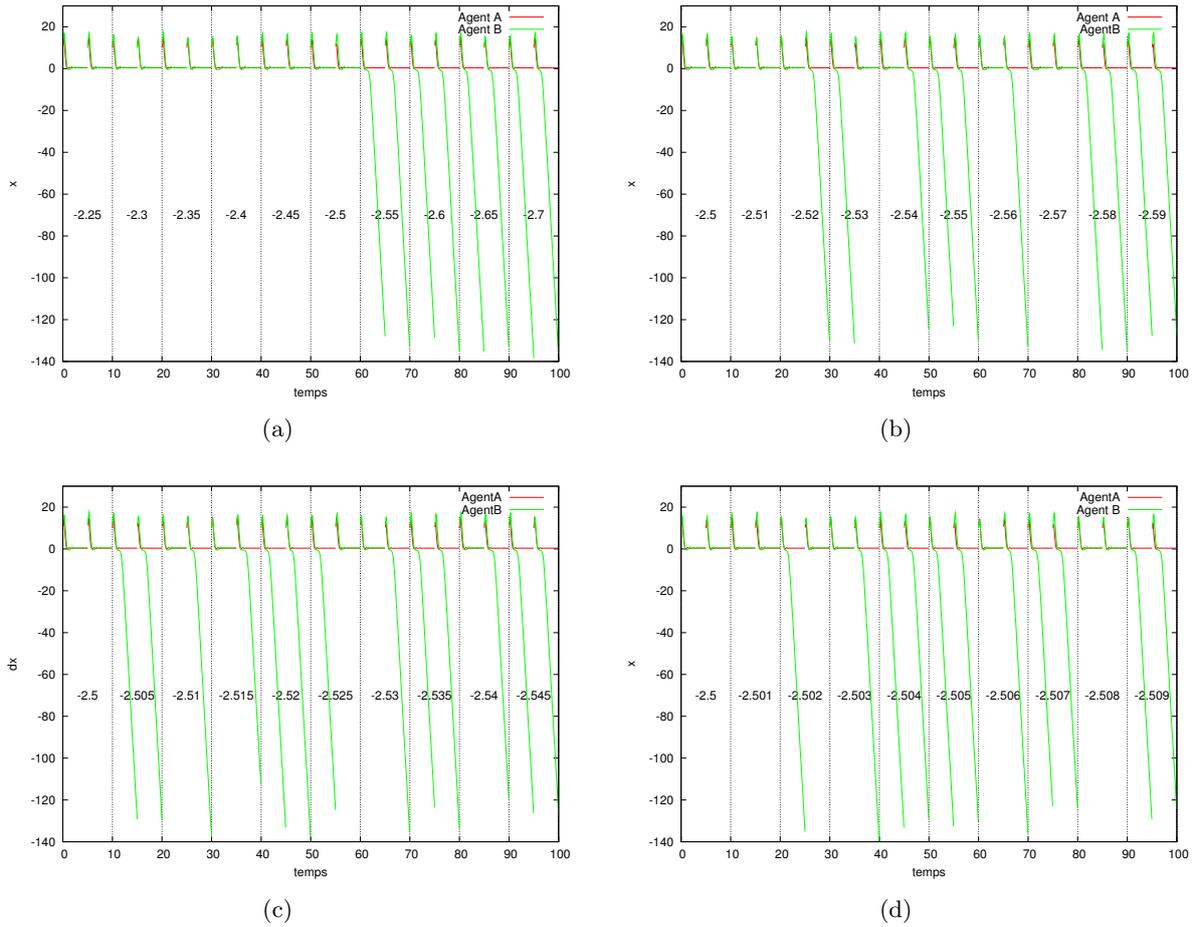


FIG. 5.10 – Trajectoires de deux agents pour des essais durant lesquels la cible est en $x_A - 10$. (a) : la vitesse varie entre $v_x = -2.25$ et $v_x = -2.70$ par pas de 0.05. (b) : la vitesse varie entre $v_x = -2.5$ et $v_x = -2.59$ par pas de 0.01. (c) : la vitesse varie entre $v_x = -2.5$ et $v_x = -2.545$ par pas de 0.005. (d) : la vitesse varie entre $v_x = -2.5$ et $v_x = -2.509$ par pas de 0.001.

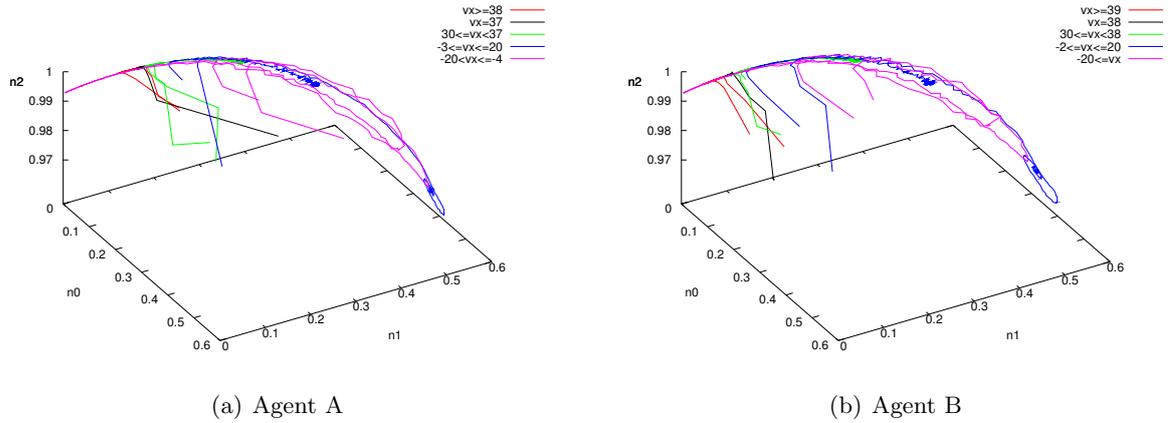


FIG. 5.11 – Évolution de l'activité du contrôleur des agents pour des essais durant lesquels la cible est en $x_A - 10$.

5.12 représentant les trajectoires dans l'espace des phases du contrôleur de l'agent B pour des essais lors desquels $v_x = -2.5$ et $v_x = -2.55$. Il s'agit d'essais pour lesquels la trajectoire de l'agent est représentée sur la figure 5.10(a). Ces deux tracés sont représentatifs du changement dans la dynamique du réseau qui conduit à un changement dans le comportement.

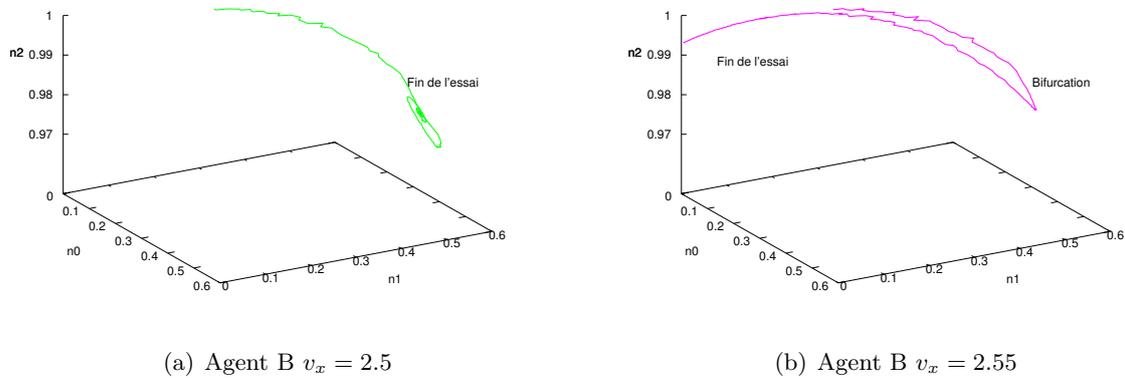


FIG. 5.12 – Activité du contrôleur de l'agent A pour des essais durant lesquels la cible est en $x_A - 10$.

Nous traçons sur la figure 5.13 l'activité du moteur de l'agent B pendant ces essais et ceux pendant laquelle v_x est proche de -2.5 . Cette activité nous indique que le changement de comportement n'est pas dû à une saturation du moteur, mais à un changement de direction. Pendant la première unité de temps, l'activité du moteur est quasiment identique, puis elle diverge totalement. Ceci se retrouve également dans l'activité du réseau comme le montre la figure 5.14.

L'observation des trajectoires comparées pour des vitesses différentes, *cf.* figure 5.15,

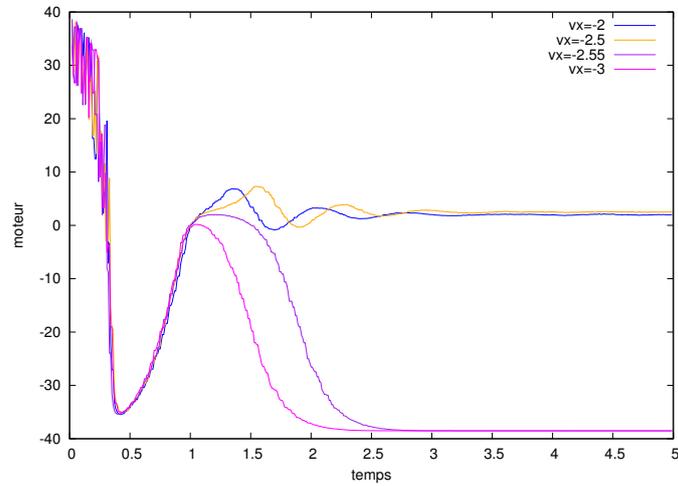


FIG. 5.13 – Activité du moteur de l’agent B pour des essais durant lesquels la cible est en $x_A - 10$.

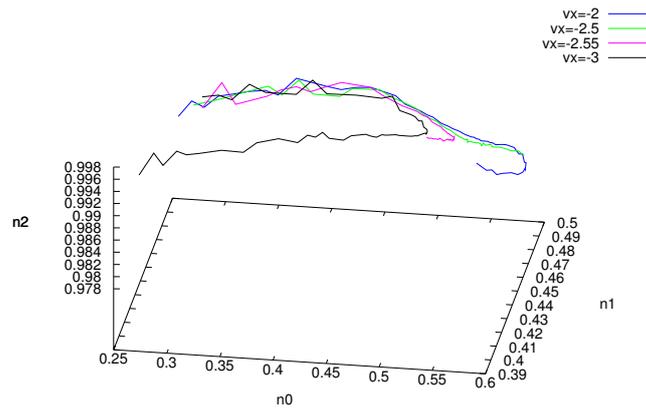


FIG. 5.14 – Activité du contrôleur de l’agent B pour des essais durant lesquels la cible est en $x_A - 10$ entre $t_0 + 0.8$ et $t_0 + 1.5$.

révèle de plus que le paramètre v_x a une influence sur la réussite de la tâche. En effet, la variation de ce paramètre implique un décalage de la distance à la cible à laquelle l'agent se stabilise. Pourtant, la tâche pour laquelle les agents A et B avaient été sélectionnés consistaient à atteindre une cible mobile et à se maintenir à son niveau. Nous avons ici donc un exemple d'évolution d'un comportement qui n'est pas exactement le comportement que nous voulions décrire avec la fonction d'évaluation, mais qui correspond à un bon score et se répand dans la population.

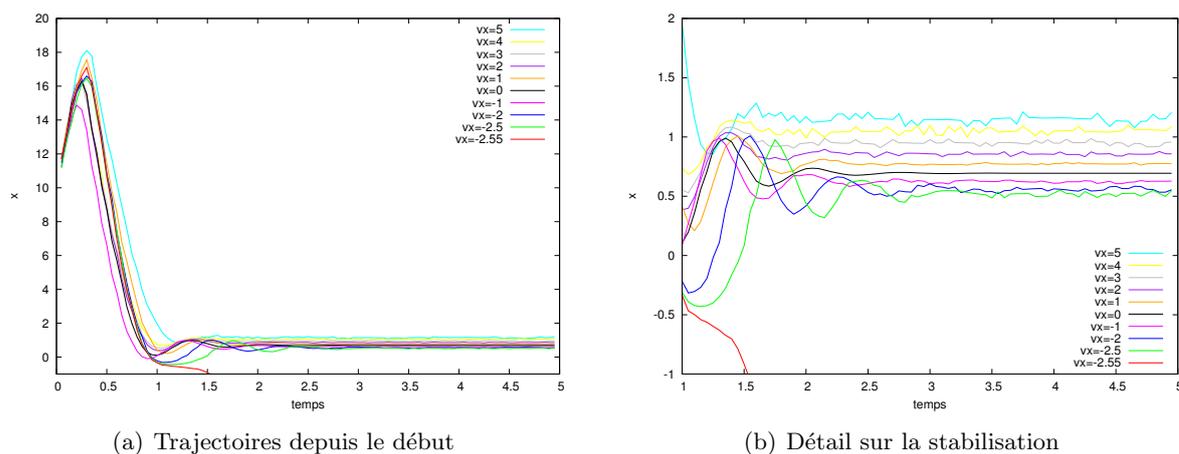


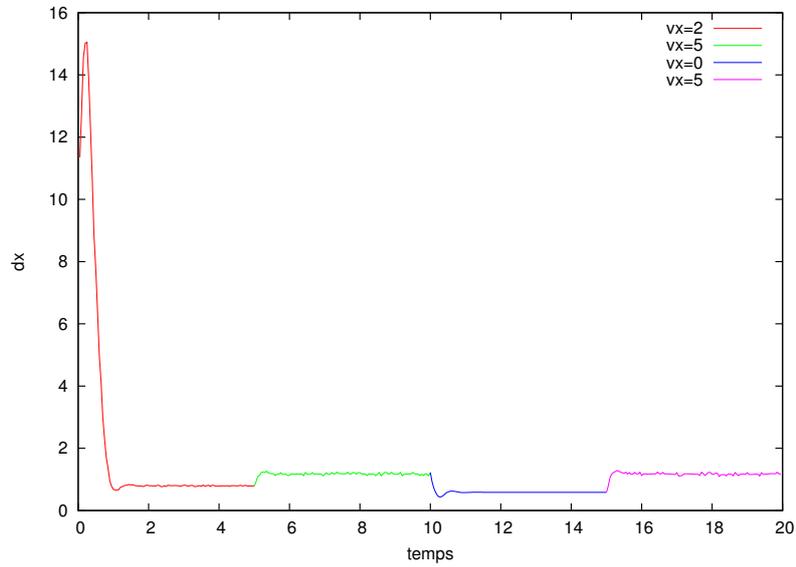
FIG. 5.15 – Trajectoires de l'agent B pour des essais durant lesquels la cible est en $x_A - 10$.

Ceci se confirme si nous faisons varier la vitesse en cours d'essai. Ainsi, la figure 5.16 a été produite à partir de la simulation de l'agent A dans ces conditions. La cible est positionnée initialement en $x_A - 10$ avec une vitesse $v_x = 2$. Après 5 unités de temps, le comportement de l'agent est stabilisé, la vitesse varie alors à $v_x = 5$. La figure 5.16(a) montre que l'agent se stabilise alors à une distance plus élevée de la cible. Quand la vitesse passe ensuite à $v_x = 0$, l'agent se place à une distance plus faible de la cible et revient exactement à la même distance quand v_x repasse à 5.

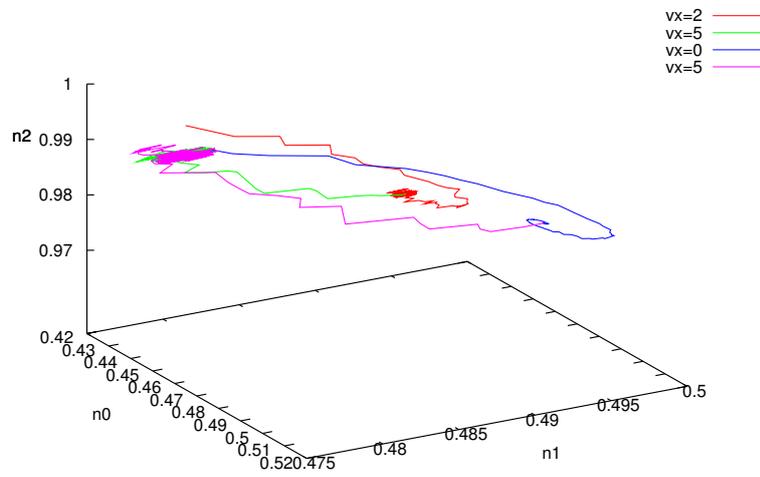
Des agents éactifs sont destinés à être capables d'évoluer et ils sont situés. Ceci implique que leur évaluation doit se faire dans des environnements dynamiques⁴. De nombreux paramètres vont varier. Néanmoins, la tâche n'est pas nécessairement liée à ces paramètres, comme c'est le cas ici dans la description de la tâche que nous donnons qui consiste à suivre une cible dont le mouvement est paramétré par v_x . En d'autres termes, la tâche doit pouvoir être décrite en faisant abstraction de certains paramètres, qui vont pourtant influencer son succès. Le comportement de l'agent doit alors être adapté aux différentes valeurs de ces paramètres, difficilement maîtrisables car pas forcément connus des expérimentateurs.

La question du score attribué se pose alors. Retenir la moyenne des scores n'est pas une solution adaptée car cela ne favorise pas les comportements stables et risque de conduire à

⁴ Le caractère dynamique de l'environnement était ici illustré de façon très simplifiée par les variations de vitesse provoquées dans les exemples.



(a) Trajectoire de l'agent



(b) Activité du réseau

FIG. 5.16 – Activité du contrôleur de l'agent A pour des essais durant lesquels la cible est en $x_A = 10$.

des solutions correspondant à des minima locaux, adaptées à certains types d'environnement c'est-à-dire à certaines valeurs des paramètres. Utiliser le score minimum comme nous l'avons fait ici permet de s'assurer qu'un agent qui obtient un score élevé est adapté aux variations du paramètre, mais ceci risque d'éliminer des agents qui sont dans l'ensemble performants mais qui ont été évalués au moins une fois dans des conditions pour lesquelles ils n'étaient pas adaptés. Un compromis consiste à réaliser une moyenne pondérée des scores en donnant plus d'importance aux essais les moins réussis. Cette méthode est décrite par Husbands et al. (1998), les scores sont triés du plus fort au plus faible, ce rang étant utilisé comme coefficient pour calculer la moyenne. Le score le moins bon a donc le coefficient le plus élevé, ce qui favorise les agents qui ont un comportement adapté de manière homogène à l'ensemble des conditions.

Néanmoins, si cette méthode peut réduire les biais, il est impossible avec une approche évolutionnaire de prouver qu'il n'en existe pas. Comme nous l'avons vu, les paramètres de l'environnement interviennent de manière imprévue sur le comportement. Nous avons montré que ce biais engendre une instabilité du comportement de l'agent lorsque l'on s'écarte des conditions de l'évolution. Or, nous cherchons à terme à "déformer" un invariant sensorimoteur en perturbant l'agent. Malgré une approche minimaliste, l'approche évolutionnaire ne peut pas nous permettre d'être assurés des relations sensorimotrices sur lesquelles repose le comportement de départ. Nous avons ici mis également en évidence la sensibilité des CTRNN, puisque dans des conditions identiques excepté un bruit réduit au minimum, le comportement d'un agent peut diverger fortement. Ceci ajoute une difficulté de taille à l'utilisation de l'approche évolutionnaire. Le bruit minimum que nous avons utilisé consiste à ne pas traiter de manière synchrone les capteurs et effecteurs de l'agent. Il s'agit donc d'une perturbation très faible du couplage entre l'agent et son environnement. Les changements spectaculaires que cette perturbation peut engendrer montrent à quel point les CTRNN sont sensibles à l'histoire de ce couplage. Si cette sensibilité est nécessaire pour permettre l'apprentissage à l'échelle ontogénétique que nous recherchons, elle reste un problème pour l'approche évolutionnaire.

5.4 Synthèse

Dans ce chapitre, nous avons simulé des agents adoptant des comportements sensorimoteurs différents. La première expérience, dans un environnement en deux dimensions montre qu'il est possible d'obtenir un comportement dépendant de deux repères dans l'espace. Elle est cependant mise de côté car développer ce type de comportements constitue un travail fastidieux que nous jugeons très intéressant mais qui diverge de notre objectif d'étude de l'évolution comportementale ontogénétique. Nous avons alors justifié notre choix d'aborder des comportements minimalistes. La difficulté de faire évoluer des agents éactifs nous encourage cependant dans notre approche, qui vise à minimiser la part phylogénétique des capacités comportementales d'un agent pour favoriser un développement à l'échelle ontogénétique de son répertoire comportemental.

Nous avons ensuite présenté une expérience dans laquelle nous simulons des agents adoptant un comportement simple qui consiste à se diriger vers une cible dans un environnement en une dimension. Une analyse qualitative du comportement nous a permis d'identifier des

invariants sensorimoteurs. à partir des agents obtenus, nous avons ensuite complexifié progressivement l'invariant sensorimoteur en utilisant une approche phylogénétique. Nous avons mis en relation les résultats obtenus avec l'analyse dynamique que nous avons proposée au chapitre 4. Ceci démontre la possibilité d'obtenir grâce à une approche phylogénétique et incrémentale des agents dont le comportement est lié à des invariants sensorimoteurs de plus en plus complexes.

Nous avons néanmoins constaté des difficultés sérieuses liées à la sensibilité des CTRNN qui est la conséquence de leur capacité à être autonomes. Cette sensibilité nous rend prudents quant à nos chances de succès et nous a amené à nuancer la sélectivité des algorithmes génétiques.

Notre objectif est d'aboutir à une complexification comportementale à l'échelle ontogénétique par un guidage. Pour cela, la prochaine étape de notre travail, qui est l'objet du chapitre 6, consiste à évaluer la possibilité d'influencer par un guidage le comportement d'un agent éactif virtuel.

Chapitre 6

Apprentissage associatif ontogénétique guidé

L' OBJECTIF de ce chapitre est d'évaluer la possibilité de modifier le comportement d'un agent éactif par l'utilisation d'un guidage. Pour cela, nous proposons une expérience dans laquelle la tâche des agents est d'apprendre une association entre des stimuli et une activité motrice. L'association n'est pas connue a priori puisqu'elle se révèle au travers du guidage. Nous commençons par décrire cette expérience dans la section 6.1, puis nous présentons les résultats obtenus dans la section 6.2.

6.1 Méthode

6.1.1 Tâche

Considérons une entité dotée de capteurs fonctionnellement équivalents à des “yeux”, un “nez” et des “oreilles”. Quand ses “oreilles” perçoivent un son aigu, elle tourne à droite. S'il s'agit d'un son grave, elle tourne à gauche. Ses “yeux” peuvent percevoir une lumière, son nez une odeur. Toutefois, ni la lumière, ni l'odeur ne provoquent un comportement particulier. Le son peut alors être utilisé comme moyen de guidage pour apprendre à cette entité une association entre un côté et la lumière ou l'odeur. Par exemple, nous voulons apprendre à l'entité à tourner à gauche quand la lumière est présente. Nous allumons cette lumière et si l'entité ne tourne pas à gauche, nous produisons un son grave. Nous supposons qu'une association peut se faire alors entre ce qui est senti (la lumière) et une activité (tourner à gauche) grâce à une influence extérieure (le son). Cet apprentissage par guidage se distingue d'un apprentissage par renforcement dans lequel une récompense binaire vient renforcer le comportement. Le guidage est, lui, un mécanisme progressif qui ne fait pas appel à la notion de récompense.

6.1.2 Agent

Un agent est modélisé comme un circuit de n neurones artificiels reliés à trois capteurs et à un moteur (voir figure 6.1). Deux capteurs de lumière, A et B, permettent de détecter la présence de deux sources de lumière. Ces capteurs sont binaires. Un capteur de guidage fournit un signal continu de guidage. Le moteur bi-directionnel de l'agent lui permet de tourner à gauche ou à droite.

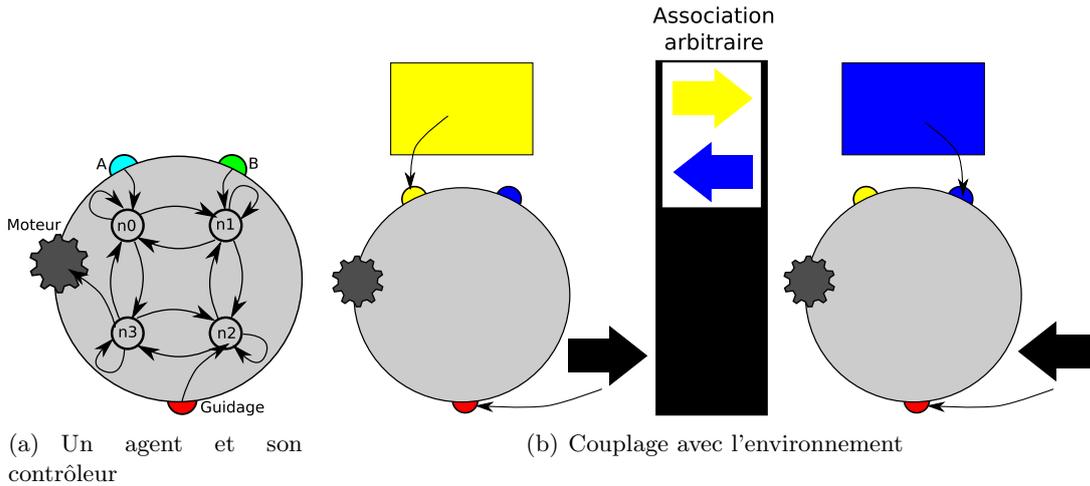


FIG. 6.1 – (a) Exemple d'architecture d'un agent contrôlé par un réseau de 4 neurones artificiels. Le réseau est très fortement récurrent : les nœuds sont tous inter-connectés avec des auto-connexions. Chaque capteur est relié à un nœud. Un des nœuds commande le moteur. (b) : couplage entre l'agent et l'environnement. Les capteurs A et B sont spécifiques à la détection d'un type de source. Les valeurs produites sont binaires. Le capteur de guidage reçoit un signal continu qui peut être positif ou négatif. Le moteur permet à l'agent de tourner dans une direction comme dans l'autre selon le signe de la commande motrice. Le guidage se fait selon une association arbitraire entre A/B et les directions, que nous représentons par les flèches colorées au centre de la figure.

La dynamique interne de l'agent est modélisée en utilisant un réseau de neurones artificiels à temps continu. Chacun des neurones artificiels évolue selon l'équation suivante :

$$\tau_i \frac{\delta y_i}{\delta t} = -y_i + \sum_j w_{j,i} z_j + I_i \quad (6.1)$$

où y_i représente l'activation du neurone i , τ_i est une constante de temps dans l'intervalle $[0.5, 2.1]$, $w_{i,j}$ est le poids de la connexion entre le neurone i et le neurone j (intervalle $[-8, 8]$), z_j est un signal provenant du neurone j , et I_i est une entrée venant d'un capteur pour les neurones sensitifs.

Les poids sur les connexions sont déterminés par des règles de plasticité de type hebbienne¹ :

$$\frac{\delta w_{i,j}}{\delta t} = \eta(\alpha y_i y_j + \beta y_j + \gamma y_i) \quad (6.2)$$

où η, α, β et γ sont des constantes dans l'intervalle $[-1, 1]$.

Le signal z_i est calculé en appliquant une fonction de transfert sigmoïde à l'activation du neurone i de la manière suivante :

$$z_i = \frac{1}{1 + e^{-(y_i + b_i)}} \quad (6.3)$$

où b_i est un biais pour le nœud i (intervalle $[-1, 1]$). L'activité du moteur est calculée en normalisant le signal provenant du neurone moteur dans l'intervalle $[-1, 1]$ et en multipliant le résultat par un gain (intervalle $[0, 40]$).

6.1.3 Algorithme évolutionnaire

Les paramètres des réseaux de neurones artificiels sont recherchés en utilisant un algorithme évolutionnaire. Ils sont directement encodés dans le génotype par des nombres réels. Ils représentent les constantes de temps, les biais, les gains et les paramètres des règles de plasticité. L'algorithme Microbial est utilisé sur des populations de 50 individus initialisés de manière aléatoire.

6.1.4 Évaluation des individus

L'évaluation des individus se fait en trois phases qui sont indépendantes.

Phase 1 : Évaluation de la réaction au guidage

Nous cherchons à obtenir des individus qui tournent à droite quand le signal de guidage est positif et à gauche quand le signal de guidage est négatif. Cette phase comporte 20 essais. à chaque essai, un signal de guidage est produit en l'absence d'autres stimuli. La valeur du signal est tirée aléatoirement dans l'ensemble $-1, 1$. La durée d'un essai varie aléatoirement entre 1,5 et 4,5. Le score d'un agent pour un essai est donné par le temps pendant lequel il tourne du côté correspondant au signal. Le score minimum obtenu sur les 20 essais est retenu comme score pour cette phase. Le calcul de ce score se fait selon l'équation suivante :

$$f_1 = \min_{i \in [0, 20]} \left\{ val_i | val_i = \frac{\int_{ts_i}^{te_i} \Delta_{Turn}(t) dt}{te_i - ts_i} \left(1 - \frac{\int_{ts_i}^{te_i} \Delta_{Energy}(t) dt}{te_i - ts_i} \right) \right\} \quad (6.4)$$

¹ Au moment où nous avons conçu cette expérience, nous n'avions pas encore établi la liste de propriétés internes que nous donnons dans le chapitre 1 comme nécessaires chez un agent éactif virtuel. Nous avons déjà choisi le CTRNN comme modèle de contrôleur. Nous nous sommes inspirés du travail de Wood et Di Paolo (2007) pour mettre au point cette expérience. Nous avons réalisés nos premiers essais en utilisant des CTRNN avec topologie fixe ou plastique. Les premiers résultats étaient alors plus concluants avec la plasticité et nous avons donc gardé ce mécanisme pour cette expérience. Nous revenons sur ce point dans la conclusion du chapitre.

où val_i est le score de l'essai i commençant au temps ts_i et finissant au temps te_i . Δ_{Turn} correspond à la corrélation entre le signal de guidage et l'activité du moteur, calculée ainsi :

$$\Delta_{Turn}(t) = \begin{cases} 1 - \|M(t) - S(t)\| & : M(t)S(t) > 0 \\ 0 & : else \end{cases} \quad (6.5)$$

où $M(t)$ est l'activité du neurone moteur ramenée dans l'intervalle $[-1,1]$ et $S(t)$ le signal de guidage. Δ_{Energy} est utilisée pour éviter les oscillations sur la sortie du moteur. Elle est calculée de la manière suivante :

$$\Delta_{Energy}(t) = \begin{cases} 0 & : M(t-1)M(t) > 0 \\ 1 & : else \end{cases} \quad (6.6)$$

Cette composante est particulièrement utile au début de l'évolution pour éliminer des solutions qui ignorent la valeur du signal de guidage et font simplement osciller la sortie du moteur. Ce comportement assure en effet un score aux alentours de 0,5 mais correspond à un minimum local qui ne nous intéresse pas. Nous retrouvons ici la nécessité d'appliquer une pression sélective adaptée à la complexité de la tâche que nous évoquions dans le chapitre 1.

Nous avons commencé par faire évoluer des populations d'agents contrôlés par des CTRNN de 4 à 7 neurones avec une topologie fixe ou plastique. Rapidement, nous avons fait le choix de nous concentrer sur des CTRNN de 6 neurones avec des connexions plastiques, qui semblaient donner de meilleurs résultats. Après avoir ajusté les paramètres de l'expérience et la fonction d'évaluation, nous n'avons donc lancé l'algorithme évolutionnaire qu'avec des CTRNN plastiques de 6 neurones. Notre choix ultérieur de privilégier les CTRNN sans aménagements, nous a ensuite conduits à réaliser des évolutions pour des CTRNN sans plasticité. Par manque de temps, nous n'avons cependant pas effectué les évolutions pour les CTRNN plastiques de taille différente. Le tableau 6.1 indique les scores obtenus.

		Meilleur score	Score moyen
Plastiques	6 neurones	0.91	0.75
Fixes	4 neurones	0.95	0.47
	5 neurones	0.96	0.51
	6 neurones	0.95	0.54
	7 neurones	0.95	0.63

TAB. 6.1 – Scores selon la taille et le type de réseau pour réaliser la première phase.

Phase 2 : Évaluation de l'apprentissage de deux associations

L'objectif est d'obtenir des agents capables d'apprendre une association entre deux stimuli. Cette phase comporte 50 essais qui correspondent chacun à la présentation d'une lumière. Nous décidons arbitrairement à chaque génération d'une association entre chaque lumière et un côté. La même association est utilisée pendant les 50 essais pour évaluer le comportement et produire le signal de guidage si nécessaire. L'ensemble des 50 essais constitue

un exercice d'apprentissage de l'association. La structure d'un exercice est décrite par la figure 6.2.

La durée d'un essai varie aléatoirement entre 1.5 et 4.5 unités de temps. Pendant un essai, une seule lumière est présentée. Les 10 premières présentations servent à stabiliser le comportement de l'agent. Le guidage commence ensuite : un signal de guidage est émis si l'agent ne tourne pas du côté associé pour cet exercice à la lumière présentée. Ce guidage n'est pas immédiat : il n'est produit que si l'agent ne tourne pas du bon côté depuis au moins 0.1 unité de temps, soit 10 pas de temps. Il est progressif, par pas de 0.1.

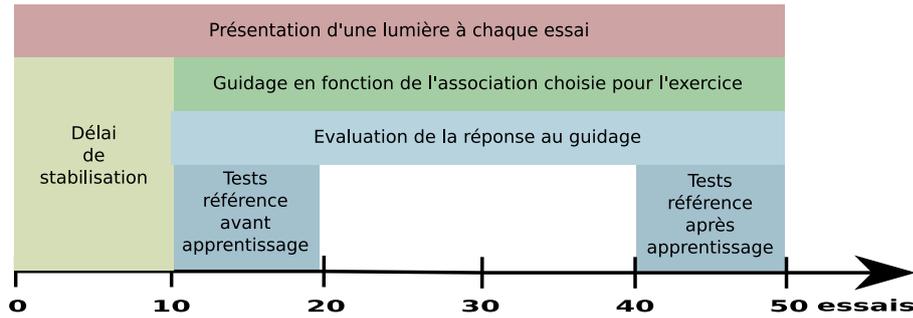


FIG. 6.2 – Structure d'un exercice d'apprentissage associatif. Les périodes bleues correspondent aux périodes pendant lesquelles l'agent est évalué.

Durant l'exercice, l'agent est donc guidé avec l'objectif qu'il fasse deux associations. Le score f_2 d'un agent pour l'exercice est donné par l'équation suivante :

$$f_2 = \min_{i \in [10.50]} \{val_i\} \times \min\{Score_A, Score_B\} \quad (6.7)$$

où val_i est un score pour l'essai i et $Score_A$ et $Score_B$ correspondent au progrès sur l'ensemble de l'exercice pour chacune des associations. A la manière de l'évaluation de la première phase, le score pour un essai est donné par le temps pendant lequel l'agent tourne du bon côté (Δ_{Turn2}) et dépend également de la stabilité de son comportement (Δ_{Energy}) :

$$val_i = \frac{\int_{ts_i}^{te_i} \Delta_{Turn2}(t) dt}{te_i - ts_i} \left(1 - \frac{\int_{ts_i}^{te_i} \Delta_{Energy}(t) dt}{te_i - ts_i} \right) \quad (6.8)$$

L'équation 6.6 permet de calculer Δ_{Energy} . Pour Δ_{Turn2} , le calcul est le suivant :

$$\Delta_{Turn2}(t) = \begin{cases} 1 & : M(t)C(t) > 0 \\ 0 & : \text{sinon} \end{cases} \quad (6.9)$$

dans laquelle $C(t)$ indique le côté associé à la lumière présentée. Δ_{Turn2} vaut donc 1 quand le comportement de l'agent correspond à l'association.

$$C(t) = \begin{cases} 1 & : \text{gauche} \\ -1 & : \text{droite} \end{cases} \quad (6.10)$$

Le progrès de l'agent $Score_X$ dans l'apprentissage de l'association d'un côté à la lumière x est calculé comme suit :

$$Score_X = \frac{Es_X - Ee_X}{Es_X}, X \in \{A, B\} \quad (6.11)$$

où Es_X et Ee_X sont respectivement le guidage moyen nécessaire au début de l'exercice et le guidage moyen nécessaire à la fin de l'exercice.

$$Es_X = \frac{\sum_{i=10}^{20} \frac{1}{te_i - ts_i} \int_{ts_i}^{te_i} \Delta_{Light}^X(t) dt}{\sum_{i=10}^{20} X(t)}, X \in \{A, B\} \quad (6.12)$$

$$Ee_X = \frac{\sum_{i=40}^{50} \frac{1}{te_i - ts_i} \int_{ts_i}^{te_i} \Delta_{Light}^X(t) dt}{\sum_{i=40}^{50} X(t)}, X \in \{A, B\} \quad (6.13)$$

où $X(t)$ vaut 1 quand la lumière X est dans l'environnement, 0 sinon. $\Delta_{Light}^X(t)$ correspond à la valeur du signal de guidage si la lumière X est dans l'environnement et que le guidage est nécessaire, à zéro sinon, comme l'indique l'équation suivante :

$$\Delta_{Light}^X(t) = \begin{cases} S(t) & : C(t)M(t) < 0 \text{ et } X(t) = 1, X \in \{A, B\} \\ 0 & : \text{sinon} \end{cases} \quad (6.14)$$

Ceci permet de mesurer le guidage nécessaire pour chaque type de lumière.

Phase 3 : Évaluation de l'apprentissage d'une autre association

Pour s'assurer que l'association apprise se fait bien grâce au guidage, les agents sont également évalués sur leur capacité à associer lumière et côté différemment. Nous permutons donc les liens à faire entre lumière et côté. L'évaluation est effectuée exactement de la manière décrite plus haut. Ceci nous donne un score f_3 .

Entre chacune de ces phases, le réseau de neurones artificiels est réinitialisé. Le score final d'un agent est une combinaison des scores des différentes phases. Il est calculé de la manière suivante :

$$f = a * f_1 + b * \min(f_2, f_3) \quad (6.15)$$

avec $a = 0.6$ et $b = 0.4$.

Le tableau 6.2 indique les scores obtenus dans l'évolution d'agents contrôlés par des réseaux de neurones artificiels pour différentes tailles et types de réseaux. Aucune des évolutions pour des CTRNN à poids fixes n'a donné de résultats positifs pour l'apprentissage associatifs. Le meilleur agent contrôlé par un CTRNN plastique n'obtient pas un score très élevé, mais son évaluation semble tout de même indiquer que le guidage influence son comportement.

6.2 Analyses

6.2.1 Étude comportementale

Dans la suite, nous nous intéressons au meilleur agent que nous avons obtenu par évolution. Il s'agit d'un agent contrôlé par un CTRNN de 6 neurones avec des connexions

Taille du réseau	Plastiques	Fixes			
	6 neurones	4 neurones	5 neurones	6 neurones	7 neurones
Meilleur f	0.72	0.59	0.57	0.58	0.57
f moyen	0.58	0.57	0.54	0.57	0.56
Meilleur f_1	0.93	0.98	0.95	0.97	0.95
f_1 moyen	0.86	0.44	0.46	0.40	0.49
Meilleur $\min(f_2, f_3)$	0.56	0	0	0	0
$\min(f_2, f_3)$ moyen	0.15	0	0	0	0

TAB. 6.2 – Résultats des algorithmes génétiques pour des réseaux de 4 à 7 neurones.

plastiques. Nous avons simulé cet agent et enregistré son activité sous différentes conditions. Tout d’abord, le comportement de l’agent en réaction à un guidage en l’absence de lumière a été observé. Pour cela, nous avons simulé cet agent alors qu’un scénario de guidage arbitraire était appliqué. L’activité de l’agent pendant cet épisode est représentée par la figure 6.3. En début de simulation, une direction est choisie. Le guidage se fait selon cette direction pendant une durée aléatoire. La direction est ensuite inversée pour une autre durée, etc... La figure 6.3 montre que l’agent observé à une réaction appropriée au guidage. L’émission d’un signal de guidage modifie toujours le moteur de la manière attendue. La plupart du temps, cette modification perdure. Il arrive cependant que le moteur bascule aussitôt et qu’il soit nécessaire d’insister sur le guidage avant d’obtenir une stabilisation du moteur.

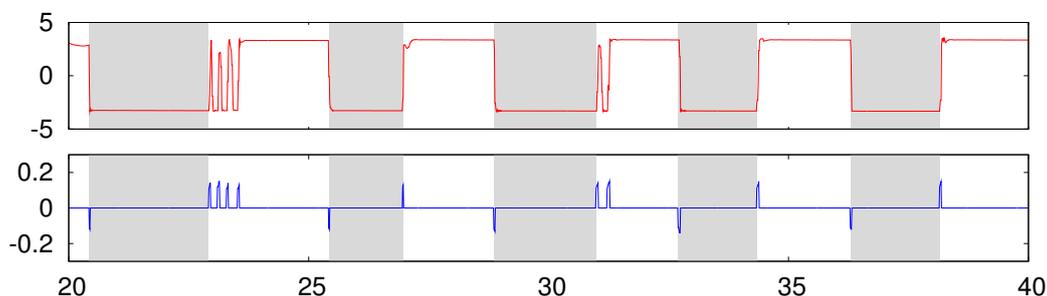


FIG. 6.3 – Le graphe du haut, en rouge, trace l’activité du moteur de l’agent. Le second graphe, en bleu, trace le signal de guidage. Les zones grisées sont les périodes durant lesquelles le guidage se fait vers la droite, les zones blanches vers la gauche.

Nous nous intéressons maintenant à la capacité de cet agent à apprendre une association. Pour cela, nous avons simulé l’agent en contrôlant l’environnement selon le protocole décrit à la figure 6.4. Après un délai, la phase A commence, lors de laquelle les lumières A et B sont introduites dans l’environnement alternativement pendant des durées aléatoires. Pendant cette phase A, aucun guidage n’est produit. Au temps $t = 40$, la phase B débute. La présentation de la lumière A alterne toujours avec celle de la lumière B pendant des durées aléatoires. Une association est fixée entre chaque type de lumière et un côté. Un signal de guidage est émis selon que le comportement de l’agent correspond ou non à cette association.

Pendant cette simulation, la tâche consiste à associer une activité motrice positive à la lumière A et une activité motrice négative à la lumière B. Lors de l’évolution, les agents

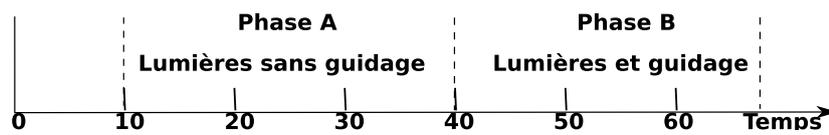


FIG. 6.4 – Protocole de test de la capacité à apprendre une association par guidage.

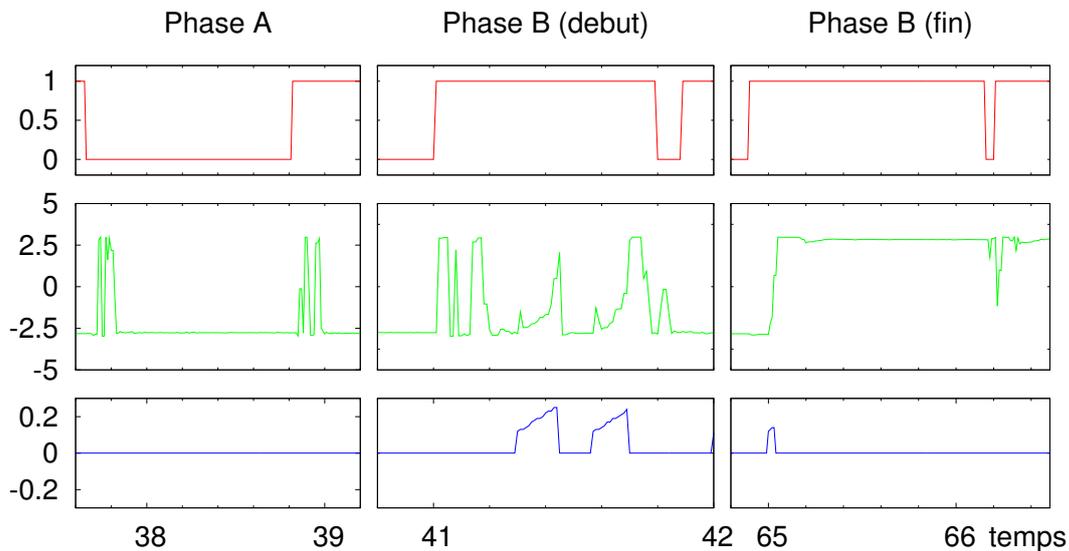


FIG. 6.5 – Simulation d'un apprentissage associatif par guidage. Les courbes représentent l'activité sensorimotrice de l'agent à la fin de la phase A, au début de la phase B et à la fin de la phase B. En rouge, en haut, l'activité du capteur A est tracée. La courbe du milieu, en vert, correspond à l'activité du moteur. L'activité du capteur de guidage est représentée par la courbe du bas en bleu.

ont été évalué pour cette association, mais aussi l'association inverse. L'agent que nous étudions ici obtient des scores satisfaisants pour les deux types d'association, ce qui montre la possibilité d'ontogenèse comportementale. La figure 6.5 représente l'activité sensorimotrice de l'agent à différentes périodes de cette simulation. Nous avons sélectionné des échantillons représentatifs de l'évolution de cette activité. Le premier échantillon, correspondant à la fin de la phase A, montre que des changements du type de lampe induisent des changements dans l'activité motrice. Cependant, il n'y a pas de corrélation directe entre l'activité motrice et les lumières. Nous remarquons aussi une tendance à avoir une activité motrice négative, c'est-à-dire vers la gauche. Le second échantillon est pris au début de la phase B. Ici, le guidage influence l'activité motrice, moyennant une certaine intensité. Il est aussi nécessaire de réitérer le guidage, car la tendance à aller vers la gauche est toujours présente. Le dernier échantillon correspond à la fin de la phase B. Le guidage n'est alors plus nécessaire puisque l'activité motrice est corrélée aux lumières. Il y a cependant une petite impulsion de guidage que nous relevons au changement de lumière. Il semble que le changement de lumière ne soit pas encore capté alors qu'il est déjà effectif pour le "guideur virtuel".

La notion d'ontogenèse comportementales implique une possibilité de généralisation de l'apprentissage. Nous venons de constater qu'il est possible d'apprendre deux associations

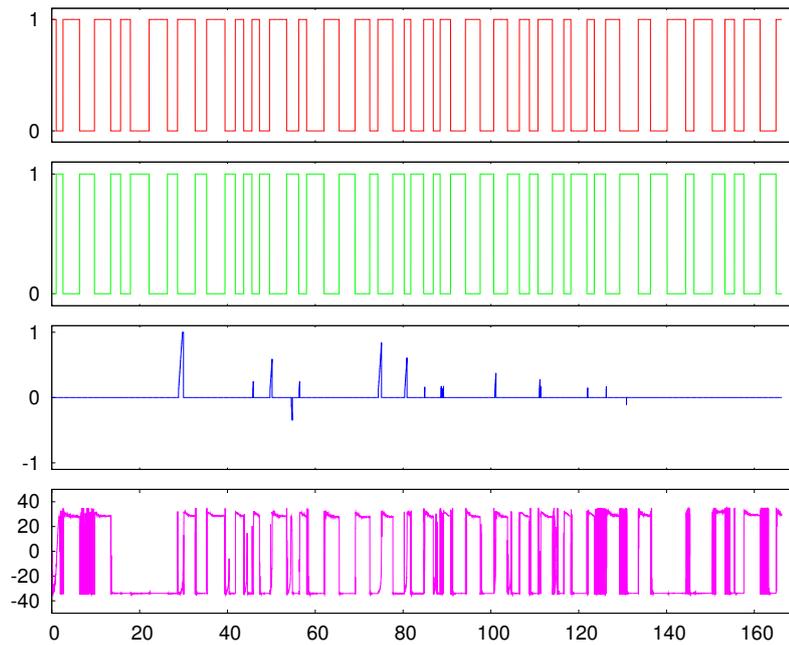
différentes par le même agent en fonction du guidage qu'il subit. Nous avons également souhaité tester si ceci mène à une telle possibilité de généralisation. Dans notre cas, cela revient à obtenir un agent capable de faire une association pour laquelle il n'a pas été évalué par l'algorithme évolutionnaire.

La figure 6.6 trace cette activité pour deux scénarios de guidage pour un même agent, d'une part l'association gauche/A droite/B et d'autre part l'association gauche/ni A ni B droite/A et B. Pour les deux associations, le guidage diminue avec le temps. Le même agent apprend les deux associations. En ce qui concerne l'association gauche/A droite/B sur laquelle l'agent a été évalué par l'algorithme génétique, l'association se fait relativement rapidement, *cf.* figure 6.6(a), mais n'est pas très stable une fois le guidage désactivé. Dans le cas de l'association gauche/ni A ni B droite/A et B, le guidage est beaucoup plus fort au début de l'apprentissage, *cf.* figure 6.6(b). L'activité du moteur est très instable. Néanmoins, l'association semble être faite puisque même après la désactivation du guidage, l'activité motrice est corrélée aux sensations selon l'association choisie. Pourtant, l'apprentissage de cette association n'a pas été un critère d'évaluation des agents par l'algorithme génétique.

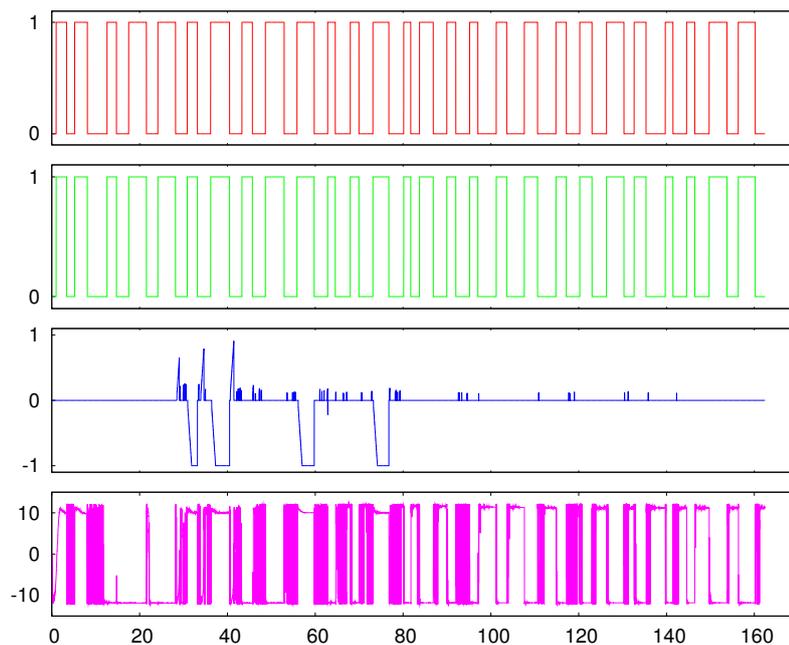
6.2.2 Étude dynamique

La simulation d'un agent nous permet également d'aborder son comportement lors de l'apprentissage en étudiant sa dynamique. Ainsi, nous représentons dans la figure 6.7 l'évolution de la relation entre l'activité motrice de l'agent et le guidage pendant l'apprentissage. Les échantillons considérés correspondent exclusivement à des périodes pendant lesquelles la lumière A est dans l'environnement. La lumière A est pour cet apprentissage à associer à droite, soit à une activité motrice positive. Cette représentation met en évidence l'existence de relations sensorimotrices.

Nous pouvons observer que pendant la phase A, soit avant l'apprentissage, l'activité du moteur se répartit entre des valeurs positives et des valeurs négatives (ensemble A1 en bleu). Au début de la phase B (ensemble B1 en orange), le guidage est très présent. Il y a une distribution de l'activité motrice sur un intervalle relativement large entre des valeurs positives et des valeurs négatives. Le guidage est distribué essentiellement entre 0.1 et 0.2. Nous rappelons que le guidage se fait au minimum avec une valeur de 0.1, puis avec des valeurs de plus en plus fortes s'il n'est pas efficace. Les motifs qui apparaissent indiquent que l'activité du moteur n'est pas immédiatement "corrigée" par le guidage. En effet, la présence de "serpentins" montre d'une part que le guidage augmente, donc qu'il est nécessaire plusieurs pas de temps de suite, d'autre part que l'activité du moteur n'est pas stable. Elle est a priori perturbée par le guidage. à la fin de la phase B (ensemble B2 en rouge), des points isolés apparaissent. L'activité motrice est majoritairement positive. Le guidage semble efficace puisque bien qu'il soit toujours présent, il est plus faible. Enfin, l'ensemble C, en vert, correspond à l'activité sensorimotrice pendant la première présentation de la lumière A après la phase de guidage. Nous retrouvons la relation sensorimotrice correspondant à l'association que nous cherchons à apprendre à l'agent, à savoir une activité motrice positive. Quelques valeurs négatives ont pourtant été relevées. Ici aussi, il peut s'agir d'un temps de latence nécessaire à la prise en compte du changement de situation de l'environnement par l'agent.



(a) Association gauche / A



(b) Association gauche / ni A ni B

FIG. 6.6 – Activité sensorimotrice d'un même agent guidé selon deux associations différentes. En rouge, en haut, l'activité du capteur A est tracée. La courbe verte correspond à l'activité du capteur B. L'activité du capteur de guidage est représentée par la courbe bleue. Enfin, la courbe rose en bas trace l'activité du moteur. Pour les deux types d'association, le guidage diminue alors que l'association se fait, ce qui se traduit par une corrélation en fin de simulation entre l'activité des capteurs et l'activité motrice.

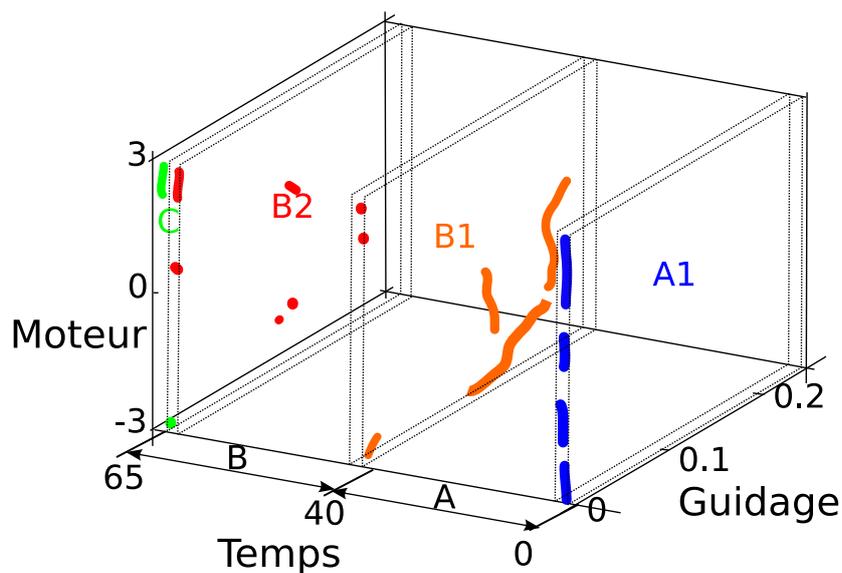


FIG. 6.7 – Dynamique sensorimotrice de l'agent. Activité du moteur en fonction de l'activité du capteur de guidage isolées à différentes périodes de la simulation.

De manière évidente, les valeurs positives dominent.

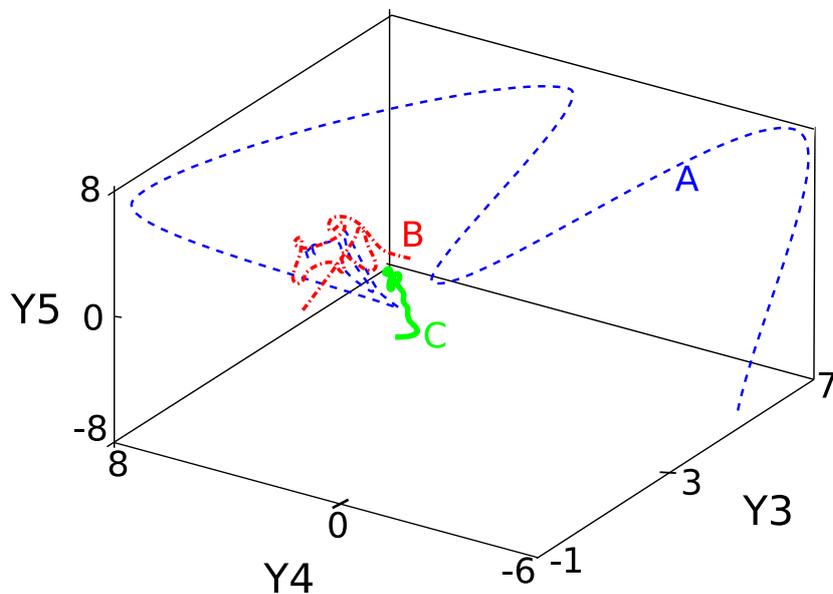


FIG. 6.8 – Dynamique interne de l'agent en fonction du guidage. Activité des neurones internes (Y4 et Y5) en fonction de l'activité du neurone lié au capteur de guidage (Y3).

En suivant l'évolution de la dynamique interne de l'agent, à travers l'activité des neurones, nous observons un réel changement. Ainsi, la figure 6.8 trace l'activité des neurones 4 et 5 en fonction de l'activité du neurone 3, qui reçoit le signal de guidage entre autres entrées. De la même manière que pour les précédentes représentations, nous avons échantillonné les données enregistrées et nous ne traçons cette activité que pour certaines périodes. Ainsi, la

courbe bleue a été construite à partir de données relevées pendant la phase A. La dynamique interne durant la phase B est tracée en rouge. Enfin, la courbe C, en vert, correspond à la présentation de la lumière A après la phase d'apprentissage, c'est-à-dire une fois que le guidage est désactivé.

6.3 Synthèse

Dans ce chapitre, nous avons présenté une expérience avec l'objectif de démontrer la possibilité d'apprendre un comportement par guidage à un agent énaactif. La tâche que nous donnons aux agents dans cette expérience est de réaliser une association entre un stimulus et une activité motrice. Nous utilisons une approche phylogénétique pour obtenir des agents capables de réaliser l'association à une échelle ontogénétique. Ainsi, nous proposons une méthode d'évaluation des agents qui permette de sélectionner ceux qui apprennent de leur interaction.

L'apprentissage au sens où nous l'entendons se traduit par une transformation de la dynamique sensorimotrice de l'agent. Nous avons montré, par une analyse de l'évolution de cette dynamique chez un agent, qu'il est possible de modifier le comportement de l'agent en interagissant avec celui-ci et que ceci se traduit par une modification de sa dynamique sensorimotrice de manière durable. En effet, la dynamique interne de l'agent est qualitativement différente, même après que le guidage ait été désactivé.

Nous relevons cependant quelques limites propres à l'expérience que nous avons menée. Ainsi, l'activité des agents est parfois interprétable de différentes manières, qui pourraient remettre en question la réalisation d'un apprentissage à l'échelle ontogénétique. Par exemple, nous remarquons que l'activité motrice est de manière "innée" influencée par les lumières. Comme nous l'avons vu précédemment une prédisposition au stimulus est toutefois nécessaire pour que le comportement se mette en place. De plus, il n'y a pas de véritable corrélation, mais une simple perturbation de l'activité motrice provoquée par la variation des lumières. La corrélation apparaît avec le guidage.

Ensuite, bien que le guidage diminue de manière évidente, des impulsions de guidage continuent à être émises par intermittence tout au long de l'épisode de guidage. Nous pensons que ce phénomène est principalement dû à un décalage de la prise en compte des changements dans l'environnement entre l'agent et le "guideur". En effet, pour introduire du bruit dans l'expérience, la mise à jour des capteurs et des moteurs se fait de manière asynchrone. De plus, un temps d'adaptation au changement peut être nécessaire à l'agent. Le "guideur" émet alors un signal alors que l'agent est déjà en train de mettre en place le comportement adapté. Ceci semble être confirmé par le comportement de l'agent après la désactivation du guidage.

L'hypothèse selon laquelle le comportement de l'agent est conditionné au changement de l'environnement plutôt qu'à une véritable association entre un côté et un type de sensations ne peut cependant pas être écartée. En particulier, nous remarquons la grande instabilité du comportement de l'agent lors de certains essais. Par exemple, lors de la simulation représentée sur la figure 6.6(a), après la désactivation du guidage, l'agent change de côté lorsque les

lumières changent, mais pas toujours selon l'association à apprendre. Il est envisageable que l'association n'ait pas été apprise, ce qui n'empêche pas de constater un changement du comportement de l'agent qui a minima associe un changement de sensation à un changement d'activité motrice. D'un autre côté, dans le cas de l'apprentissage de l'autre association, l'agent maintient la bonne corrélation, bien qu'il existe encore des oscillations sur le moteur, ce qui laisse à penser que l'association est réellement faite. Pour s'assurer que l'association soit réellement faite, cette expérience pourrait être étendue sans l'alternance systématique des lumières. La tâche serait alors d'associer le guidage à un type de lumière afin de conditionner le comportement à la présence ou l'absence de ce type de lumière.

En outre, la dimension sensorimotrice est trop limitée pour considérer que les agents artificiels que nous obtenons soient des agents situés dans un environnement et que leur couplage initie une dynamique dont il est possible d'extraire des invariants sensorimoteurs. Les sensations liées à la lumière sont binaires et surtout l'influence de l'activité de l'agent sur ses sensations est extrêmement pauvre. En l'absence du guidage, elle est totalement nulle. Pendant l'apprentissage, l'activité de l'agent a une influence sur le guidage. Néanmoins, elle n'en a aucune sur la manière dont sont captées les lumières A et B. L'invariant sensorimoteur ne peut exister que pendant l'apprentissage, seule période où il y ait une réelle interaction entre l'agent et l'environnement.

Enfin, dans cette expérience les contrôleurs utilisés ne sont pas de simples CTRNN puisqu'un mécanisme de plasticité a été utilisé. Ceci n'est pas un problème en soi, mais au regard de notre définition d'un agent éactif virtuel, ce mécanisme n'est théoriquement pas nécessaire. Au vu des scores moyens finaux présentés dans le tableau 6.2, la différence entre les populations est faible. Ceci nous laisse envisager qu'il n'y a pas de différence forte au niveau comportemental entre les agents selon le type de contrôleur et que les échecs des CTRNN à poids fixe (tout comme le relatif succès des CTRNN plastiques) sont dûs à la difficulté de trouver une solution par algorithme évolutionnaire dans ce cas. Afin d'éprouver notre choix du CTRNN, les expériences qui suivent n'utilisent donc pas cette plasticité.

Dans le chapitre 7, nous tentons de réaliser un apprentissage par guidage chez un agent éactif qui résulte en une véritable transformation d'un invariant sensorimoteur pour l'agent. Dans cette expérience, le couplage entre l'agent et son environnement devra permettre la mise en place d'une co-évolution par une interaction qui s'actualise dans le comportement à la fois de l'agent et de l'environnement. Ceci veut dire que l'agent sera situé dans son environnement, dans le sens où ses actions auront des conséquences sur ses sensations et que l'environnement aura également une influence sur le comportement de l'agent.

Apprentissage ontogénétique guidé d'invariants sensorimoteurs

NOUS avons, dans le chapitre 5, fait évoluer des agents pour qu'ils adoptent différents invariants sensorimoteurs. La transformation du comportement était alors obtenue en utilisant un algorithme évolutionnaire. Nous cherchons ici à réaliser une transformation du comportement à l'échelle d'un agent en nous appuyant sur un guidage. Dans le chapitre 6, nous avons vu que le guidage peut influencer le comportement de l'agent et amener à un apprentissage associatif. L'objectif est donc de valider cette influence du guidage sur des comportements véritablement sensorimoteurs. Le comportement étudié est très proche de ceux exposés au chapitre 5, à ceci près que le guidage intervient, tel qu'il a été introduit au chapitre 6. C'est pourquoi, les descriptions de la tâche et de l'agent que nous donnons dans la suite reprennent des éléments des chapitres précédents.

7.1 Méthode

7.1.1 Tâche

Les agents évoluent dans un environnement en une dimension. Leur comportement "inné" consiste à suivre une cible placée dans l'environnement. De plus, ils ont une réaction "innée" également à un guidage : l'intensité de ce guidage influence leur activité motrice. La tâche des agents est de modifier leur comportement par rapport à la cible, c'est-à-dire de passer du comportement de suivi à un comportement de maintien d'une distance. Cette distance est arbitraire et ne peut être connue en l'absence du guidage. La stratégie de guidage consiste à émettre un signal de plus en plus fort si l'agent s'écarte de la distance qu'il doit maintenir à la cible.

7.1.2 Agent

Un agent est doté d'un moteur qui lui permet de se déplacer sur une ligne en avant ou en arrière. Un capteur lui permet de détecter la cible. Un autre capteur reçoit le signal de guidage. Le contrôleur est un CTRNN dont certains neurones sont reliés au moteur et aux capteurs.

Chacun des neurones artificiels évolue selon l'équation suivante :

$$\tau_i \frac{\delta y_i}{\delta t} = -y_i + \sum_j w_{j,i} z_j + I_i \quad (7.1)$$

où y_i représente l'activation du neurone i , τ_i est une constante de temps dans l'intervalle $[\exp 0, \exp 5]$, $w_{i,j}$ est le poids de la connexion entre le neurone i et le neurone j (intervalle $[-10,10]$), z_j est un signal provenant du neurone j , et I_i est une entrée venant d'un capteur pour les neurones sensitifs.

Le signal z_i est calculé en appliquant une fonction de transfert sigmoïde à l'activation du neurone i de la manière suivante, où b_i est un biais pour le nœud i (intervalle $[-10,10]$) :

$$z_i = \frac{1}{1 + e^{-(y_i + b_i)}} \quad (7.2)$$

L'activité d'un moteur est calculée en normalisant le signal provenant du neurone moteur dans l'intervalle $[-1,1]$ et en multipliant le résultat par un gain (intervalle $[0,40]$). La vitesse de l'agent est donnée par cette activité.

La valeur sur le capteur de la cible est calculée par la formule, où le $gain_i$ est dans l'intervalle $[0,40]$, d_i est la distance entre l'agent et la cible et $D = 100$ est la portée du capteur :

$$IC_i(t) = gain_i * \left(1 - \frac{d_i(t)}{D}\right) \quad (7.3)$$

La valeur sur le capteur de guidage est calculée en multipliant un gain, dans l'intervalle $[0,40]$, au signal de guidage qui est dans l'intervalle $[0,1]$. Une position est choisie vers laquelle l'agent va être guidé. Le signal de guidage est émis si l'agent s'éloigne de cette position. Un accumulateur assouplit légèrement le guidage. En effet, le signal de guidage à l'instant t vaut :

$$IG_i(t) = accu(t) * \| d_i(t) - \Delta \| \quad (7.4)$$

Ceci signifie que l'intensité du guidage augmente progressivement si l'agent s'éloigne pendant plusieurs pas de temps.

$$accu(t) = \begin{cases} \max\{1, accu(t-1) + 0.01\} & : d_i(t) > d_i(t-1) \\ 0 & : sinon \end{cases} \quad (7.5)$$

7.1.3 Algorithme évolutionnaire

Les paramètres d'un agent sont déterminés en utilisant l'algorithme Microbial. Il s'agit des gains pour les capteurs et le moteur, ainsi que les biais et constantes de temps des neurones, de même que les poids des connexions.

7.1.3.1 Évolution du comportement "inné"

Dans un premier temps, l'algorithme évolutionnaire est utilisé pour obtenir des agents qui, en l'absence de guidage, se dirigent vers la cible et qui, en l'absence de la cible, réagissent au guidage. Pour cela, nous effectuons deux types d'essais. D'une part, nous simulons les agents dans un environnement du même type que celui décrit à la section 5.2.1 : une cible est placée dans un environnement en une dimension. Initialement, elle est placée en x_c avec $x_A - 11 < x_c < x_A - 8$ ou $x_A + 8 < x_c < x_A + 11$, x_A étant la position de l'agent. Dans le repère de l'agent, la cible est mobile avec une vitesse constante pendant chaque essai, mais qui varie d'un essai à l'autre.

Le score pour un essai est donné par l'équation suivante :

$$s_i = \max\left\{0, 1 - \frac{d_e}{d_i}\right\} \quad (7.6)$$

dans laquelle d_e est la distance finale de l'agent à la cible, soit à la position vers laquelle l'agent est guidé, d_i est la distance initiale à la cible.

Les scores s_i sont ensuite triés du plus fort au plus faible, leur rang étant alors utilisé comme coefficient pour calculer une moyenne pondérée qui nous donne le score final : $s_1 = \frac{2}{n(n+1)} \sum_{j=1}^n j * s_j$, où j est donc le rang du score et n le nombre d'essais avec la cible.

D'autre part, nous simulons également ces agents dans un environnement en une dimension sans cible. Dans ce cas, une position à atteindre est fixée en x_p avec $x_A - 11 < x_p < x_A - 8$ ou $x_A + 8 < x_p < x_A + 11$, x_A étant la position de l'agent. L'agent est ensuite guidé vers cette position en utilisant la méthode décrite ci-dessus.

Le score pour un essai est donné par l'équation suivante :

$$s_i = \max\left\{0, 1 - \frac{d_e}{d_i}\right\} \quad (7.7)$$

dans laquelle d_e est la distance finale de l'agent à la position vers laquelle l'agent est guidé, d_i est la distance initiale à la position vers laquelle l'agent est guidé.

Les scores s_i sont ensuite triés du plus fort au plus faible, leur rang étant alors utilisé comme coefficient pour calculer une moyenne pondérée qui nous donne le score final : $s_2 = \frac{2}{n(n+1)} \sum_{j=1}^n j * s_j$, où j est donc le rang du score et n le nombre d'essais avec le guidage.

Le score final d’un agent pour cette phase est donné par l’équation suivante :

$$s = a * s_1 + b * s_2 \tag{7.8}$$

avec $a = 0.6$ et $b = 0.4$.

Nous avons lancé, pour des CTRNN de 3 à 5 neurones, 10 fois l’algorithme évolutionnaire avec les populations obtenues au chapitre 5 et avec des populations initiales arbitraires. L’évolution s’arrête si un agent obtient un score supérieur à 0.95 et a été sélectionné au moins cinq fois ou si la génération 10000 été atteinte.

Nous avons simulé la meilleure population finale issue de ces différents types d’évolutions pendant des essais type utilisés pour l’évaluation d’une génération et enregistré leur activité.

		Meilleur score f	Score moyen f	Meilleur score s_1	Score moyen s_1	Meilleur score s_2	Score moyen s_2
Évolution incrémentale	3 neurones	0.45	0.26	0.18	0.02	0.94	0.54
	4 neurones	0.45	0.30	0.15	0.05	0.96	0.65
	5 neurones	0.90	0.33	0.89	0.11	0.97	0.61
Évolution directe	3 neurones	0.88	0.35	0.89	0.18	0.99	0.57
	4 neurones	0.51	0.31	0.33	0.08	0.96	0.60
	5 neurones	0.49	0.33	0.19	0.13	0.96	0.61

TAB. 7.1 – Évolution incrémentale : scores selon le nombre de neurones pour la première phase.

à partir des résultats présentés dans le tableau 7.1, nous pouvons remarquer que si les scores moyens sont sensiblement les mêmes, nous avons des agents, issus de l’évolution incrémentale de CTRNN à 5 neurones et de l’évolution directe de CTRNN à 3 neurones, qui réussissent mieux la tâche. La figure 7.1 représente l’évolution des scores (courbes lissées) pour les deux essais dont sont issus ces agents.

L’évolution incrémentale est beaucoup plus rapide et donne de meilleurs résultats¹. Les scores sont ceux du meilleur agent à chaque génération et si autour de la génération 300, un basculement s’opère entre s_1 et s_2 pour le meilleur agent, d’autres agents dans la population

¹ L’évaluation reportée dans le tableau 7.1 montre cependant des résultats bien meilleurs pour le CTRNN à 3 neurones que ceux obtenus lors de l’évolution, ce qui souligne encore une fois l’instabilité potentielle du comportement des agents.

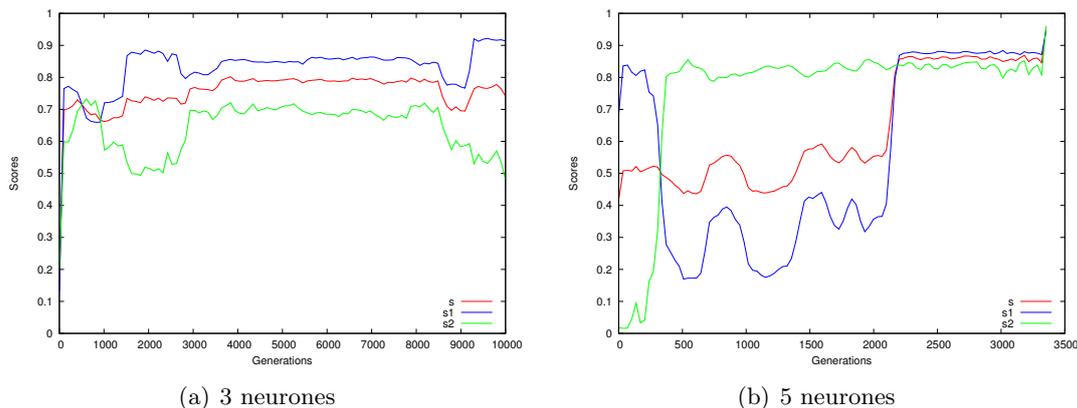


FIG. 7.1 – Évolutions des scores pour les comportements “innés” du meilleur agent par génération. (a) Évolution de CTRNN de 3 neurones à partir d’une population aléatoire. (b) Évolution de CTRNN de 5 neurones à partir d’une population obtenue par algorithme évolutionnaire avec s_1 comme critère de sélection.

conservent un bon score pour s_1 . Nous présentons dans la figure 7.2 les simulations des meilleurs agents issus de ces deux essais. Les comportements obtenus sont clairement différents en présence de la cible. En effet, la trajectoire de l’agent à trois neurones est très oscillante, il s’approche progressivement de l’objectif. Alors que l’agent à cinq neurones procède en deux étapes, il avance systématiquement jusqu’à la position $x = 18$ environ, puis approche de manière directe l’objectif.

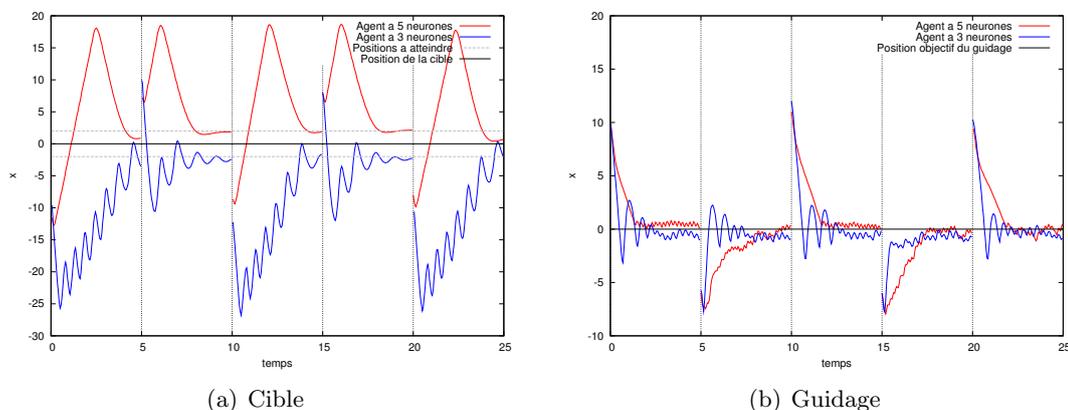


FIG. 7.2 – Comportements “innés” des deux meilleurs agents. (a) Trajectoires en présence de la cible. (b) Trajectoires en présence du guidage.

7.1.3.2 Évolution de l’apprentissage ontogénétique

Dans cette phase, nous utilisons Microbial pour faire évoluer des agents de manière à ce qu’ils soient capables de modifier leur comportement sensorimoteur en fonction d’un guidage.

Ces agents sont simulés dans un environnement en une dimension dans lequel est placé une cible. x_A étant la position initiale de l’agent, la cible est placée au début d’un essai en x_c avec $x_A - 11 < x_c < x_A - 6$ ou $x_A + 6 < x_c < x_A + 11$, x_A . Dans le repère de l’agent, la cible est mobile avec une vitesse constante pendant chaque essai, mais qui varie d’un essai à l’autre. Nous fixons une distance $< d_c <$ que l’agent doit respecter par rapport à la cible. L’agent est alors guidé s’il s’éloigne de la cible ou s’il s’en approche trop. De plus, le guidage est désactivé en cours d’essai (ici à $t_i + 6$, la fin de l’essai étant fixé à $t_e = t_i + 10$).

Le score pour un essai est donné par l’équation suivante :

$$s_i = \max\{0, f_g * (1 - \frac{\|d_e - \Delta\|}{\|d_i - \Delta\|})\} \quad (7.9)$$

dans laquelle d_e est la distance finale de l’agent soit à la cible, soit à la position vers laquelle l’agent est guidé, d_i est la distance initiale soit à la cible, soit à la position vers laquelle l’agent est guidé et Δ est la distance à maintenir que l’agent doit “apprendre”. f_g est fonction du temps de guidage sur un essai :

$$f_g = \frac{\int_{ts_i}^{te_i} G(t) dt}{te_i - ts_i} \quad (7.10)$$

$$G(t) = \begin{cases} 1 & : IG(t) = 0 \\ 0 & : sinon \end{cases} \quad (7.11)$$

s_i est donc maximum si l’agent atteint la position objectif et que le temps de guidage a été minimum. Les scores s_i sont ensuite triés du plus fort au plus faible, leur rang étant alors utilisé comme coefficient pour calculer une moyenne pondérée qui nous donne le score s_3 : $s_3 = \frac{2}{n(n+1)} \sum_{j=1}^n j * s_j$, où j est donc le rang du score et n le nombre d’essais.

Pour tenter de préserver le comportement inné que nous avons obtenu par évolution à la phase précédente, nous continuons d’évaluer s_1 et s_2 pour chaque génération. Le score final d’un agent est alors donné par :

$$s = a * s_1 + b * s_2 + c * s_3 \quad (7.12)$$

avec $a = 0.3$, $b = 0.3$ et $c = 0.4$.

7.2 Analyses

7.2.1 Résultats de l’évolution

Nous avons simulé les meilleurs agents obtenus dans les conditions utilisées par l’algorithme évolutionnaire et enregistré leur activité. Le tableau 7.2 présente les scores obtenus par ces agents lors de ces simulations.

	Meilleur score s	Score moyen s	Meilleur score s_1	Score moyen s_1	Meilleur score s_2	Score moyen s_2	Meilleur score s_3	Score moyen s_3
3 neurones	0.78	0.56	0.9	0.54	0.84	0.69	0.89	0.47
5 neurones	0.26	0.14	0.83	0.74	0.58	0.33	0.41	0.22

TAB. 7.2 – Évolution incrémentale : scores selon le nombre de neurones pour la deuxième phase.

7.2.2 Analyse comportementale

Nous proposons ici un aperçu du comportement des agents sélectionnés par l’algorithme évolutionnaire. Les figures 7.3, 7.4 et 7.5 représentent les trajectoires de trois agents pendant la simulation que nous avons enregistré. Les 4 premiers essais correspondent aux conditions de l’évaluation de s_1 . La cible est placée dans l’environnement à une distance aléatoire de l’agent. Elle se déplace à une vitesse constante différente à chaque essai. Le comportement “inné” de l’agent doit être de suivre la cible en mouvement tout en maintenant une distance $\Delta = 2$. Les essais 5 à 8 correspondent aux conditions de l’évaluation de s_2 . Il n’y a pas de cible dans l’environnement. Une position arbitraire est choisie comme objectif pour chaque essai. L’agent est alors guidé s’il s’éloigne de cette position. Le comportement “inné” de l’agent doit être d’atteindre cette position. Enfin, les essais 9 à 12 correspondent aux conditions de l’évaluation de s_3 . La cible est placée dans l’environnement à une distance aléatoire de l’agent. Elle se déplace à une vitesse constante différente à chaque essai. Un Δ arbitraire est fixé et l’agent est guidé s’il s’approche de moins de Δ de la cible ou s’il s’en éloigne de plus de cette valeur. Le comportement “acquis” de l’agent doit être de suivre la cible en mouvement tout en maintenant la distance Δ auquel il a accès uniquement par le guidage. Il doit donc en résulter un apprentissage ontogénétique comportemental de l’agent grâce au guidage.

Nous étudions ces trois agents car ils sont représentatifs de différents types de comportements qui se retrouvent dans la population. Ils sont tous les trois contrôlés par un CTRNN à 3 neurones.

L’agent dont nous traçons la trajectoire sur la figure 7.3 obtient pour ces essais les scores suivants : $s_1 = 0.90$, $s_2 = 0.82$ et $s_3 = 0.66$, ce qui donne $s = 0.78$. L’agent semble avoir de manière “innée” les comportements recherchés par rapport à la cible aussi bien que par rapport au guidage, même si pour celui-ci il reste à une certaine distance de la position objectif. Cependant, en ce qui concerne l’apprentissage ontogénétique, il semble que le comportement lié à la cible prenne le pas sur la réaction au guidage. Ainsi, le comportement est semblable entre les essais 10 et 12, pour lesquels la cible est à la même distance initiale de l’agent, mais qui diffèrent par la vitesse de la cible ainsi que par le Δ que le guidage doit introduire comme distance à respecter. Cet agent ne semble donc pas avoir la sensibilité au guidage nécessaire pour pouvoir apprendre le comportement recherché.

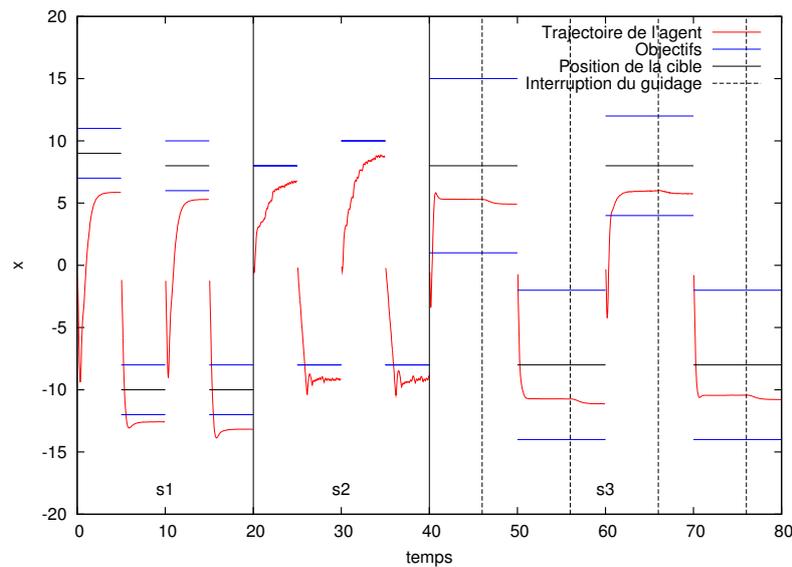


FIG. 7.3 – Trajectoires d’un agent pendant une simulation correspond aux essais effectués par génération. Le comportement de cet agent semble être conditionné essentiellement par la cible.

L’agent dont nous traçons la trajectoire sur la figure 7.4 obtient pour ces essais les scores suivants : $s_1 = 0.1$, $s_2 = 0.71$ et $s_3 = 0.44$, ce qui donne $s = 0.41$. L’agent ne semble pas avoir de réaction spécifique à la cible. En particulier, sur les 4 essais représentés, pour les essais 1 et 3, la cible est à la même position, ainsi que pour les essais 2 et 4. Pourtant, l’agent ne montre pas du tout le même comportement. Surtout, il a le même comportement durant les essais 1 et 4, ou 2 et 3. En adoptant l’un ou l’autre de ces comportements, l’agent a une chance sur deux de s’approcher de la cible. Par contre, cet agent a un comportement satisfaisant par rapport au guidage (phase s_2), bien qu’il reste éloigné de la position objectif. De manière assez prévisible puisque l’agent ne maîtrise pas les deux comportements “innés”, les essais correspondant à l’apprentissage (phase s_3) ne sont pas réussis. En effet, l’agent réagit au guidage et s’approche de la position recherchée, mais quand le guidage est interrompu, il ne maintient pas la distance à la cible qui est en mouvement nous le rappelons. Cet agent ne semble donc pas avoir la sensibilité à la cible nécessaire pour pouvoir apprendre le comportement recherché.

L’agent dont nous traçons la trajectoire sur la figure 7.5 obtient pour ces essais les scores suivants : $s_1 = 0.75$, $s_2 = 0.76$ et $s_3 = 0.64$, ce qui donne $s = 0.71$. L’agent semble avoir de manière “innée” les comportements recherchés par rapport à la cible (phase s_1) aussi bien que par rapport au guidage (phase s_2), même si dans les deux cas il reste à une certaine distance de la position objectif. En ce qui concerne l’apprentissage ontogénétique (phase s_3), il semble que le comportement lié à la cible soit transformé par la réaction au guidage. Ainsi, le comportement est sensiblement différent entre les essais 10 et 12, pour lesquels la cible est à la même distance initiale de l’agent, mais qui diffèrent par la vitesse de la cible ainsi que par le Δ que le guidage doit introduire comme distance à respecter. Lors de l’essai 10 Δ est plus élevé que lors de l’essai 12. L’agent se positionne en effet plus loin de la cible. Une fois le guidage interrompu, il y a un léger décrochage, mais l’agent maintient globalement la même distance à la cible. Bien que le comportement obtenu ne soit pas totalement celui

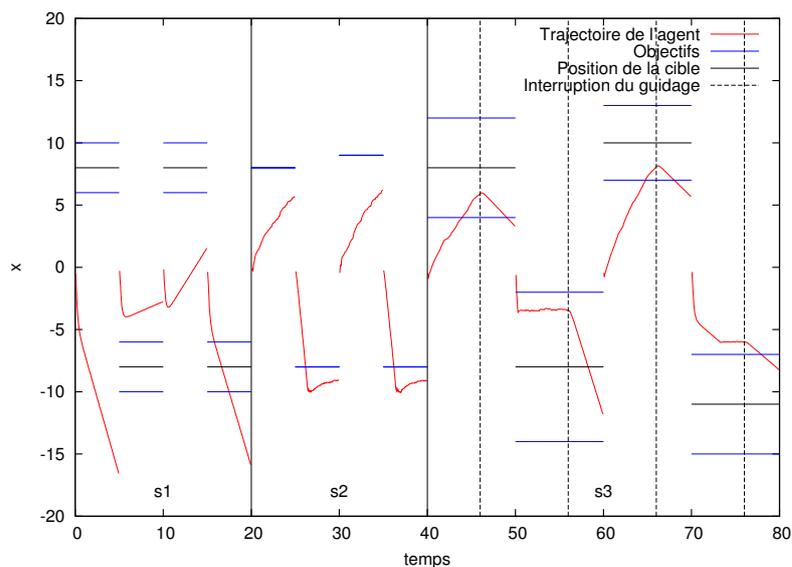


FIG. 7.4 – Trajectoires d'un agent pendant une simulation correspond aux essais effectués par génération. Le comportement de cet agent semble être conditionné essentiellement par le guidage.

que nous recherchions, il semble donc que cet agent ait la sensibilité au guidage et à la cible nécessaires pour que son comportement reflète les interactions particulières qu'il a eues avec l'environnement et soit donc capable d'assimiler des perturbations de manière à produire un comportement transformé à l'échelle ontogénétique.

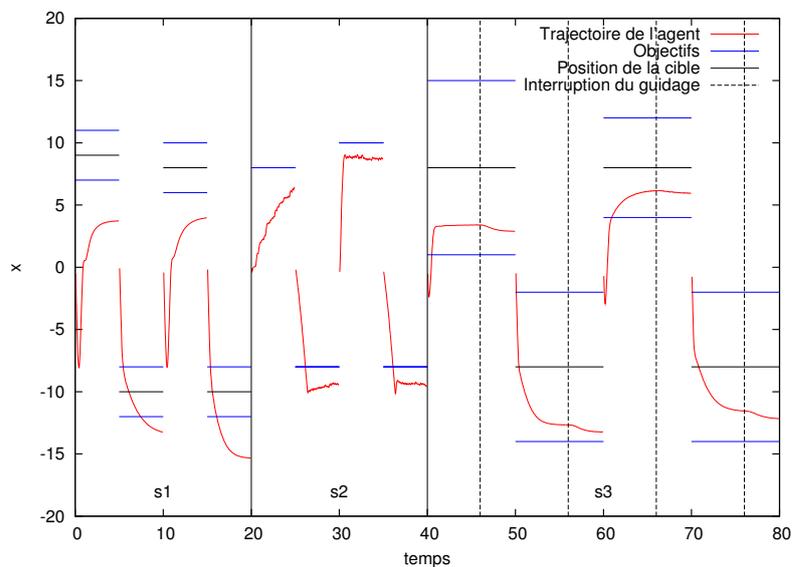


FIG. 7.5 – Trajectoires d'un agent pendant une simulation correspond aux essais effectués par génération. Le comportement de cet agent semble dépendre à la fois de la cible et du guidage, il reflète un début de développement du comportement à l'échelle ontogénétique.

7.2.3 Analyse dynamique

Dans cette section, nous nous intéressons à l’activité interne des trois agents étudiés précédemment.

La figure 7.6 représente l’activité interne de l’agent qui est celui que nous avons identifié comme étant essentiellement conditionné par la cible. Les courbes en rouge correspondent aux conditions de l’évaluation de s_1 (cible seule), les courbes bleues correspondent aux conditions de l’évaluation de s_2 (guidage seul) et enfin les courbes vertes correspondent aux conditions de l’évaluation de s_3 (apprentissage avec cible et guidage). Il existe deux attracteurs bien séparés correspondant à la poursuite de la cible d’une part et à la présence du guidage seul d’autre part. Pendant les essais liés à l’apprentissage, la dynamique interne est légèrement différente, même si l’on peut y déceler une faible part se rapprochant de la dynamique de guidage car elle présente des variations.

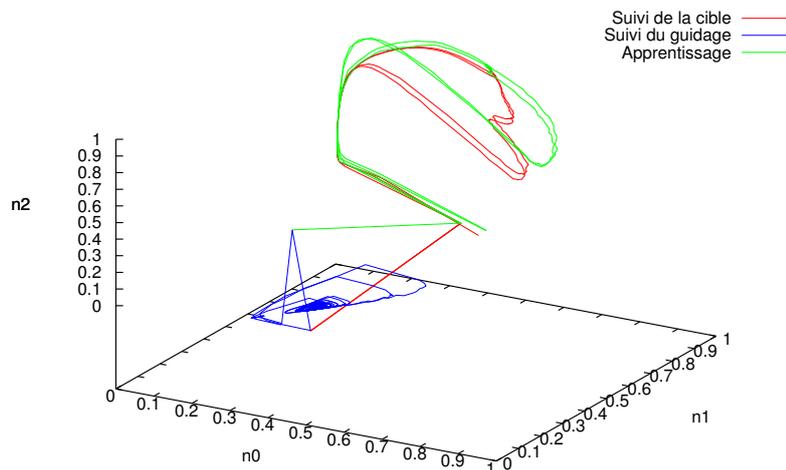


FIG. 7.6 – Dynamique interne d’un agent conditionné par la cible pendant une simulation correspond aux essais effectués par génération.

La figure 7.7 représente l’activité interne d’un agent pendant la simulation que nous avons enregistré. Cet agent est celui que nous avons identifié comme étant essentiellement conditionné par le guidage. Il existe également deux attracteurs bien séparés correspondant à la poursuite de la cible d’une part et à la présence du guidage seul d’autre part. Pendant les essais liés à l’apprentissage, la dynamique interne est très proche de celle que l’on retrouve dans les essais de poursuite de la cible, mais on retrouve pendant les essais une dynamique dont le motif ressemble à celui correspondant au guidage. En effet, la dynamique ne se stabilise pas aussi nettement que lorsqu’il n’y a que la cible.

La figure 7.8 représente l’activité interne d’un agent pendant la simulation que nous avons enregistré. Cet agent est celui que nous avons identifié comme montrant un début de développement du comportement à l’échelle ontogénétique. Il existe également deux

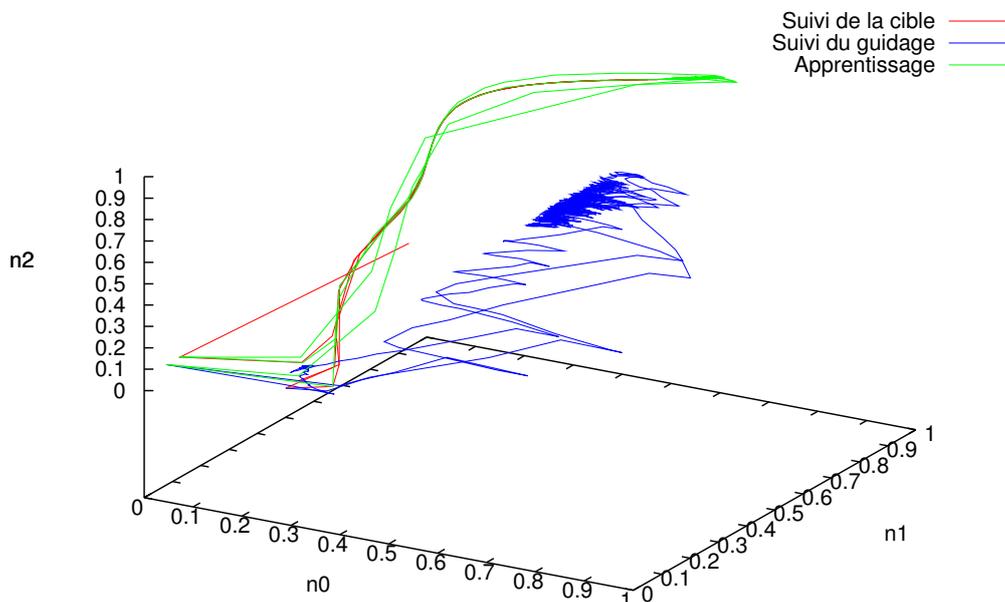


FIG. 7.7 – Dynamique interne d’un agent conditionné par le guidage pendant une simulation correspond aux essais effectués par génération.

attracteurs bien correspondant à la poursuite de la cible d’une part et à la présence du guidage seul d’autre part. Pendant les essais liés à l’apprentissage, la dynamique interne est proche de celle que l’on retrouve dans les essais de poursuite de la cible, mais bien distincte cependant, car en quelque sorte “décalée”. La dynamique de cet agent, caractéristique de son comportement à la cible, a donc été perturbée par le guidage, sans être pour autant “écrasée”, ce qui est un premier pas vers notre objectif de transformation du comportement à l’échelle ontogénétique.

7.3 Synthèse

Dans ce chapitre, nous avons présenté une expérience dont l’objectif est de réaliser un apprentissage comportemental ontogénétique par guidage. Il s’agit d’une expérience minimaliste dans un environnement en une dimension, mais qui préserve l’aspect sensorimoteur du comportement et qui utilise des CTRNN sans aménagements, théoriquement suffisants pour des agents éactifs virtuels, comme contrôleurs d’agents. La tâche des agents consiste à “apprendre”, par l’interaction avec un guide virtuel, une distance qu’ils doivent maintenir avec une cible.

Nous nous sommes basés en partie sur des agents obtenus dans les expériences présentées dans le chapitre 5 et bien qu’un nouveau capteur soit ajouté, l’évolution incrémentale s’est

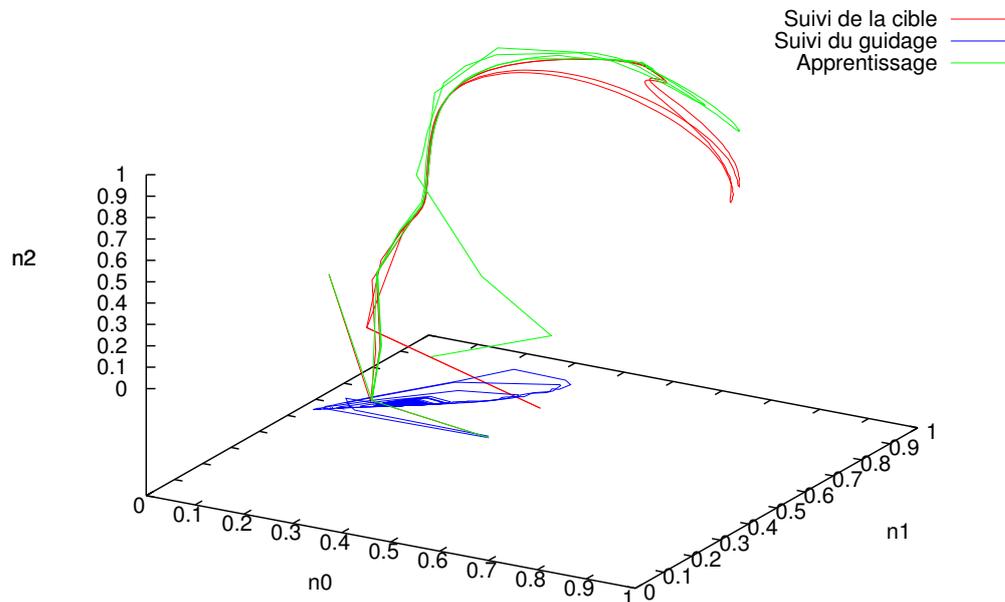


FIG. 7.8 – Dynamique interne d’un agent perturbé par le guidage pendant une simulation correspond aux essais effectués par génération.

révélee efficace pour la première étape. Par évolution phylogénétique, nous obtenons alors des agents capables de deux comportements “innés” relativement stables, qui consistent pour l’un à maintenir une distance $d = 2$ à une cible qu’ils captent dans l’environnement ; pour l’autre à atteindre un objectif qui n’est pas matérialisé, l’agent ne captant que des signaux de guidage qui “corrige” son comportement.

Nous avons ensuite cherché à faire évoluer les populations obtenues pour obtenir un apprentissage à l’échelle ontogénétique. Une étude comportementale de certains agents obtenus semble montrer la possibilité d’une évolution du comportement liée au guidage, même si celle-ci ne correspond pas complètement à l’apprentissage du comportement recherché. En particulier, une évolution de la dynamique interne est mise en avant chez le meilleur agent obtenu.

Ce résultat montre que le guidage peut orienter l’évolution à l’échelle ontogénétique du comportement d’un agent contrôlé par un CTRNN et paramétré selon une approche évolutionnaire. Les différentes expériences que nous avons menées illustrent la difficulté de l’application d’une approche évolutionnaire pour obtenir des agents éactifs virtuels, due en partie à l’instabilité des CTRNN et également au non déterminisme des simulations dont nous avons besoin pour obtenir un comportement spécifique à l’interaction entre agent et environnement.

Conclusions

LA ligne directrice qui m'a été donnée au début de ce travail était de creuser la notion d'autonomie telle qu'elle est abordée par les sciences cognitives, et en particulier par le paradigme de l'énaction, pour dégager des pistes applicables en réalité virtuelle qui renforcent le principe d'autonomie. Deux points de vue ont alors été envisagés : celui de l'homme et celui des modèles. Nous avons choisi de nous centrer sur l'autonomisation de modèles artificiels à un niveau comportemental, ce qui implique des capacités de résistance mais aussi d'interaction, ainsi qu'une variabilité mais également de la régularité.

Dans le cadre de l'énaction, un système autonome et cognitif peut être défini comme un système soumis à une causalité circulaire qui génère et maintient de manière active une identité.

Dans le domaine de l'artificiel, des travaux en vie artificielle modélisent et simulent une autonomie constitutive à partir de l'étude de la constitution d'un organisme, mais laissent peu de place au comportement, c'est-à-dire à l'action de l'organisme. L'intelligence artificielle et la robotique proposent quant à elle des réflexions pour produire des comportements autonomes. Les courants récents dans ce domaine écartent la notion de représentation interne de l'environnement et ont tendance à se rapprocher de considérations éenactives sur la cognition artificielle.

à partir de ces travaux, nous avons identifié un modèle qui nous a semblé le plus adapté à la conception d'agents éenactifs : les CTRNN (Continuous Time Recurrent Neural Networks). Ces réseaux permettent de simuler un système dynamique, qui peut être couplé à un environnement ouvert. L'approche évolutionnaire que nous avons également exposée est une méthodologie appropriée pour paramétrer ce type d'agents.

Nous avons ensuite étudiés les travaux basés sur l'usage de ces modèles, concernant des propriétés caractéristiques de l'autonomie que nous recherchons. Cette étude, associée à notre objectif final, a montré qu'il fallait aborder la capacité d'évolution comportementale de ces systèmes, suite ou au cours de leurs interactions avec leur environnement.

Pour Varela, la cognition est l'énaction, c'est-à-dire *une histoire de couplage structurel qui fait émerger un monde* (Varela et al., 1993). L'apprentissage à l'échelle d'un individu résulte alors de l'histoire de ses interactions avec l'environnement à cette même échelle. Pour aboutir à cet apprentissage à l'échelle ontogénétique, il faut que l'individu co-évolue avec son environnement². Pour faire évoluer des agents énatifs à l'échelle ontogénétique, nous avons proposé d'utiliser un guidage interactif.

Cette proposition de guidage a été justifiée par le fait que nos agents sont situés, qu'ils adoptent des invariants sensorimoteurs et que les capteurs peuvent prendre le rôle de perturbateurs particuliers provoquant une transformation du comportement global qui reflétera l'histoire des interactions. Pour revenir à la réalité virtuelle, l'intervention d'un observateur humain sur l'évolution de la dynamique de l'environnement peut permettre à celui-ci de donner du sens à la transformation que l'interaction peut amener dans le comportement de l'agent.

Nous avons rencontré des difficultés pour obtenir des comportements interactifs ontogénétiques avec des agents contrôlés par des CTRNN. Nous avons pourtant adopté une approche progressive et minimaliste pour tenter d'y arriver. Ainsi, la première expérience décrite au chapitre 5 s'étant révélée trop complexe par rapport à notre objectif, nous avons réalisé ensuite une série d'expériences minimalistes afin de démontrer que les outils que nous utilisons permettent de synthétiser des agents qui adoptent des comportements relevant d'invariants sensorimoteur différents en utilisant une approche phylogénétique. Dans le chapitre 6, nous avons utilisé un guidage pour influencer la transformation de la dynamique d'un agent à une échelle ontogénétique. Nous avons illustré l'effet du guidage sur un agent en proposant une étude comportementale et dynamique de cet agent. Cette expérience présente toutefois des limites et bien que le guidage influence le comportement, il est délicat d'affirmer que l'association soit faite de façon définitive et irréversible, en particulier à cause de la sensibilité des CTRNN. Enfin, dans le chapitre 7, le guidage est appliqué afin d'effectuer la transformation de comportements sensorimoteurs obtenue dans le chapitre 5 cette fois-ci à une échelle ontogénétique. Après avoir présenté l'évolution des comportements "innés" sur la base desquels l'apprentissage doit se faire, nous avons étudié le comportement et la dynamique des agents obtenus après la deuxième phase de l'évolution. Le comportement que nous obtenons n'est pas tout à fait celui que nous recherchons. Cependant, nous montrons à travers l'analyse comportementale et dynamique de certains agents, qu'il est possible d'obtenir une modification comportementale induite par le guidage à l'échelle ontogénétique.

Discussion et perspectives

L'objectif à long terme, annoncé au début du mémoire, était d'explorer le principe d'autonomie pour la réalité virtuelle. Il nous a mené à une recherche exploratoire pour associer autonomie constitutive et développement ontogénétique comportemental. Notre démarche n'a été illustrée que par des exemples minimalistes et, malgré de nombreuses tentatives et

² Une intelligence artificielle basée sur l'énaction doit respecter les principes d'irréversibilité, de développement ontogénétique et de signification par l'interaction avec un observateur humain (De Loor et al., 2009)

le temps passé, nous restons très prudents quant aux résultats obtenus. Ces constatations amènent deux réflexions. Tout d'abord, une analyse des échecs rencontrés nous permet de formuler des conséquences et de proposer des perspectives pour les dépasser. Ensuite, nous discutons de l'exploitation de l'autonomie constitutive en réalité virtuelle.

Leçons sur un travail exploratoire

Nous avons proposé une méthode pour permettre un développement comportemental sensorimoteur à l'échelle ontogénétique pour des agents éactifs virtuels, dont l'organisation est circulaire. Dans ce contexte, le choix du CTRNN était parfaitement justifié et la plupart des travaux se positionnant dans le même cadre que le nôtre utilisent ces modèles. D'un point de vue théorique, il n'était pas nécessaire de changer de modèle, ni même de l'aménager. Notre démarche a été fondée sur la préservation des caractéristiques "fondamentales" tout en acceptant des simplifications acceptables, qui ont consisté à s'attaquer à des exemples minimalistes.

Le minimalisme des environnements et des tâches que nous avons choisis pour illustrer notre propos n'en rendent pour autant pas la mise en œuvre simple. En effet, comme nous l'avons illustré, la sensibilité des CTRNN rend la sélection par algorithme génétique délicate. La mise au point des conditions dans lesquelles vont être évalués les agents ainsi que le choix d'une fonction d'évaluation efficace requièrent de nombreux ajustements qui prennent beaucoup de temps. De plus, l'évolution en elle-même est également longue. à chaque étape de nos expériences, nous avons ainsi dû faire de nombreux aller-retours entre mise au point des expériences et évolution des agents. Malgré cela, les résultats sont assez frustrants. Ceci nous permet de tirer quelques leçons sur l'approche évolutionnaire appliquée à des agents contrôlés par des CTRNN pour obtenir des comportements complexes.

Tout d'abord, les difficultés que nous avons rencontrées peuvent paraître étonnantes au vu des résultats obtenus par Beer et Gallagher (1992); Di Paolo (2000); Iizuka et Di Paolo (2007); Tuci et al. (2002) ou encore Izquierdo et al. (2008). Cependant, dans ces travaux, seul le meilleur agent est analysé et il semble qu'il ne soit pas obtenu si aisément. De plus, bien que ces expériences montrent des propriétés tout à fait intéressantes des CTRNN, la tâche est souvent très simple, ne faisant pas appel à la notion de situation aussi fortement que dans les expériences que nous avons présenté au début du chapitre 5 et surtout ne mettant pas en œuvre de développement ontogénétique comme nous avons cherché à faire dans les autres expériences. Les comportements les plus complexes sont d'ailleurs obtenus en modifiant le modèle de manière à greffer des mécanismes supplémentaires. C'est le cas de la plasticité homéostatique par exemple qui permet de maintenir plusieurs buts. D'un point de vue théorique, il était important de tester notre démarche sans aménagements du modèle.

Un problème-clé de ce type d'approches est lié à la maîtrise du rapport entre la part de l'inné et de l'acquis. Plus l'ontogenèse est complexe, plus cette maîtrise est difficile. Le *développement* implique l'existence d'une base sur laquelle viennent se greffer de nouvelles capacités. Cette base constitue un savoir-faire *inné* dont il est impossible de se passer. Une partie de ce savoir-faire est inscrite dans le corps. Cependant, une autre partie de la base

innée des connaissances d'un agent doit se trouver dans son contrôleur, que nous pouvons comparer au système nerveux d'un organisme³. Le développement potentiel de l'agent dépend de cette base et des interactions avec l'environnement.

Les expériences que nous avons menées portent sur des comportements minimalistes et pourtant, même si nous avons cherché à obtenir un développement comportemental ontogénétique, les connaissances "innées" que nous avons utilisées sont très proches du comportement développé final. La question du lien entre la complexité du potentiel développemental et la complexité de la base de connaissances pré-requises se pose. Comme nous l'avons vu avec l'expérience de Tuci et al. (2002) au chapitre 1, cette question est déjà présente pour le développement d'un comportement de manière évolutionnaire. Ainsi, il est nécessaire de déterminer les réflexes, les traits adaptatifs, qui permettent à tel comportement d'apparaître.

Pour poursuivre les objectifs que nous nous sommes fixés, nous envisageons deux possibilités. La première consiste à faire des concessions sur le modèle en introduisant de la plasticité. La seconde est de proposer une alternative ou un complément aux algorithmes évolutionnaires permettant l'analyse des modèles et ainsi une meilleure maîtrise de la dynamique.

Développement et plasticité

La plasticité structurelle d'un système est ce qui peut permettre à ce système de faire évoluer son comportement propre (Varela, 1989). En effet, en réagissant à des perturbations, le système effectue des ajustements pour compenser ces perturbations. Les propriétés dynamiques des CTRNN leur permettent de s'adapter⁴. Ils sont théoriquement capables de reproduire n'importe quelle dynamique si le nombre de neurones est suffisant.

Pourtant, en pratique, les CTRNN ne sont utilisés qu'avec très peu de neurones. Nous pouvons raisonnablement penser que plus les comportements recherchés seront complexes, plus le nombre de neurones requis sera grand. De plus, Beer (2006) montre qu'il peut être plus efficace dans la pratique d'utiliser plus de neurones que strictement nécessaire pour arriver à un résultat satisfaisant plus rapidement. Dans cette situation, introduire un mécanisme qui renforce la plasticité du système peut être intéressant.

La plasticité homéostatique de Iizuka et Di Paolo (2007), *cf.* section 3.1.3, permet par exemple à un agent de maintenir en concurrence deux comportements différents. Le choix du comportement adopté à un instant donné est déterminé à la fois par la dynamique interne et par la dynamique externe. Selon ce modèle, le comportement que l'agent adopte correspond à une *habitude*, concrétisée par des zones de stabilité pour l'activité neuronale. Il semble envisageable de multiplier les habitudes d'un agent et donc les comportements qu'il est susceptible d'adopter. Cependant, introduire de nouveaux comportements n'est pas

³ Pour Piaget, l'intelligence sensorimotrice commence chez l'enfant par l'exercice des réflexes. Ces réflexes sont installés dans le système nerveux, ce ne sont pas des propriétés *matérielles* de l'organisme. Ils constituent une base de savoir-faire modulables pour la construction d'un répertoire comportemental.

⁴ Il ne s'agit toutefois pas d'une modification de la topologie du réseau, mais d'une évolution en terme de réaction aux perturbations.

directement possible avec ce mécanisme. Il faudrait pour cela que de nouvelles zones de stabilité puissent être créées dynamiquement et par conséquent que l'agent ait un moyen de régulation de ses règles de plasticité. Ceci permettrait également d'affiner les habitudes ou encore de les *oublier*.

Pour rendre le système plus souple, un mécanisme de modulation du réseau par son activité même, inspiré du travail de Soltoggio (2008), pourrait par exemple être envisagé. Dans ce travail, des neurones sont utilisés spécifiquement pour moduler les connexions. Toutefois, Izquierdo (2008) relève le danger à considérer que puisqu'avec de la plasticité, il est possible d'obtenir un apprentissage, la plasticité est alors le mécanisme *responsable* de l'apprentissage. Il faut donc faire attention à ne pas introduire en tant qu'élément explicatif du comportement du système ce qui n'est qu'un outil de régulation supplémentaire.

En ce sens, de tels *aménagements* des CTRNN ont le désavantage, par l'introduction d'hypothèses, de réduire la part de son autonomie constitutive. Néanmoins, elles peuvent se révéler être des pistes à creuser pour obtenir un développement ontogénétique plus abouti. Il s'agit d'être capable de trouver une solution qui facilite le développement, tout en conservant une autonomie dans la production du comportement qui doit rester le fruit de la dynamique interne d'un agent, perturbée et non pas entièrement déterminée par des commandes provenant de l'environnement.

Nous présentons ici quelques idées pour augmenter la souplesse du contrôleur :

- ▷ Utiliser des règles de plasticité sur les connexions du réseau de neurones. Une plasticité de type hebbien par exemple module la force d'une connexion entre deux neurones en fonction de la corrélation de leurs activités⁵.
- ▷ Le guidage pourrait être la source de la modulation de l'activité du réseau, alors que pour l'instant il est une perturbation. Ou bien de manière plus intégrée, la modulation pourrait être globale, chaque neurone modulant potentiellement chaque connexion.
- ▷ La plasticité peut être envisagée au niveau des poids des connexions, mais aussi au niveau des gains, des constantes de temps ou encore des biais. Ceci complique évidemment encore grandement la complexité du système.

Analyse in virtuo

Nous pouvons imaginer que l'analyse des CTRNN progressera facilitant la tâche du concepteur ou encore que d'autres modèles théoriques seront un jour proposés permettant de mieux maîtriser la dynamique des systèmes et d'éviter ou du moins de réduire l'usage d'approches évolutionnaires. Ceci est une perspective théorique à long terme qu'il est nécessaire de creuser pour atteindre nos objectifs initiaux. La théorie de la viabilité fait partie de ce type d'approches (Aubin, 1991). Elle est toutefois pour l'instant trop limitée en puissance opérationnelle pour être applicable dans notre cas. Cependant, même si le paramétrage d'un contrôleur était simplifié, résolvant le problème de la conception d'agents, celui du développement ontogénétique interactif donc fortement variable, resterait.

Une piste intéressante face à la complexité des CTRNN est l'expérience *in virtuo*

⁵ Nous avons utilisé cette plasticité dans les expériences présentées dans le chapitre 6.

(Tisseau, 2001). Permettre cette expérience consiste à donner au concepteur des moyens d'expérimentation interactive du modèle qu'il crée. Il s'agit donc ici de construire des outils permettant le prototypage interactif des agents. Modifier les paramètres d'un agent ne constitue pas une difficulté, mais n'a pas de sens si les conséquences de ces modifications ne peuvent pas être analysées. Pour cela, il est nécessaire d'imaginer des moyens de visualisation, qu'ils soient évidents comme la visualisation du comportement de l'agent ou bien plus complexes comme la visualisation de la dynamique de l'agent. De plus, ces outils devront permettre d'interagir avec les agents pour initier des perturbations et ainsi permettre à l'expérimentateur de voir comment l'agent y réagit et éventuellement les compense. En effet, cet aspect est déterminant pour provoquer un développement ontogénétique.

Exploitation en réalité virtuelle

Engagement, consistance et résistance

Nous avons exposé dans le chapitre 1 le point de vue de De Loor (2006) qui considère que pour un observateur la pertinence d'un comportement autonome nécessite que celui-ci soit suffisamment consistant pour pouvoir être compris, anticipé et suffisamment résistant pour surprendre et susciter un engagement. Cette question de la créativité dans le comportement autonome est cruciale dans l'instauration d'une co-évolution entre l'homme et le système, qui fasse sens pour l'homme.

Elle n'a pas été traitée explicitement dans ce travail. Toutefois, nous constatons que les agents que nous avons obtenus montrent une grande résistance! Ceci nous a posé problème dans la réalisation d'expériences même minimalistes. Alors que nous voulions mettre de côté ce que nous appelions indéterminisme des comportements, en ne cherchant à obtenir qu'un comportement, cet indéterminisme apparaît tout de même à travers une imprévisibilité du comportement certainement dûe à la nature chaotique du contrôleur. Cependant, cette imprévisibilité du comportement laisse à penser que les CTRNN sont de bons candidats pour qui s'intéresse à la *créativité artificielle*. Ils sont d'ailleurs utilisés dans des domaines artistiques, notamment par Bown et Wiggins (2005) dans le cadre de la musique.

La plasticité homéostatique de Iizuka et Di Paolo (2007) ouvre également des pistes de ce côté. Comme nous l'expliquions précédemment, elle associe des comportements habituels à des zones de stabilité de l'activité neuronale. Bien sûr, plus ces comportements seront nombreux, plus la zone de stabilité associée à chacun sera réduite. En prenant comme hypothèse que la force de l'habitude dépend de la taille de cette zone, on peut supposer que des zones très réduites pourraient entraîner des changements de comportements plus fréquents provoqués par la dynamique interne. Il pourrait être intéressant de tester cette hypothèse pour tenter de produire des agents plus ou moins prévisibles dans leur changement de comportement ou bien qui ont des habitudes plus prononcées pour un comportement que pour d'autres.

Du guideur virtuel au guideur humain

Lors de ce travail, nous avons été amenés à faire des choix parmi différentes pistes identifiées. Ainsi, nous nous sommes intéressés à l'autonomisation de modèles artificiels, là où nous aurions pu nous concentrer sur le point de vue humain. Cependant, nous avons été amenés à intégrer ce point de vue de l'homme dans la construction de notre proposition à travers l'utilisateur comme initiateur d'un guidage qui rend pertinente l'évolution d'un agent artificiel dans le cadre d'applications en réalité virtuelle.

Étant donné le caractère exploratoire de ce travail, l'utilisateur humain n'a pas été pourtant acteur de nos tentatives d'apprentissage par guidage. Nous avons été contraints de le remplacer par un guideur virtuel, ce qui implique des restrictions dans la souplesse des interactions possibles. Les perspectives possibles pour obtenir un guidage plus fin sont soit de permettre à un utilisateur humain de guider l'agent, soit de faire co-évoluer plusieurs agents qui se guideraient les uns les autres. Cependant, comme nous l'avons déjà évoqué, il y a ici un risque de perte de sens du comportement des agents artificiels pour des observateurs humains.

Le passage d'un guideur virtuel à un guideur humain n'est néanmoins pas trivial. En effet, nous pouvons supposer que chacun adopte une stratégie de guidage particulière et peut-être même propre à un épisode de guidage. Or, dans les expériences que nous avons présentées, les agents sont sélectionnés en partie pour leur réaction à un type de guidage, utilisé ensuite lors de l'apprentissage. Tant que l'approche évolutionnaire restera nécessaire au paramétrage des contrôleurs d'agents, il paraît difficile d'envisager de manière directe une interaction avec un utilisateur humain.

Ceci pourrait toutefois se tenter en conservant une approche évolutionnaire, mais en travaillant de manière plus approfondie sur la modélisation du guidage qu'un utilisateur pourrait mettre en œuvre. Cette modélisation nécessiterait a priori de s'intéresser aux temps de réaction de l'utilisateur, c'est-à-dire aux délais entre l'observation du comportement et l'émission d'un signal de guidage. La stratégie de guidage en elle-même devrait aussi être étudiée, en déterminant les "marges d'erreurs" autorisées ou encore en étudiant l'influence du progrès de l'agent sur la manière dont l'utilisateur le guide.

Une fois cette modélisation obtenue pour une population, des sous-groupes ou des individus spécifiques, il serait possible d'implémenter un guideur virtuel sur la base de ce modèle et de faire évoluer des agents pour qu'ils réagissent à ce type de guidage et soient capables d'interagir avec la population, les sous-groupes ou les individus spécifiques dont sont tirés les modèles.

Pour conclure, il semble que bien qu'une autonomie faible restera sans doute encore longtemps le seul moyen d'obtenir des interactions homme-machine utilisables de manière pragmatique, il nous apparaît comme nécessaire de mieux étudier la notion d'autonomie constitutive pour la synthèse d'agents artificiels. Elle offre en effet des perspectives formidables en ce qui concerne des applications associant vie artificielle et réalité virtuelle. Nous la jugeons également indispensable pour la réalisation d'applications orientées vers la créativité.

Publications

Cette section énumère les articles et communications produits pendant cette thèse.

De Loor, P., Lannuzel, A., Manac'h, K., et Harrouet, F. (2007). Cognition dynamique artificielle appliquée à l'autonomisation d'un humanoïde virtuel. Dans *ARCo'07, Colloque de l'Association pour la Recherche Cognitive. Cognition-Complexité-Collectif. Communications affichées*.

De Loor, P., Manac'h, K., Fronville, A., et Tisseau, J. (2008). Requirement for an enactive machine : Ontogenesis, interaction and human in the loop. Dans *5th International Conference on Enactive Interfaces, ENACTIVE08*.

De Loor, P., Manac'h, K., Fronville, A., et Tisseau, J. (2009). Simulation of abstract autopoietic machine. Dans Kampis, G. et Szathmary, E., éditeurs, *ECAL : 10th European Conference on Artificial Life*, LNCS/LNAI 5777,5778. Springer.

De Loor, P., Manac'h, K., Parenthoën, M., et Tisseau, J. (2010). Approche énaactive de la cognition dynamique artificielle : que devient le contexte? Journées de Rochebrune : systèmes complexes naturels et artificiels.

De Loor, P., Manac'h, K., et Tisseau, J. (2009). Enaction-based artificial intelligence : Toward co-evolution with human in the loop. *Minds and Machines*, 19 :319–343.

Fronville, A., Manac'h, K., et De Loor, P. (2007). L'énaaction, une appropriation par des exemples. Séminaires du Laboratoire d'Informatique des Systèmes Complexes.

Manac'h, K. et De Loor, P. (2007). Passage du discret au continu : identification des verrous pour la simulation de systèmes autopoïétiques. Dans *ARCo'07, Colloque de l'Association pour la Recherche Cognitive. Cognition-Complexité-Collectif. Communications affichées*.

Manac'h, K. et De Loor, P. (2008). De l'autonomie à l'apprentissage interactif appliqué à un modèle de cognition artificielle dynamique plastique. Dans *ARCo'08 Colloque de l'Association pour la Recherche Cognitive. Connaissances : Genèse, Nature et Fonction*, pages 44–49.

Manac'h, K. et De Loor, P. (2008). La cognition dynamique artificielle appliquée à l'apprentissage interactif. Actes du séminaire annuel du Laboratoire d'Informatique des Systèmes Complexes.

Manac'h, K. et De Loor, P. (2008). A study on the origin of anticipation by guidance for artificial dynamic cognition. Dans *fourth workshop on : Anticipatory Behaviour in Adaptive Learning Systems (ABiALS 2008)*.

Manac'h, K. et De Loor, P. (2009). Guiding for associative learning : How to shape dynamical cognition. Dans Kampis, G. et Szathmary, E., éditeurs, *ECAL : 10th European Conference on Artificial Life*, LNCS/LNAI 5777,5778. Springer.

Manac'h, K., De Loor, P., et Tisseau, J. (2007). L'inscription virtuelle de l'autonomie. Actes du séminaire annuel du Laboratoire d'Informatique des Systèmes Complexes.

Références bibliographiques

- Andry, P., Gaussier, P., Moga, S., Banquet, J., et Nadel, J. (2001). Learning and communication via imitation : An autonomous robot perspective. *IEEE Transactions on Systems, Man, and Cybernetics, Part A : Systems and Humans*, 31(5) :431–442.
- Ashby, W. R. (1960). *Design for a Brain*. Chapman and Hall Ltd, London, second édition. First edition published 1952.
- Aubin, J.-P. (1991). *Viability theory*. Birkhäuser Boston Inc., Cambridge, MA, USA.
- Beer, R. et Gallagher, J. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, 1(1) :91–122.
- Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4) :471–511.
- Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3) :91 – 99.
- Beer, R. D. (2004). Autopoiesis and cognition in the game of life. *Artificial Life*, 10(3) :309–326.
- Beer, R. D. (2006). Parameter space structure of continuous-time recurrent neural networks. *Neural Computation*, 18 :3009–3051.
- Beurier, G., Michel, F., et Ferber, J. (2006). A morphogenesis model for multiagent embryogeny. Dans Rocha, Luis, M., Yaeger, Larry, S., Bedau, Mark, A., Floreano, D., Goldstone, Robert, L., et Vespignani, A., éditeurs, *ALIFE X, Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 84–90, Bloomington, Indiana, USA. MIT Press.
- Beurier, G., Simonin, O., et Ferber, J. (2003). Un modèle de système multi-agent pour l'émergence multi-niveau. *Technique et Science Informatiques*, 22(4) :235–247.
- Bossard, C., Kermarrec, G., Buche, C., et Tisseau, J. (2008). Transfer of learning in virtual environments : a new challenge? *Virtual Reality*.

- Bourgine, P. et Stewart, J. (2004). Autopoiesis and cognition. *Artificial Life*, 10 :327–345.
- Bown, O. et Wiggins, G. (2005). Modelling musical behaviour in a cultural-evolutionary system. Dans *Computational Creativity Workshop, IJCAI*, Edinburgh.
- Braitenberg, V. (1986). *Vehicles : Experiments in synthetic psychology*. Bradford Books. The MIT press.
- Bratman, M., Israel, D., et Pollack, M. (1988). Plans and resource-bounded practical reasoning. *Computational intelligence*, 4(4) :349–355.
- Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of robotics and automation*, 2(1) :14–23.
- Brooks, R. A. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1&2) :3–15.
- Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence*, 47 :139–159.
- Buche, C., Querrec, R., De Loor, P., Chevaillier, P., et Tisseau, J. (2009). PEGASE : un système tutoriel intelligent générique et adaptatif en environnement virtuel. *Revue des Sciences et Technologies de l'Information, série Techniques et Sciences Informatiques (RSTI-TSI)*, 28(8) :1051–1076.
- Carabelea, C., Boissier, O., et Florea, A. (2003). Autonomie dans les systèmes multi-agents : Essai de classification. *Revue des sciences et technologies de l'information*, 22 :191–204.
- Collins, C. et Bach-Y-Rita, P. (1973). Transmission of pictorial information through the skin. *Advances in biological and medical physics*, 14 :285.
- Cornuéjols, A. et Miclet, L. (2002). *Apprentissage artificiel. Concepts et algorithmes*. Eyrolles.
- Cussat-Blanc, S., Luga, H., et Duthen, Y. (2009). *Artificial Life : Borrowing from Biology*, volume 5865 of *Lecture Notes in Computer Science*, chapitre Making a Self-feeding Structure by Assembly of Digital Organs, pages 116–125. Springer LNCS.
- De Loor, P. (2006). *Autonomisation de modèles pour les simulations participatives*. Mémoire d'habilitation à diriger des recherches, Université de Bretagne Occidentale.
- De Loor, P., Lannuzel, A., Manac'h, K., et Harrouet, F. (2007). Cognition dynamique artificielle appliquée à l'autonomisation d'un humanoïde virtuel. Dans *ARCo'07, Colloque de l'Association pour la Recherche Cognitive. Cognition-Complexité-Collectif. Communications affichées*.
- De Loor, P., Manac'h, K., et Tisseau, J. (2009). Enaction-based artificial intelligence : Toward co-evolution with human in the loop. *Minds and Machines*, 19 :319–343.
- De Loor, P., Windelschmidt, C., Martinaud, K., et Cabioch, V. (2010). Connecting theater and virtual reality with cognitive sciences. Dans Richir, S. et Shirai, A., éditeurs, *Proceedings of Virtual Reality International Conference (VRIC 2010)*.
- Deb, K. (1999). Multi-objective genetic algorithms : Problem difficulties and construction of test problems. *Evolutionary computation*, 7(3) :205–230.

- Deb, K., Pratap, A., Agarwal, S., et Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm : Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2) :182–197.
- Declerck, G. (2010). *Phénoménologie et psychologie du tangible. Éléments pour une théorie de la valeur cognitive et pratique de la résistance*. Thèse de doctorat, Université de Technologie de Compiègne.
- Dellaert, F. et Beer, R. (1994). Toward an evolvable model of development for autonomous agent synthesis. *Artificial Life IV, Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*.
- Demazeau, Y. (1995). From interactions to collective behaviour in agent-based systems. Dans *Proceedings of the 1st European Conference on Cognitive Science, Saint-Malo*. Citeseer.
- Di Paolo, E. (2000). Homeostatic adaptation to inversion in the visual field and other sensorimotor disruptions. Dans Meyer, J., Berthoz, A., Floreano, D., Roitblat, H., et Wilson, S., éditeurs, *From Animals to Animats 6. Proceedings of the VI International Conference on Simulation of Adaptive Behavior*, pages 440–449.
- Di Paolo, E. (2003). *Dynamical System Approach to Embodiment and Sociality*, chapitre Organismically-inspired Robotics : Homeostatic Adaptation and Natural Teleology Beyond the Closed Sensorimotor Loop, pages 19–42. Adelaide, Australia : Advanced Knowledge international.
- Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4) :429–452.
- Di Paolo, E. et Harvey, I. (2003). Decisions and noise : the scope of evolutionary synthesis and dynamical analysis. *Adaptive Behavior*, 11(4) :284.
- Di Paolo, E., Rohde, M., et De Jaegher, H. (2010). *Enaction : Towards a New Paradigm for Cognitive Science*, chapitre Horizons for the Enactive Mind : Values, Social Interaction, and Play. MIT Press, Cambridge, MA.
- Di Paolo, E. A. et Iizuka, H. (2008). How (not) to model autonomous behaviour. *BioSystems*, Special issue on Modelling Autonomy.
- Doncieux, S., Mouret, J.-B., et Bredeche, N. (2009). Exploring new horizons in evolutionary design of robots. Dans *IROS Workshop "Exploring New Horizons in Evolutionary Design of Robots"*, pages 5–12, Saint Louis, USA.
- Doursat, R. (2009). Facilitating evolutionary innovation by developmental modularity and variability. Dans *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, pages 683–690. ACM.
- Dreyfus, H. (1984). *Intelligence artificielle : mythes et limites*. Flammarion.
- Egbert, M. D. et Di Paolo, E. (2009). Integrating autopoiesis and behavior : An exploration in computational chemo-ethology. *Adaptive Behavior*, 17(5) :367–386.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14 :179–211.

- Fahlman, S. et Lebiere, C. (1991). The cascade-correlation learning architecture.
- Froese, T., Virgo, N., et Izquierdo, E. (2007). Autonomy : a review and a reappraisal. Dans *Proceedings of the 9th European conference on Advances in artificial life*, pages 455–464. Springer-Verlag.
- Froese, T. et Ziemke, T. (2009). Enactive artificial intelligence : Investigating the systemic organization of life and mind. *Artif. Intell.*, 173(3-4) :466–500.
- Funahashi, K. et Nakamura, N. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6 :801–806.
- Gaussier, P. et Zrehen, S. (1995). Perac : A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16(2-4) :291 – 320. Moving the Frontiers between Robotics and Biology.
- Gigliotta, O. et Nolfi, S. (2008). On the coupling between agent internal and agent/environmental dynamics : Development of spatial representations in evolving autonomous robots. *Adaptive Behavior*, 16(2-3) :148–165.
- Giraud, M.-D. (2006). *Le traité de la réalité virtuelle, Volume 1 : L’homme et l’environnement virtuel*, chapitre Invariants et variabilité : antinomie ou complémentarité, pages 151–170. Mines, Paris.
- Gruau, F., Whitley, D., et Pyeatt, L. (1996). A Comparison between Cellular Encoding and Direct Encoding for Genetic Neural Networks. Dans *Proceedings of the First Annual Conference on Genetic Programming*, pages 81–89. Genetic Programming, MIT Press.
- Guillot, A. (1999). *Intelligence artificielle située, cerveau, corps et environnement*, chapitre Pour une approche dynamique des animats, pages 33–58. Alexis Drogoul et Jean-Arcady Meyer.
- Harvey, I. (2001). Artificial evolution : A continuing saga. Dans *Evolutionary Robotics. From Intelligent Robotics to Artificial Life*, volume 2217/2001 of *Lecture Notes in Computer Science*, pages 94–109. Springer Berlin / Heidelberg.
- Harvey, I. (2009). The microbial genetic algorithm. Dans et al., G. K., éditeur, *Proceedings of the Tenth European Conference on Artificial Life*. Springer LNCS.
- Harvey, I., Husnabds, P., et Cliff, D. (1994). Seeing the light : artificial evolution, real vision. Dans Meyer, J.-A. et Wilson, S., éditeurs, *From Animals to Animats 3*, Proceedings of third International Conference on Simulation of Adaptive Behavior, SAB94, pages 392–401, Boston MA. MIT Press/Bradford Books.
- Harvey, I., Paolo, E. D., Wood, R., Quinn, M., et Tuci, E. (2005). Evolutionary robotics : A new scientific tool for studying cognition. *Artificial Life*, 11(1-2) :79–98.
- Hassas, S. (2005). Engineering complex adaptive systems using situated mas : Some selected works and contributions. Dans *Sixth International Workshop on Engineering Societies in the Agents’ World (ESAW 2005)*, number 3963, Kusadasi, Aydin, Turkey. LNCS Springer Verlag, LNAI.
- Haykin, S. (2009). *Neural Networks and Learning Machines*. Pearson Education, third édition.

- Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press. Republished by the MIT press, 1992.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8) :2554.
- Husbands, P., Smith, T., Jakobi, N., et O´Shea, M. (1998). Better living through chemistry : Evolving GasNets for robot control. *Connection Science*, 10(3) :185–210.
- Iizuka, H. et Di Paolo, E. (2007). Toward spinozist robotics : Exploring the minimal dynamics of behavioural preference. *Adaptive Behavior*, 15 :359 :376.
- Izquierdo, E. (2008). *The dynamics of learning behavior : a situated, embodied, and dynamical systems approach*. Thèse de doctorat, University of Sussex, Brighton, United Kingdom.
- Izquierdo, E., Harvey, I., et Beer, R. (2008). Associative learning on a continuum in evolved dynamical neural network. *Adaptive Behavior*, 16(6) :361–384.
- Jonas, H. (1966). *The phenomenon of life : Towards philosophical biology*. Northwestern University Press, Evanston, IL.
- Jordan, M. (1986). Serial order : A parallel distributed processing approach. Rapport technique, Institute for Cognitive Science, University of California, San Diego.
- Kuo, A. D. (1999). Stabilization of lateral motion in passive dynamic walking. *The International journal of robotics research*, 18(9) :917.
- Kupiec, J.-J. (2008). *L’origine des individus*. Le temps des sciences. Fayard.
- Langton, C. G., éditeur (1989). *Artificial Life*, volume VI of *Series : Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley Publishing Company, Inc., Redwood City, California. Proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems held September, 1987, in Los Alamos, New Mexico.
- Laughlin, R. (2005). *A Different Universe : Reinventing Physics From the Bottom Down*. Basic Book.
- Lenay, C. (1997). Mouvement et perception : médiation technique et constitution de la spatialisation. *Communication à l’école d’été de l’Association pour la Recherche Cognitive sur le mouvement*, pages 69–80.
- Lenay, C. (2002). *Epistémologie et technologies cognitives : la question de l’espace*. Habilitation à Diriger des Recherches - UTC.
- Maes, P. (1989). The dynamics of action selection. Dans *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, volume 2, pages 991–997.
- Manac’h, K. et De Loor, P. (2007). Passage du discret au continu : identification des verrous pour la simulation de systèmes autopoïétiques. Dans *ARCo’07, Colloque de l’Association pour la Recherche Cognitive. Cognition-Complexité-Collectif. Communications affichées*.
- Maturana, H. R. et Varela, F. J. (1973). *Autopoiesis : The Organization of the Living*. Reidel, Dordrecht.

- Maturana, M. et Varela, F. (1984). *Autopoiesis and Cognition*. Reidel, Dordrecht.
- McCulloch, W. et Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5(4) :115–133.
- McGann, M. (2010). Perceptual modalities : Modes of presentation or modes of action? *Journal of Consciousness Studies*, 17(1-2).
- McGeer, T. (1990). Passive dynamic walking. *International Journal of Robotics Research*, 9(2) :62–82.
- McMullin, B. (2004). 30 years of computational autopoiesis : A review. *Artificial Life*, 10(3) :347–361.
- Meyer, J. (1996). Pour une approche complémentaire de l'ia traditionnelle : le manifeste animat. *In Cognito*, 6 :1–4.
- Mouret, J. et Doncieux, S. (2009). Using behavioral exploration objectives to solve deceptive problems in neuro-evolution. Dans *Proceedings of the 11th Annual conference on Genetic and evolutionary computation, GECCO'09*, pages 627–634. ACM.
- Müller, J. P. et Pischel, M. (1994). An architecture for dynamically interacting agents. *International Journal of Cooperative Information Systems (IJCIS)*, 3(1) :25–45.
- Nolfi, S. et Floreano, D. (1998). How co-evolution can enhance the adaptive power of artificial evolution : Implications for evolutionary robotics. Dans *Proceedings of the First European Workshop on Evolutionary Robotics*, Lecture Notes in Computer Science, pages 22–38. Springer-Verlag.
- Pearlmutter, B. (1990). Dynamic recurrent neural networks. Rapport technique CMU-CS-90-196, School of Computer Science, Carnegie Mellon University.
- Pfeifer, R., Iida, F., et Bongard, J. (2005). New robotics : Design principles for intelligent systems. *Artif. Life*, 11(1-2) :99–120.
- Pfeifer, R. et Scheier, C. (1999). *Understanding Intelligence*. MIT Press.
- Prepin, K. et Revel, A. (2007). Human-Machine interaction as a model of Machine-Machine interaction : how to make machines interact as humans do. *Advanced Robotics*, 21(15) :1709–1723.
- Revel, A. et Gaussier, P. (2003). *Biologically inspired robot behavior engineering*, volume 109 of *Studies in fuzziness and soft computing*, chapitre Designing neural control architectures for an autonomous robot using vision to solve complex learning tasks, pages 299–350. Physica-Verlag Heidelberg, Heidelberg, Germany.
- Rohde, M. et Stewart, J. (2008). Ascriptional and 'genuine' autonomy. *Biosystems*, 91(2) :424–433.
- Rumelhart, D., Hinton, G., et Williams, R. (1986). Learning internal representations by error propagation. *Parallel distributed processing : Explorations in the microstructure of cognition*, 1 :319.

- Schaffer, J. et Grefenstette, J. (1985). Multi-objective learning via genetic algorithms. Dans *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 593–595.
- Simon, H. (1996). *The sciences of the artificial*. The MIT Press, third édition.
- Simonin, O. et Ferber, J. (2000). Modeling Self Satisfaction and Altruism to handle Action Selection and Reactive Cooperation. Dans *The Sixth International Conference on the Simulation of Adaptive Behavior FROM ANIMALS TO ANIMATS*, volume 6, pages 314–323.
- Sims, K. (1994). Evolving 3D morphology and behavior by competition. *Artificial Life*, 1(4) :353–372.
- Soltoggio, A. (2008). Neuromodulation increases decision speed in dynamics environments. Dans Schlesinger, M., Berthouze, L., et Balkenius, C., éditeurs, *Proceedings of the Eighth International Conference on Epigenetic Robotics : Modeling Cognitive Development in Robotic Systems*, pages 119–126. Lund University Cognitive Studies.
- Spector, L. et Klein, J. (2005). *Genetic programming theory and practice III*, chapitre Trivial geography in genetic programming, pages 109–123. Genetic Programming. Springer.
- Srinivas, N. et Deb, K. (1995). Multiobjective function optimization using nondominated sorting genetic algorithms. *Evol. Comput.*, 2(3) :221–248.
- Stanley, K. et Miikkulainen, R. (2002). Evolving Neural Networks through Augmenting Topologies. *Evolutionary Computation*, 10(2) :99–127.
- Stratton, G. (1897). Upright vision and the retinal image. *Psychological Review*, 4(2) :182–187.
- Sutton, R. et Barto, A. (1998). *Reinforcement learning : An introduction*. The MIT press.
- Tani, J. (1996). Model based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Trans. Syst. Man and Cybern. B*, 26(3) :421–436.
- Taylor, J. (1962). *The behavioural basis of perception*. New Haven : Yale University press.
- Tencé, F., Buche, C., De Loor, P., et Marc, O. (2010). The challenge of believability in video games : Definitions, the challenge of believability in video games : Definitions, agents models and imitation learning. Dans Mao, W. et Vermeersch, L., éditeurs, *2nd Asian Conference on Simulation and AI in Computer Games (GAMEON-ASIA '10)*, pages 38–45. Eurosis.
- Tisseau, J. (2001). *Réalité Virtuelle - autonomie in virtuo*. Habilitation à diriger des recherche, Université de Rennes I.
- Tuci, E., Quinn, M., et Harvey, I. (2002). An Evolutionary Ecological Approach to the Study of Learning Behavior Using a Robot-Based Model. *Adaptive Behavior*, 10(3-4) :201–221.
- Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society London*, 237 :37–72.
- Varela, F. (1988). *Invitation aux sciences cognitives*. Editions du seuil, Paris.

- Varela, F. (1989). *Autonomie et connaissance*. Seuil, traduction de "Principles of Biological Autonomy" par P. Bourguin et P. Dumouchel.
- Varela, F., Thompson, E., et Rosch, E. (1993). L'inscription corporelle de l'esprit. *Le Seuil*, 235 :214.
- Von Neumann, J., Burks, A., et al. (1966). Theory of self-reproducing automata.
- Waddington, C. H. (1940). *Organisers and Genes*. Cambridge : Cambridge University Press.
- Weber, A. et Varela, F. (2002). Life after Kant : Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2) :97–125.
- Wood, R. et Di Paolo, E. (2007). New models for old questions : Evolutionary robotics and the 'a not b' error. Dans *Proceedings of the 9th European Conference on Artificial Life ECAL 2007*. Springer-Verlag.
- Yamauchi, B. et Beer, R. (1994a). Integrating reactive, sequential, and learning behavior using dynamical neural networks. Dans *From Animals to Animats 3 : Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, pages 382–391.

Annexe A

Comportement et invariants sensori-moteurs

NOUS donnons deux exemples de comportements sensori-moteurs pour lesquels nous cherchons des invariants sensori-moteurs. Pour obtenir, ces deux exemples nous programmons le déplacement d'un agent. Nous les présentons dans un but illustratif pour expliciter ce que nous cherchons à obtenir.

Déplacement selon une figure carrée

Nous avons simulé un agent artificiel dont nous contrôlons les déplacements de manière à ce qu'il effectue un carré dans un espace en deux dimensions. Il ne s'agit ici pas d'un agent éactif, son comportement est programmé pour suivre une trajectoire précise. L'environnement contient deux repères visuels. La figure 4.3 page 69 représente la trajectoire réalisée par l'agent relativement aux repères.

	A	B	C	D
Temps de passage	137	548	944	1340
	1737	2133	2529	2926

TAB. A.1 – Temps de passage de l'agent aux positions A, B, C et D.

L'agent a un capteur pour chacune des balises et deux moteurs. Pendant la simulation, les données captées sont enregistrées ainsi que les commandes des moteurs. Les valeurs captées sont calculées en fonction de la distance aux balises de la manière suivante : $I_i = gain * (1 - d_i/D)$, où I_i est la valeur du capteur i , d_i la distance de l'agent à la balise i , D la portée du capteur (ici 100) et le *gain* des capteurs est le même (ici 5). La figure A.1 trace les valeurs enregistrées sur les capteurs en fonction de l'activité des moteurs. Chacune des sous-figures correspond à un intervalle temporel différent. Nous avons relevé dans le tableau A.1 les instants auxquels l'agent passe aux positions notées A,B,C et D sur la

figure, correspondant aux coins du carré. Les intervalles de temps choisis pour les sous-figures correspondent aux étapes suivantes :

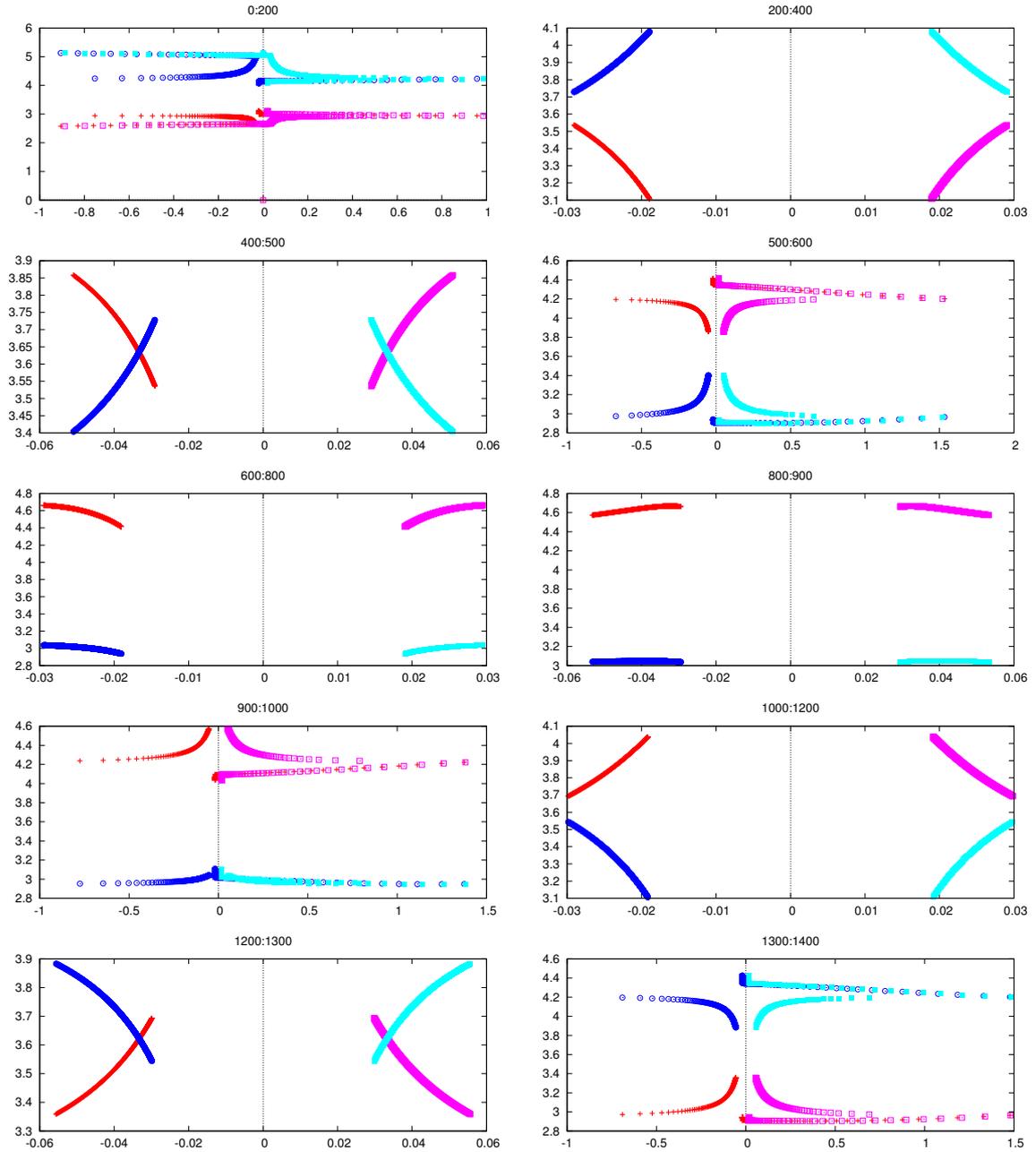
- ▷ Passage d'une position. Par exemple, pour le passage de la position B au temps 548, voir la figure A.1(a) ligne 2 colonne 2 correspondant à l'intervalle $[t_0 + 500; t_0 + 600]$.
- ▷ Déplacement vers la position suivante. Par exemple, pour le premier déplacement de la position B à la position C, voir la figure A.1(a) ligne 3 colonne 1 correspondant à l'intervalle $[t_0 + 600; t_0 + 800]$.
- ▷ Arrivée à proximité de la position suivante. Par exemple, pour l'arrivée à proximité de la position C, voir la figure A.1(a) ligne 3 colonne 2 correspondant à l'intervalle $[t_0 + 800; t_0 + 900]$.

La légende de la figure est donnée ci-contre. Les valeurs des moteurs sont en abscisses, les valeurs des capteurs en ordonnées. Pour chaque sous-figure, l'intervalle de temps utilisé est indiqué au dessus du graphique. Les échelles sont ajustées pour chacune des figures, il y a donc des écarts importants ($\times 10^{-2}$ pour les abscisses entre la première et la deuxième figure par exemple).

+	Valeurs enregistrées par le capteur de la balise placée en $(50, 0)$ en fonction de la commande du moteur gauche.
⊙	Valeurs enregistrées par le capteur de la balise placée en $(0, 0)$ en fonction de la commande du moteur gauche.
□	Valeurs enregistrées par le capteur de la balise placée en $(50, 0)$ en fonction de la commande du moteur droit.
■	Valeurs enregistrées par le capteur de la balise placée en $(0, 0)$ en fonction de la commande du moteur droit.

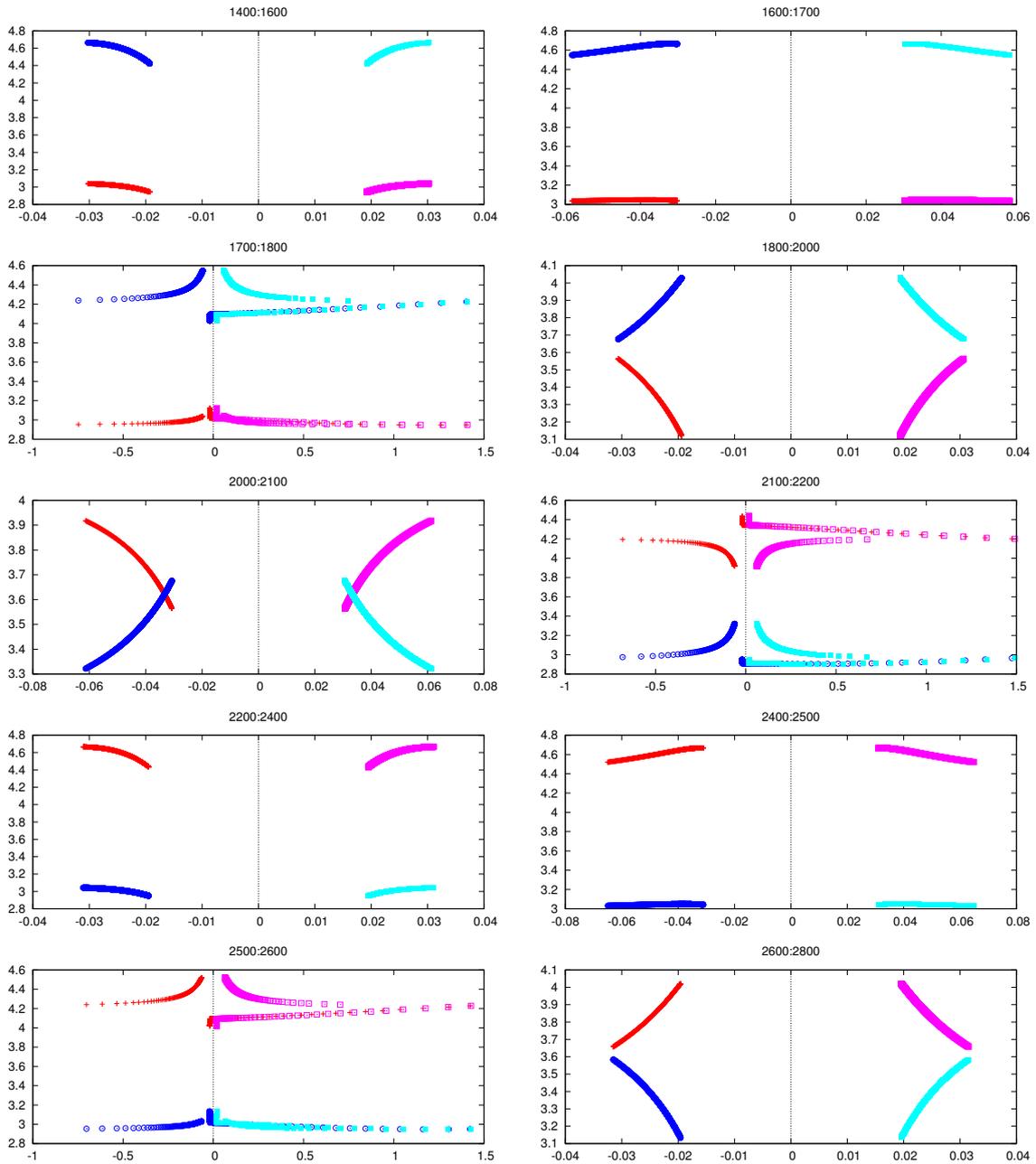
Nous remarquons que des motifs particuliers apparaissent dans le lien entre les actions et les sensations de l'agent. Ces motifs peuvent être mis en relation avec le comportement de l'agent. Ainsi, à chaque étape de la figure correspond un type d'invariant sensori-moteur. Nous retrouvons les mêmes motifs aux différents instants correspondant au passage aux mêmes positions (A, B, C et D), ainsi que des motifs pour les déplacements correspondants aux côtés du carré.

Ainsi par exemple, entre $t_0 + 600$ et $t_0 + 800$ (figure A.1(a) ligne 3, colonne 1), l'agent vient de passer à la position B, les variations sont très faibles, le déplacement est régulier. Les valeurs sur les capteurs augmentent ce qui indique que l'agent se rapproche des balises. Le moteur droit accélère légèrement, les commandes motrices sont symétriques. Ensuite, entre $t_0 + 800$ et $t_0 + 900$ (figure A.1(a) ligne 3, colonne 2), l'agent s'éloigne des balises. Au temps $t_0 + 944$, l'agent atteint la position C. Nous remarquons que les relations sensori-motrices autour de cet instant changent considérablement ($[t_0 + 900; t_0 + 1000]$ figure A.1(a) ligne 4, colonne 1). Il y a plus de variations sur les moteurs et sur les capteurs. Plus globalement, il y a une symétrie des relations sensori-motrices pendant le déplacement de l'agent le long des côtés opposés du carré. Nous retrouvons cette symétrie pour les virages. Tout ceci est bien entendu attendu, la régularité étant d'autant plus forte que le déplacement est contrôlé. Cela met en évidence l'existence d'invariants sensori-moteur qui signent un comportement.



(a) Partie 1

FIG. A.1 – Valeurs sur les capteurs en fonction des commandes des moteurs sur différents intervalles de temps pendant le déplacement d'un agent effectuant la figure 4.3. La légende est donnée dans le texte.



(b) Partie 2

FIG. A.1 – Valeurs sur les capteurs en fonction des commandes des moteurs.

Déplacement selon une figure en forme de sablier

Nous avons ensuite simulé un agent réalisant une figure en forme de sablier. L’environnement et le type d’agent sont les mêmes que pour la trajectoire précédente. Nous contrôlons toujours le déplacement de l’agent et enregistrons les valeurs sur ses capteurs et les commandes motrices pendant tout l’épisode.

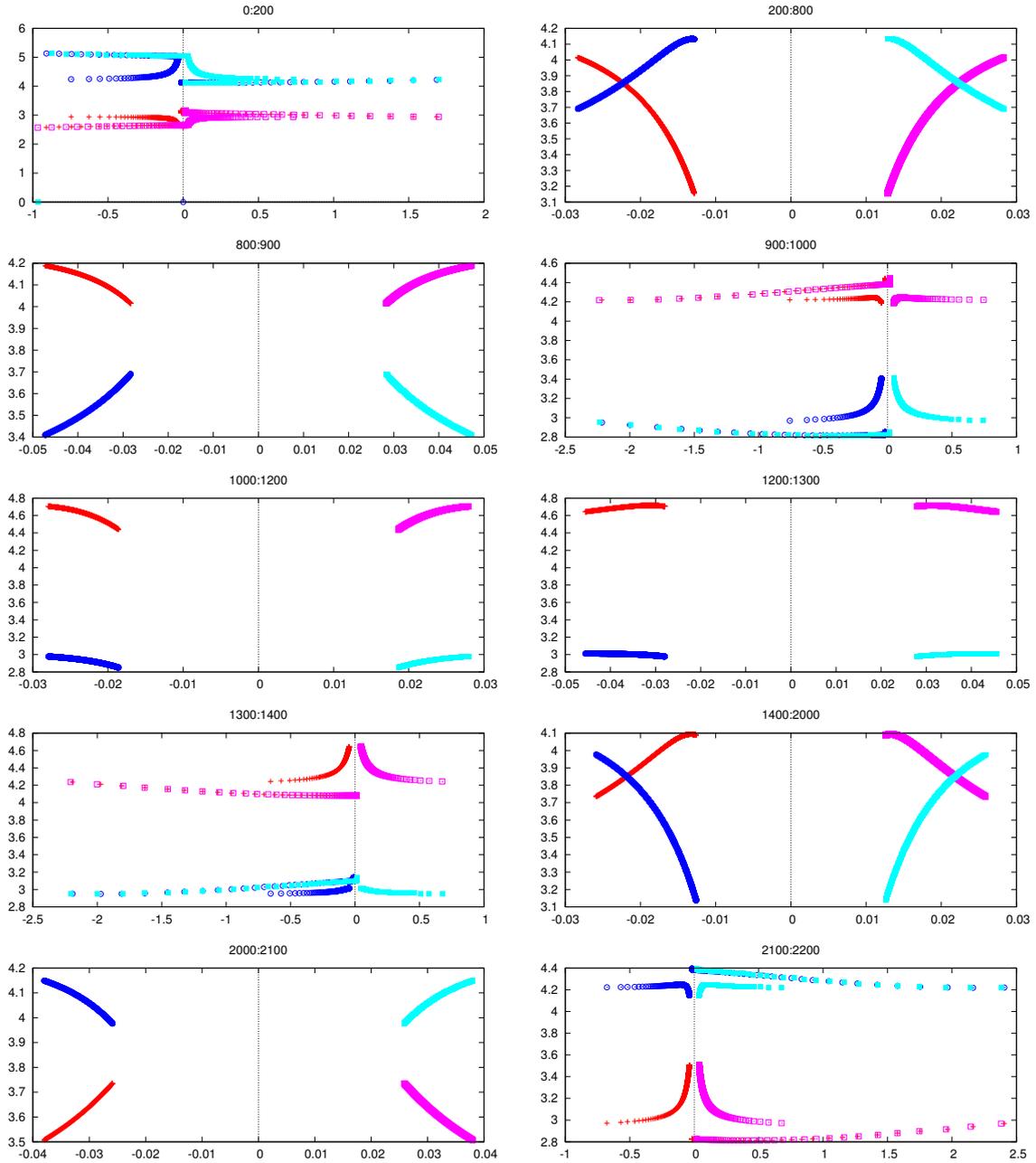
La figure 4.4 page 69 représente la trajectoire réalisée par l’agent relativement aux repères. Les instants où l’agent passe à chacun des “coins” du sablier, à savoir les positions A, B, C et D, sont recensés dans le tableau A.2. La figure A.2 trace les données enregistrées

	A	B	C	D
Temps de pas-	137	956	1360	2186
sage	2591			

TAB. A.2 – Temps de passage de l’agent aux positions A, B, C et D.

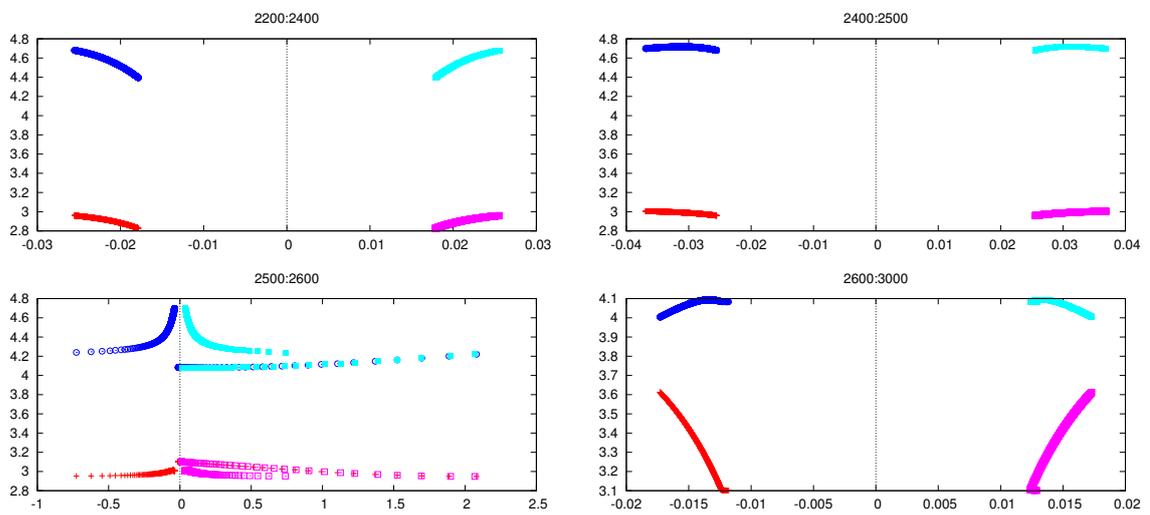
sur les capteurs en fonction de l’activité des moteurs. Chacune des sous-figures correspond à un intervalle temporel différent. La légende est la même que celle utilisée pour la figure A.1. Comme pour le carré, nous remarquons des motifs particuliers dans le lien entre les actions et les sensations de l’agent. Ces motifs correspondent aux passages aux positions A, B, C et D et aux trajets entre ces positions.

Les relations sensori-motrices mises en évidence ici ne sont pas très différentes de celles identifiées pour la trajectoire du carré. En particulier, nous retrouvons globalement les mêmes motifs pour les déplacements entre les différentes positions. Mais si au niveau global, l’aspect des relations est le même pour le sablier que pour le carré, il existe pourtant des différences. Les plus remarquables sont observables surtout au moment des virages, et pour cause, pour le carré, l’agent ne tourne que vers la gauche alors que, pour le sablier, il fait des virages à droite et à gauche. Nous pouvons aussi observer que les déplacements sur les diagonales correspondent à des relations sensori-motrices plus complexes. La symétrie n’est plus respectée. En effet, les relations entre la commande des moteurs et les valeurs de chacun des capteurs sont bien distinctes sur les trajets A-B et C-D. Par exemple, entre $t_0 + 200$ et $t_0 + 800$ (figure A.2(a) ligne 1 colonne 2), qui correspond à une partie du trajet A-B, les valeurs sur le capteur A augmentent bien plus que ne diminuent celles sur le capteur B.



(a) Partie 1

FIG. A.2 – Valeurs enregistrées sur les capteurs en fonction des commandes des moteurs sur différents intervalles de temps pendant le déplacement d'un agent contrôlé pour effectuer la figure 4.4. La légende est donnée dans le texte.



(b) Partie 2

FIG. A.2 – Variations des capteurs en fonction des variations des moteur.

Annexe B

Simulations dans un environnement en deux dimensions

NOUS présentons ici le résultat de quelques simulations liées à l'expérience décrite en détail dans la section 5.1. Ici, nous modifions uniquement les conditions initiales ou la durée des essais.

La figure B.1 représente la version non lissée de la figure 5.3. L'agent tourne sur place au début d'un essai et à la fin. Les déplacements se font en "zig-zag". Cela provient du fait que l'un de ses moteurs oscille en permanence comme le montre la figure B.2.

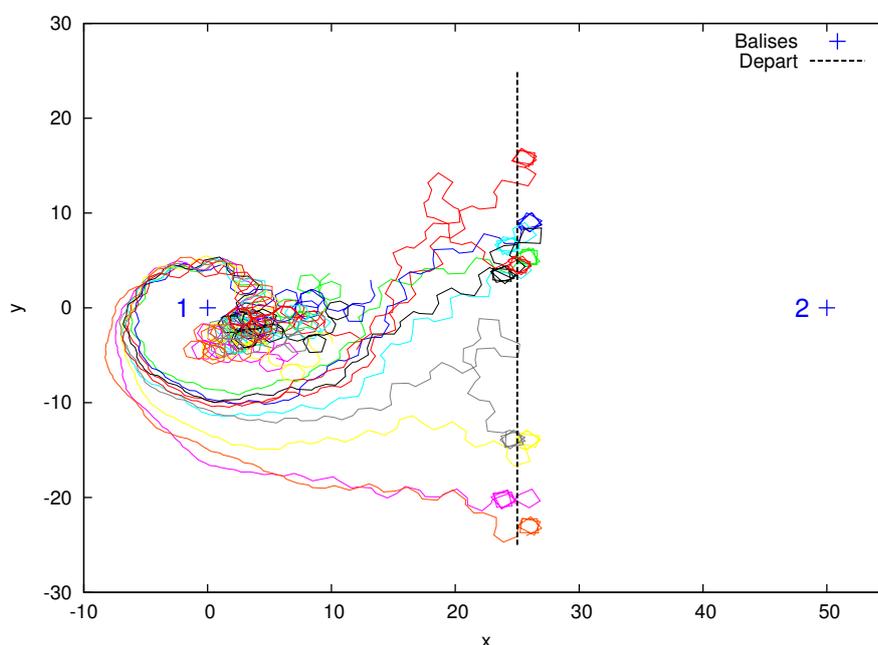


FIG. B.1 – Trajectoires non lissées du meilleur agent lors de 10 essais dans les conditions de la phase 3.

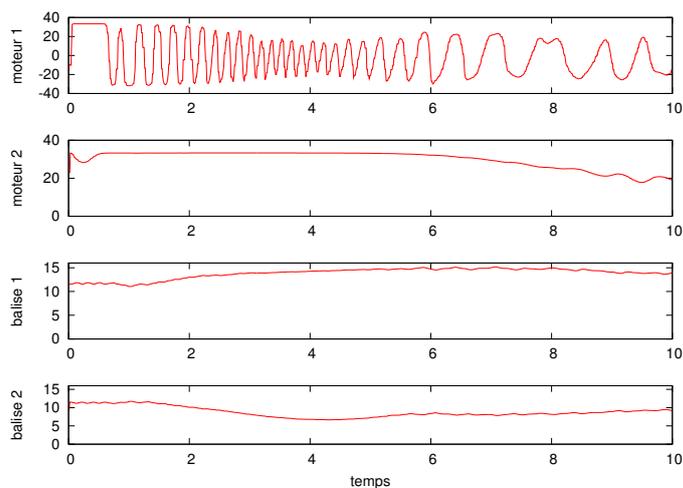


FIG. B.2 – Activité des moteurs et des capteurs de l’agent pendant l’essai représenté en noir sur la figure B.1.

La figure B.3 représente la version non lissée de la figure 5.4. Comme précédemment, l’agent tourne sur place au début d’un essai, il continue ensuite à se déplacer en faisant des cercles. L’observation de son activité sensori-motrice lors d’un essai (cf. figure B.4) nous montre un moteur qui oscille alors que l’autre est saturé. Il est à noter que la balise 2 est toujours captée.

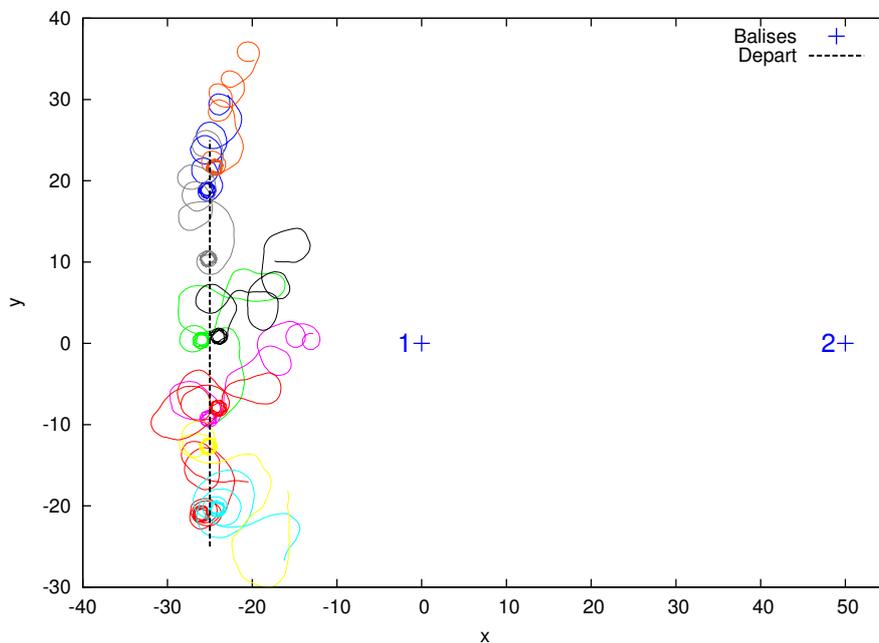


FIG. B.3 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3 en modifiant uniquement les conditions initiales liées à la balise 2.

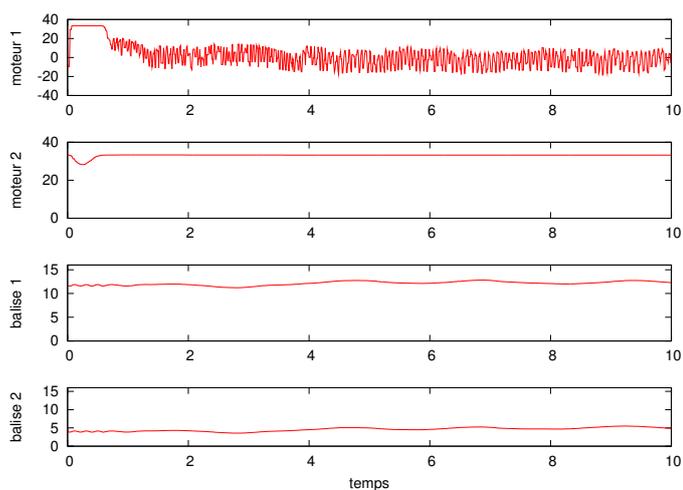


FIG. B.4 – Activité des moteurs et des capteurs de l’agent pendant l’essai représenté en noir sur la figure B.3.

Les trajectoires représentées sur la figure B.5 correspondent à des simulations lors desquelles les conditions initiales par rapport aux deux balises sont modifiées. L’agent est en effet placé en (x, y) avec $0 < x < 50$ et $-25 < y < 25$. Dans la majorité des essais, l’agent approche la balise 1, excepté lors d’un essai pour lequel il est placé “loin” de la balise 2.

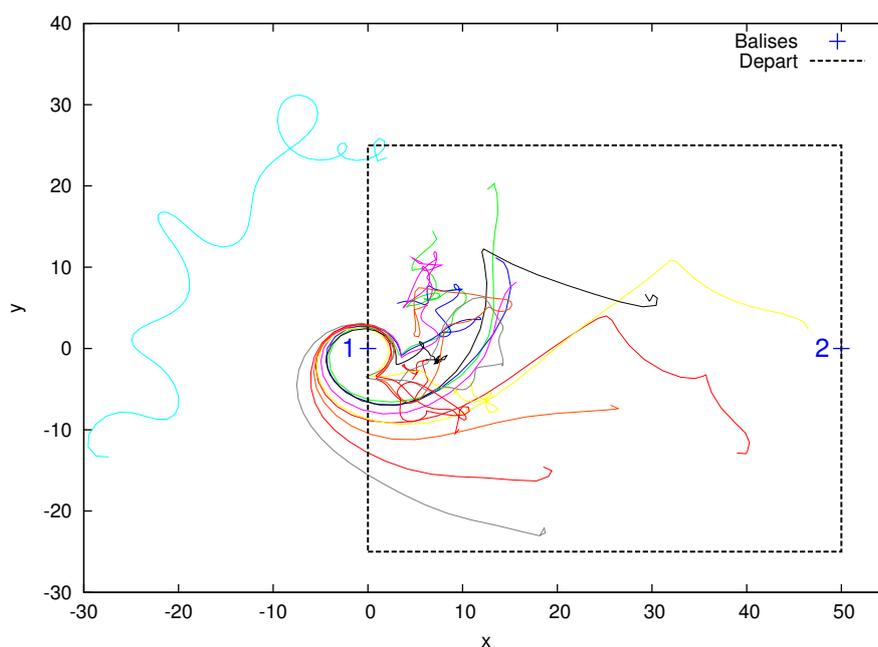


FIG. B.5 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3 en modifiant les conditions initiales liées aux deux balises.

Lors des essais représentés sur la figure B.6, l'agent est initialement placé en (x, y) avec $-10 < x < 60$ et $-25 < y < 25$. Dans la majorité des essais, l'agent approche la balise 1, excepté lors des essais pour lequel il est placé "loin" de l'une des balises.

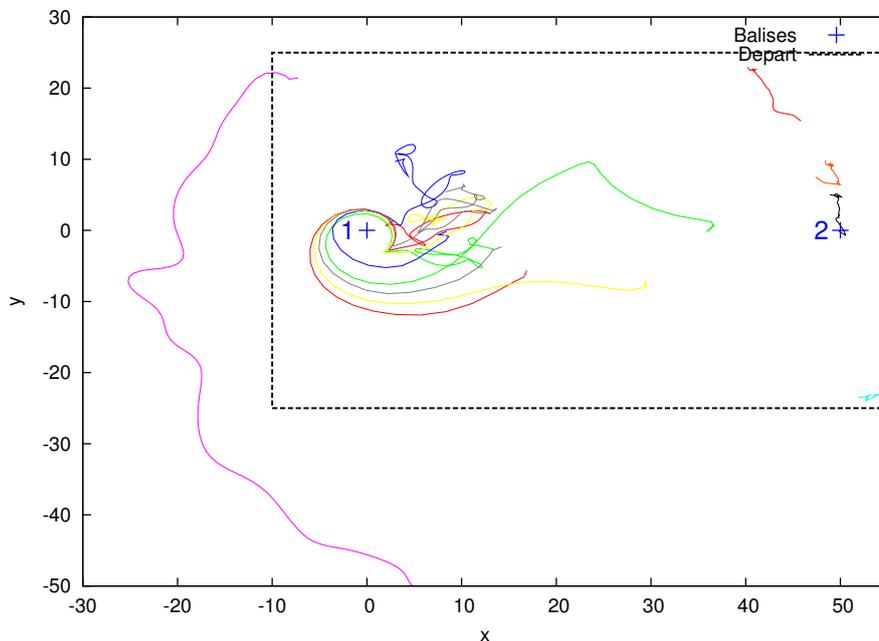


FIG. B.6 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3 en modifiant les conditions initiales liées aux deux balises.

La figure B.7 représente les trajectoires d'un agent pour des essais au début desquels il est placé dans les mêmes conditions que pour ceux représentés sur la figure 5.3 (qui correspondent aux conditions initiales utilisées par l'algorithme évolutionnaire). Le temps de simulation est ici doublé et nous observons qu'à la fin des essais l'agent est toujours proche de la balise 1, ce qui semble indiquer qu'après avoir approché la balise 1, il ne se dirige pas vers la balise 2.

La figure B.8 représente les trajectoires d'un agent pour des essais au début desquels il est placé en (x, y) avec $x = 0$ et $y = -25$ ou $y = 25$. Pour tous les essais effectués où l'agent part de la position $(0, -25)$, il se dirige vers la balise 1. Lorsqu'il est placé en $(0, 25)$, il approche de la balise 1 lors de deux essais, alors que lors des autres essais ce n'est pas le cas, il part dans une autre direction qui reste toutefois qualitativement toujours la même. La seule variable est ici l'orientation. Il semble que la limite de la distance initiale à la balise 2 permettant à l'agent d'approcher la balise 1 soit proche de celle utilisée ici. La direction dans laquelle l'agent part est alors importante.

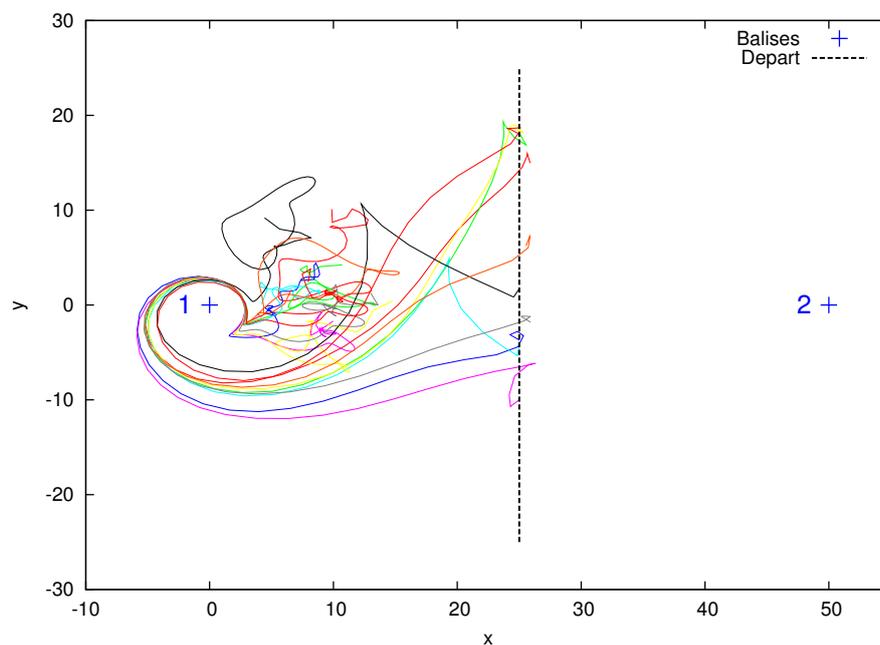


FIG. B.7 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3. Le temps de simulation est doublé par rapport à la figure B.1.

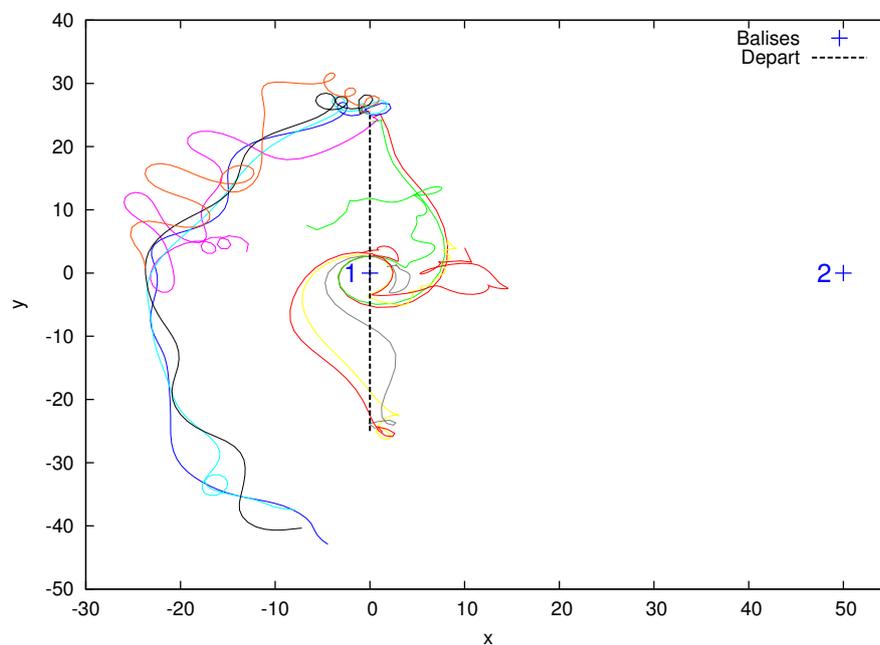


FIG. B.8 – Trajectoires du meilleur agent lors de 10 essais dans les conditions de la phase 3 en modifiant uniquement les conditions initiales liées à la balise 2.

Cette thèse propose de s'appuyer sur le paradigme de l'énaction des sciences cognitives pour aborder la modélisation de l'autonomie d'entités artificielles. Dans un contexte d'interaction entre des humains et des entités autonomes en réalité virtuelle (thème fédérateur de mon équipe d'accueil), les propriétés comportementales résultantes de ce type d'autonomie dite *constitutive*, pourrait favoriser la créativité et l'engagement en proposant à l'utilisateur un compromis entre surprise et régularité. Cette thèse est structurée en deux parties, la première est une proposition théorique s'appuyant sur les liens entre l'énaction et l'autonomie d'agents artificiels interactifs qui mène à une définition d'agents virtuels situés dits énatifs. La deuxième partie est une tentative d'implémentation issue de cette réflexion à l'aide d'un modèle de cognition dynamique paramétré selon une approche évolutionnaire.

Partie I : Proposition théorique

Chapitre 1 Autonomie constitutive : de la biologie à l'artificiel

L'objectif de ce chapitre est de présenter comment la notion d'autonomie constitutive peut être abordée dans le domaine de l'artificiel.

Pour cela, la première partie présente les fondements biologiques de l'autonomie selon la théorie de l'énaction. Cette théorie ancre la cognition dans l'autonomie d'un système en mettant l'accent sur l'organisation circulaire des organismes vivants : pour Varela, les systèmes autonomes sont caractérisés par le fait qu'ils maintiennent de manière active leur organisation dans des conditions précaires. L'organisation d'un système autonome n'est définie qu'en lien avec des éléments internes, il n'existe donc pas d'entrées et de sorties du point de vue du système, mais des perturbations pour une organisation qui résiste et se maintient. La conséquence est que le comportement d'un système autonome, assimilé à ses facultés cognitives, est indissociable de l'histoire de son couplage avec l'environnement.

Dans la deuxième partie du chapitre, nous nous intéressons à la vie artificielle qui est une piste pour la simulation d'entités dotées d'une autonomie constitutive car elle reproduit des mécanismes biologiques qui permettent la mise en place d'une organisation circulaire. Ainsi, les automates autopoïétiques reproduisent le maintien d'une organisation malgré des perturbations. La production d'une entité à partir de la définition d'une organisation est quant

à elle abordée par l'embryogenèse artificielle. Cependant, si le problème de la constitution et du maintien de l'identité est adressé, les systèmes de ce type ne peuvent pas être qualifiés d'agents dans la mesure où ils "n'agissent" pas et surtout où l'interaction avec l'environnement reste très faible.

La troisième partie développe une seconde piste, celle de la conception d'agents, capables de comportements sensorimoteurs. Le courant de la *Nouvelle Intelligence Artificielle* préconise de concevoir des agents incarnés, c'est-à-dire pour lesquels l'environnement est son propre modèle. L'approche *Animat* introduit l'idée d'une causalité circulaire comme nécessaire à la simulation d'entités véritablement situées. Ceci justifie l'adoption d'une approche dynamique, dans laquelle la modélisation se fait selon l'évolution temporelle du système et non pas selon une analyse fonctionnelle ou même selon une analyse des interactions entre composants, pour l'étude des comportements autonomes, qui permet de donner une place importante à l'histoire du couplage de l'agent avec son environnement. La *Nouvelle Robotique* complète cette démarche en réhabilitant le corps dans la production du comportement. Enfin, le courant de la *Robotique inspirée organisme* applique les idées éenactives à la construction d'agents autonomes artificiels en insistant sur la nécessité d'une organisation autodéterminée, capable de générer et modifier ses propres buts.

En conclusion de ce chapitre, nous proposons une première caractérisation d'un agent éenactif, c'est-à-dire d'un agent simulé pour lequel nous chercherons à nous approcher le plus d'une autonomie constitutive. Nous retenons également que l'inter-activité que nous recherchons pour obtenir de véritables agents nous amène à faire une concession forte par rapport à la définition de l'autonomie constitutive telle que la donne Varela qui consiste à mettre de côté la génération de l'identité du système.

Chapitre 2 Choix techniques pour l'autonomie constitutive

Dans ce chapitre, nous cherchons à définir quels modèles et outils informatiques sont les plus adaptés à la synthèse d'agents dotés d'une autonomie constitutive. L'objectif est donc de proposer une caractérisation plus concrète d'un agent éenactif capable de s'intégrer dans une simulation interactive.

La première partie du chapitre s'intéresse au contrôleur de l'agent. Les architectures classiques d'agents sont généralement incompatibles avec les contraintes que nous définissons pour une autonomie constitutive. Ces agents doivent adopter un comportement autodéterminé par une organisation circulaire perturbable.

En revanche, les réseaux de neurones artificiels récurrents et en particulier, les CTRNN¹, qui font l'objet de la seconde partie du chapitre, possèdent une organisation circulaire dont le fonctionnement n'est pas défini en terme d'entrées/sorties, mais par des relations entre éléments internes associées aux interactions avec l'environnement qui agissent comme des perturbations. Les CTRNN peuvent théoriquement reproduire le comportement de n'importe

¹ Pour *Continuous Time Recurrent Neural Networks*.

quel système dynamique, dont celui d'un système chaotique. Un agent contrôlé par un CTRNN et couplé à l'environnement au moyen de capteurs et d'effecteurs répond alors aux contraintes que nous nous sommes fixées, c'est pourquoi notre choix d'une architecture pour des contrôleurs d'agents éactifs se porte sur les CTRNN.

La troisième partie du chapitre présente la robotique évolutionnaire qui consiste à utiliser des algorithmes évolutionnaires pour paramétrer des contrôleurs de robots ou d'agents simulés. Elle est confrontée à des problèmes spécifiques tels que l'influence des conditions d'évolution sur la robustesse ou encore la difficulté d'obtenir des solutions pour des tâches complexes. Des méthodes sont disponibles pour résoudre ces difficultés, comme l'approche incrémentale qui consiste à faire évoluer les contrôleurs en utilisant des tâches graduellement plus complexes. La co-évolution permet d'appliquer une pression sélective qui s'adapte aux progrès faits par l'évolution. L'approche multi-objectif est un autre exemple de technique permettant de faire évoluer des agents pour satisfaire différents objectifs. Enfin, la communauté de chercheurs s'intéressant à l'évolution d'agents artificiels contrôlés par des CTRNN utilise l'algorithme Microbial, que nous choisissons également.

Nous utilisons le terme d'approche dynamique évolutionnaire pour l'utilisation conjointe des CTRNN et des méthodes évolutionnaires que nous considérons comme les plus adaptés à la synthèse d'agents éactifs.

Chapitre 3 Argumentation pour la génération d'agents éactifs virtuels avec des CTRNN

Ce chapitre présente un état de l'art de l'utilisation des CTRNN. Il a pour objectif de comparer l'approche dynamique évolutionnaire et la notion d'agent éactif afin de préciser la spécificité de notre travail qui va consister à introduire un apprentissage à l'échelle ontogénétique au sein de ces CTRNN.

Ainsi, dans la première partie du chapitre, nous illustrons comment des CTRNN ont été utilisés dans une volonté d'obtenir un comportement proche d'une autonomie constitutive. Les propriétés comportementales obtenues sont intéressantes du point de vue du principe qui permet la production du comportement.

Dans la seconde partie du chapitre, nous nous intéressons à la capacité du modèle choisi à produire un comportement qui évolue. Les travaux que nous décrivons montrent que des agents contrôlés par des CTRNN peuvent apprendre des associations sans mécanisme particulier dédié à cet apprentissage. Nous considérons cependant que ces apprentissages résultent d'un mécanisme d'adaptation mis en place à l'échelle phylogénétique. Cela nous amène à faire une distinction entre apprentissage et adaptation, c'est-à-dire entre apprentissage "construit" et apprentissage "réflexe". L'apprentissage "construit" relève d'un mécanisme opérant à l'échelle ontogénétique. Il reflète l'histoire de l'interaction de l'agent avec l'environnement et doit permettre une co-évolution à cette même échelle ontogénétique.

Nous précisons alors notre définition d'un agent éactif qui reprend les contraintes

identifiées auparavant en ajoutant la capacité à enrichir un comportement par l'interaction.

Chapitre 4 Proposition d'une démarche pour l'apprentissage ontogénétique

L'objectif de ce chapitre est d'étayer notre proposition en construisant une démarche pour la mise en place d'un apprentissage ontogénétique chez des agents éactifs. Pour cela, nous nous intéressons à l'acquisition d'invariants sensorimoteurs à l'échelle ontogénétique. En effet, l'apprentissage que nous voulons mettre en place portera sur des comportements sensorimoteurs, qui sont considérés comme la base de toute connaissance dans la théorie de l'éaction.

Dans la première partie du chapitre, nous définissons et illustrons la notion d'invariants sensorimoteurs, directement liée pour les phénoménologues à la perception du monde et que nous utiliserons en tant qu'outil d'observation du comportement de nos agents. Nous considérons que deux comportements sensorimoteurs différents correspondent à des invariants sensorimoteurs différents et donc que l'apprentissage "construit" que nous cherchons à obtenir peut résulter en la transformation d'un invariant sensorimoteur en un autre.

La deuxième partie du chapitre porte sur la transformation en elle-même. Nous montrons comment la dynamique de l'agent et celle de l'environnement doivent être couplées de manière à s'entre-influencer pour atteindre un apprentissage à l'échelle ontogénétique. La transformation des invariants sensorimoteurs doit refléter l'histoire particulière de ce couplage, ce qui implique que l'environnement perturbe la dynamique de l'agent en lui laissant une "marge de manœuvre" et que réciproquement, les actions de l'agent modifient les réactions de l'environnement. L'interaction doit alors se faire en contraignant sans imposer le comportement de l'agent éactif. En conformité avec ces critères, nous proposons d'utiliser la notion de guidage.

Le guidage s'il est efficace induira une co-évolution puisque l'interaction sera adaptée au comportement de l'agent au fur et à mesure de son évolution. Ce chapitre se conclue par la mise en place d'un programme de validation de notre proposition en trois expériences.

Partie II : Expérimentations exploratoires

Chapitre 5 Évolution phylogénétique incrémentale de comportements sensorimoteurs

Ce chapitre correspond à la première étape qui consiste à complexifier un comportement sensorimoteur en utilisant une approche phylogénétique. L'enjeu est d'obtenir un comportement "régulier" le plus complexe possible en préservant les principes d'une autonomie constitutive qui implique donc des comportements difficiles à maîtriser.

Nous commençons par essayer d’obtenir des agents qui se déplacent de manière à dessiner des figures dans un environnement en 2D. Cette expérience n’aboutit pas. Ceci est surprenant au vu des travaux existants, nous donnons des pistes d’explications et décidons de traiter des tâches minimalistes mais qui préservent la possibilité de faire évoluer des invariants sensorimoteurs. Cette première tentative montre la difficulté que nous avons rencontré dans l’évolution des paramètres de CTRNN. Nous sommes cependant confortés dans notre démarche qui vise à minimiser la part phylogénétique de mise en place du comportement.

L’expérience minimaliste que nous réalisons ensuite montre qu’il est possible de faire évoluer de manière incrémentale un comportement sensorimoteur graduellement plus complexe en utilisant des agents contrôlés par des CTRNN. Concrètement, le comportement de départ consiste pour l’agent à s’approcher d’une cible dans un environnement en une dimension. La complexification de la tâche se fait en positionnant la cible à une distance arbitraire, en la rendant mobile et enfin en cherchant à ce que l’agent maintienne une distance fixe à cette cible. Nous mettons en évidence les invariants sensorimoteurs correspondants aux différents comportements à titre illustratif.

Nous concluons donc qu’il est possible de complexifier un comportement sensorimoteur à l’échelle phylogénétique et mettons en évidence les difficultés liées à l’approche dynamique évolutionnaire.

Chapitre 6 Apprentissage associatif ontogénétique guidé

Le chapitre 6 aborde la possibilité d’un apprentissage à l’échelle ontogénétique pour un agent contrôlé par un CTRNN.

La seconde expérience utilise donc le guidage pour réaliser un apprentissage de type associatif entre un type d’environnement et un comportement moteur. Il s’agit d’obtenir des agents qui soient capables de modifier leur comportement en fonction d’un guidage. Cela implique qu’il préexiste une réaction à ce guidage et que celle-ci provoque un “changement d’attitude” de l’agent par rapport à des stimuli indépendants. Nous décrivons la tâche à réaliser, les agents que nous simulons et le processus évolutionnaire que nous mettons en place.

Nous présentons ensuite les résultats de l’évolution. En particulier, les performances obtenues par le meilleur agent résultant du processus d’évolution sont détaillées. La dynamique de cet agent est étudiée. Nous mettons ainsi en évidence une évolution de la dynamique sensorimotrice et de la dynamique interne au cours de l’apprentissage. L’aspect sensorimoteur est cependant très faible dans cette expérience, car en l’absence du guidage il n’y a aucune relation entre les sensations et les actions de l’agent.

L’expérience présentée dans ce chapitre est encourageante dans notre démarche mais nous incite à aller plus loin en associant apprentissage ontogénétique guidé et évolution sensorimotrice.

Chapitre 7 Apprentissage ontogénétique guidé d'invariants sensorimoteurs

Ce chapitre présente une expérience qui vise à finaliser notre démarche en combinant les deux expériences précédentes, de manière à obtenir une complexification d'un comportement sensorimoteur résultant d'un apprentissage à l'échelle ontogénétique.

Il s'agit de transformer un invariant sensorimoteur. Le comportement de départ est le même que le comportement de départ de la première expérience. Cependant ici, la complexification obtenue dans l'expérience 1 par évolution artificielle est recherchée en utilisant le guidage et non pas par phylogénèse.

La méthode utilisée est présentée à travers une description de la tâche, de l'agent et du processus évolutif. Le comportement que nous obtenons n'est pas tout à fait celui que nous recherchons. Cependant, nous montrons à travers l'analyse comportementale et dynamique de certains agents, qu'il est possible d'obtenir une modification comportementale induite par le guidage à l'échelle ontogénétique.

Conclusions

La conclusion de ce mémoire présente un bilan critique des résultats obtenus et de la démarche utilisée. Les résultats illustrent la difficulté de la manipulation d'un système autonome. En effet, le modèle est réellement résistant. Nous avons montré certaines limites de l'approche évolutive dans un tel contexte. Les perspectives tiennent compte de la nécessité à faire des compromis entre complexité comportementale et autonomie constitutive. Nous discutons également de l'exploitation potentielle de ce travail en réalité virtuelle.

Mots clefs : autonomie constitutive, agents situés, éaction, cognition dynamique artificielle, apprentissage ontogénétique, approche évolutive, CTRNN

Towards the notion of enactive virtual agent: Application to the evolutionary dynamic approach

This PhD thesis relies on enaction paradigm to address the modeling of artificial entities autonomy. In a context of interaction involving human and autonomous entities in virtual reality, behavioral properties resulting from this kind of autonomy, called *constitutive*, may favor creativity and involvement by offering the user a compromise between surprise and regularity. This thesis is structured in two parts, the first one is a theoretical proposition based on relationships between enaction and autonomy of artificial interactive agents, which leads to a definition of virtual *enactive* agents. The second part is an attempt of implementation following this reflection, in which a dynamic cognition model parameters are set using an evolutionary approach.

Part I : Theoretical proposition

Chapter 1 Constitutive autonomy : from biology to artificiality

This chapter aims at presenting how the notion of constitutive autonomy can be addressed in the artificial domain.

The first part presents the biological foundations of autonomy according to the enaction theory. The latter grounds the cognition in the autonomy of a system and insists on the circular organization of living organisms. According to Varela, the active conservation of the organization in precarious conditions characterizes them. Autonomous system organization is only defined in reference to internal elements. From the system perspective, there are no inputs nor outputs, but perturbations of an organization which resists and maintains itself. Consequently, the behavior of an autonomous system, considered equivalent to its cognitive faculties, is not distinguished from the coupling history with the environment.

The second part of the chapter deals with artificial life as it could lead to simulation of entities endowed with constitutive autonomy. Indeed, artificial life reproduces biological mechanisms, which allow a circular organization. Thus, autopoietic automata can resist perturbations to maintain an organization. As for artificial embryogenesis, it addresses the production of an entity from the definition of an organization. The problem of constitution and maintaining of the identity is addressed, but these kind of systems cannot be considered

as agents, because they do not really act and, more importantly, the interaction with the environment is limited.

The third part concerns artificial agents, able to exhibit sensorimotor behaviors. The *New Artificial Intelligence* advocates for embodied agents which use world as its own model. The *Animat* approach introduces a circular causality as a necessity to simulate situated entities. A dynamic approach, in which modeling is made according to a temporal evolution of the system, is then more suited for the study of autonomous behaviors, than a functional analysis or even an analysis of components interactions. The dynamic approach can provide an account of the coupling history between the agent and the environment. The *New Robotics* goes further by redeeming the participation of the body in the production of the behavior. Finally, the *Organism Inspired Robotics* applies enactive ideas to the synthesize of autonomous artificial agents, by putting the emphasis on the necessity of a self-determined organization, able to produce and alter its goals.

This chapter concludes with a first characterization of a virtual enactive agent, i.e. a simulated agent for which we aim at get as close as possible to a constitutive autonomy. It should also be noted, as Varela defines it, that interactivity we are seeking in order to get real agents, leads us to accept a concession on the constitutive autonomy. Indeed, we keep aside the question of the identity production of the system.

Chapter 2 Technical choices for constitutive autonomy

This chapter aims at defining the models and the tools that are most suited to synthesize agents endowed with constitutive autonomy. The goal is then to propose a more concrete characterization of an enactive virtual agent for interactive simulations.

The first part addresses the question of the agent controller. Classical architectures are generally incompatible with constraints of a constitutive autonomy. Enactive virtual agents must indeed adopt a behavior self-determined by a disturbable circular organization.

The second part concerns a particular kind of artificial recurrent neural networks, CTRNN², that have such a circular organization, which operation is defined by relationships between internal components associated with environment interactions acting like disturbances. Theoretically, CTRNN are able to approximate any dynamical system behavior, including chaotic behavior. Constraints we defined are thus met by an agent controlled by a CTRNN coupled with environment through sensors and effectors, that is the reason why we retain this architecture for enactive virtual agents.

The third part presents evolutionary techniques used to set the controller parameters. In the case of robot controllers or simulated agents controllers, specific issues occur such as the evolution setup influence on robustness or the difficulty to find solutions for complex tasks. Methods have been proposed to overcome these problems. For instance, incremental

² For *Continuous Time Recurrent Neural Networks*.

evolution consists in defining more and more complex tasks based on results of previous evolution. Another proposition is given by co-evolution which applies a selective pressure fitted to evolution progress. Multi-objective algorithms provide a technique to evolve solutions that satisfy several criteria.

The chapter concludes with the choice of an evolutionary dynamic method, involving CTRNN and evolutionary methods, as a suited technique to synthesize enactive virtual agents.

Chapter 3 Confrontation of CTRNN with enactive virtual agents

This chapter present works involving CTRNN and compare evolutionary dynamic approach with enactive virtual agent notion, in order to point the specificity of our work. That work consists of introducing learning process at the ontogenetic scale.

Thus, the first part illustrates the willingness to get a constitutive-autonomous close behavior. Interest of the behavioral properties obtained comes from the principle of the behavior production rather than from the complexity of the behaviors.

The second part addresses the model capacity to behave in an evolving way. Works mentioned here show the ability for CTRNN-controlled agents to learn associations without any learning-specific mechanism.

However, we consider this learning results from an adaptation mechanism arising at a phylogenetic scale. This leads us to distinguish adaptation from learning, i.e. “reflex” learning and “constructed” learning. The “constructed” learning originates in a mechanism at an ontogenetic scale and renders the agent-environment interaction history. This learning must then allow agent-environment co-evolution at this phylogenetic scale.

Enactive virtual agent definition is extended with the capacity to develop a behavior by taking in account interaction.

Chapter 4 Process proposition to initiate learning at an ontogenetic scale

The goal of this chapter is to build a process which will initiate a learning at an ontogenetic scale in enactive virtual agents. This learning consists in sensorimotor invariants acquisition at an ontogenetic scale. Sensorimotor behaviors are indeed esteemed as the base for knowledge in enaction theory.

In the first part, the notion of sensorimotor invariants is defined and illustrated. Phenomenologies link it directly with the world perception. For our research, that will be a tool to observe agents behavior. We consider sensorimotor as different when they

correspond to different sensorimotor invariants. Then, “constructed” learning may result in the transformation of one sensorimotor invariant to another one.

The second part of this chapter deals with the transformation itself. Agents and environment dynamics have to be coupled so that they affect each other in order to get learning at the ontogenetic scale. the transformation of sensorimotor invariants must represent the coupling-specific history. This implies that environment disturbs agent dynamics while letting it a leeway and reciprocally that agent actions change environment reactions. Interaction must then constrain without laying down the behavior. In order to meet these criteria, our proposition is to use the notion of guidance.

An effective guidance will induce co-evolution as interaction will suit agent behavior at every stage of its evolution. This chapter ends by setting-up a validation program in three experiments.

Part II : Exploratory experiments

Chapter 5 Incremental phylogenetic evolution of sensorimotor behaviors

This chapter describes the first step of our set-up which consists in the change of a sensorimotor behavior by using a phylogenetic approach. Due to the principles of constitutive autonomy, which implies uneasily handable behaviors, obtaining a “regular” behavior as complex as possible is not trivial.

At first, we present an experiment to evolve agents which move in a two-dimensional environment following paths, like square-shaped path or hourglass-shaped path. Explanation hypothesis of the lack of satisfying results are formulated and we decide to treat minimal tasks preserving a sensorimotor aspect. This first unsuccessful attempt show the difficulties that appeared when evolving CTRNN parameters. Yet, it comforts us in our claim that phylogenetic part of the behavior setting-up should be minimized.

The next minimal experiment show the possibility to get using an incremental approach a sensorimotor behavior gradually more and more complex with CTRNN controlled agents. The starting behavior of an agent consists in coming close to a target in a one-dimensional environment. The complexity of the task is then increased by changing arbitrarily the initial distance to the target, then by making the target mobile and finally by modifying the task so that agents have to maintain a distance to the target. For illustration purposes, we highlight sensorimotor invariants corresponding to the different behaviors.

In this chapter, the complexity of a behavior has been increased at the phylogenetic scale. This chapter also highlights the difficulties that occur when applying the evolutionary dynamic approach in the conditions that we consider as necessary to allow a learning at the ontogenetic scale.

Chapter 6 Associative guided ontogenetic learning

Chapter 6 deals with the ontogenetic learning possibility for agents controlled by a CTRNN.

The experiment uses a guidance in order to realize an associative learning between one kind of environment and one motor behavior. The aim is to get agents able to modify their behavior contingent to a guidance. This implies that some reaction to the guidance already exists and that this reaction provokes a change in the agent “attitude” to some independent stimuli. We describe the task, the simulated agents and the evolutionary set-up.

Evolutionary results are then presented. In particular, the performances of the best agent evolved are detailed. Its dynamics are studied. This highlights an evolution in the sensorimotor dynamics and internal dynamics during the learning. However, sensorimotor aspect is very weak, because without the guidance there is no relationship between agent sensations and actions.

The experiment presented in this chapter is encouraging for our approach, but the limits incite us to go further by associating guided ontogenetic learning with sensorimotor evolution.

Chapter 7 Sensorimotor invariants guided ontogenetic learning

The experiment described in this chapter aims at combining previous experiments, in order to get a sensorimotor behavior which complexity is increased resulting from a learning at the ontogenetic scale.

The purpose is to transform a sensorimotor invariant. Starting behavior is one of the behaviors obtained in the experiment described in chapter 5. Yet here, we seek to increase the complexity of the task not by phylogenetic evolution anymore, but by using guidance.

The method we used is presented by describing the task, the agent and the evolutionary process. The resulted behavior is not exactly the expected one. However, a behavioral and dynamic analysis of agents evolved show that guidance can induce a behavioral modification at the ontogenetic scale.

Conclusions

This dissertation concludes with a critical review of results and approach. Results illustrate the difficulty in handling an autonomous system, as the model resists. We have shown a set of limitations of the evolutionary approach in such a context. The perspectives of this work take into account the need of trade-offs between behavioral complexity and

constitutive autonomy. We discuss exploitation potentiality of this work in virtual reality. This require to work on the interaction between a human and an artificial system. Creativity seems to be a promising domain to use autonomous models in virtual reality.

Key words : constitutive autonomy, situated agents, enaction, artificial dynamic cognition, ontogenetic learning, evolutionary approach, CTRNN

— Résumé —

Cette thèse s'appuie sur le paradigme de l'énaction pour aborder la modélisation de l'autonomie d'entités artificielles. Dans un contexte d'interaction entre des humains et des entités autonomes en réalité virtuelle, les propriétés comportementales résultantes d'une autonomie constitutive pourraient favoriser la créativité et l'engagement en proposant à l'utilisateur un compromis entre surprise et régularité. La première partie de cette thèse est une proposition théorique qui mène à une définition d'agents virtuels situés dits énatifs, à travers des propriétés internes, liées au fonctionnement de l'agent, et des propriétés externes, liées au comportement de l'agent. Nous nous inscrivons dans une approche dynamique évolutionnaire, qui utilise conjointement des CTRNN et des méthodes évolutionnaires. Nous considérons, parmi les propriétés externes, la capacité à enrichir un comportement par l'interaction comme un point dur. Nous proposons une démarche pour obtenir un apprentissage à l'échelle ontogénétique qui reflète l'histoire particulière des interactions de l'agent avec l'environnement en utilisant un guidage. La deuxième partie est une tentative d'implémentation issue de cette réflexion à travers trois expériences. La conclusion de ce mémoire présente un bilan critique des résultats et de notre démarche, qui illustre la difficulté de la manipulation d'un système autonome. Nous avons montré des limites de l'approche évolutionnaire dans un tel contexte. Les perspectives tiennent compte de la nécessité à faire des compromis entre complexité comportementale et autonomie constitutive. Nous discutons également de l'exploitation potentielle de ce travail en réalité virtuelle.

Mots clés : autonomie constitutive, agents situés, énaction, cognition dynamique artificielle, apprentissage ontogénétique, approche évolutionnaire, CTRNN.

— Abstract —

This thesis, based on the enaction paradigm, addresses the modeling of artificial entities autonomy. In a context of human - autonomous entities interaction in virtual reality, behavioral properties resulting from constitutive autonomy may favor creativity and involvement by offering the user a compromise between regularity and surprise. The first part of this work is a theoretical proposition which leads to a definition of virtual enactive agents, through internal properties, related to the agent operation, and through external properties, related to the agent behavior. Our theoretical framework is the dynamic evolutionary approach, using CTRNN and evolutionary methods. We consider, among external properties, the ability to enrich the behavior by interaction is a hard problem. We develop a proposition to obtain a learning at an ontogenetic scale, corresponding to the specific interactions history between an agent and an environment by using a guidance. The second part of the thesis is an implementation attempt following this reflection, made of three experiments. In conclusion, a critical analysis of the results and of the approach illustrates the difficulty of handling an autonomous system. We have shown some limitations of the evolutionary approach in our context. The outlook takes into account the need of trade-offs between behavioral complexity and constitutive autonomy. We also discuss exploitation potentiality of this work in virtual reality.

Key words : constitutive autonomy, situated agents, enaction, artificial dynamic cognition, ontogenetic learning, evolutionary approach, CTRNN.