



HAL
open science

Un modèle computationnel de l'automatisation motrice

Lionel Rigoux, Olivier Sigaud

► **To cite this version:**

Lionel Rigoux, Olivier Sigaud. Un modèle computationnel de l'automatisation motrice. Deuxième conférence française de Neurosciences Computationnelles, "Neurocomp08", Oct 2008, Marseille, France. hal-00331620

HAL Id: hal-00331620

<https://hal.science/hal-00331620>

Submitted on 17 Oct 2008

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

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

UN MODÈLE COMPUTATIONNEL DE L'AUTOMATISATION MOTRICE

Lionel Rigoux et Olivier Sigaud
Institut des Systèmes Intelligents et de Robotique
Université Pierre et Marie Curie - Paris 6, CNRS FRE 2507
4 place Jussieu, F75252 Paris Cedex 05
email : lionel.rigoux@laposte.net et Olivier.Sigaud@isir.fr

RESUME

Nous proposons un modèle computationnel de l'automatisation motrice qui rend compte à la fois du transfert de séquences d'actions dirigées par des buts vers des séquences balistiques et de la diminution du temps de réponse entre actions induite par l'automatisation. Notre modèle repose sur l'hypothèse d'un codage des séquences sous la forme d'un gradient d'activation. Les ganglions de la base effectuent une compression des séquences souvent répétées en *chunks* et le cervelet se charge d'ordonner dans le temps l'exécution de ces actions. Nous vérifions sur une expérience simple que notre modèle dispose bien des principales propriétés que l'on attribue à l'automatisation motrice.

MOTS-CLES

Automatisation, Modélisation, Ganglions de la base, Cervelet.

1 Introduction

De nombreux modèles de l'automatisation proviennent de l'étude de l'apprentissage par renforcement et se concentrent sur l'évolution du contrôle au cours de l'automatisation à travers des protocoles jouant sur la sensibilité à la récompense chez l'animal. Ces modèles reproduisent qualitativement certaines données expérimentales telles que la formation rapide de séquences en mémoire lors de la première phase de l'apprentissage ou la perte du contrôle actif lors de l'automatisation, ce qui se manifeste par des difficultés à adapter le comportement dans un environnement changeant [1, 2]. Toutefois, les variations des caractéristiques temporelles de la réalisation des séquences durant les différentes phases de l'apprentissage sont négligées et restent peu claires pour ces théories.

A l'inverse, des travaux de psychologie s'intéressent à la modification des temps de réponse au cours de l'automatisation de séquences motrices. L'un des concepts qui émerge de cette littérature est la notion de *chunk*. Il s'agit d'un groupement d'actions consécutives réalisées très rapidement et de façon automatique. Lors de la réalisation d'une séquence motrice, les actions sont dans un premier temps exécutées « une par une » puis, avec l'expérience, sous la forme d'un seul *chunk*. La distinction entre ces deux modes s'exprime par une différence significative dans la si-

gnature temporelle de l'exécution motrice : le temps de réponse d'une action simple ou de la première action d'un *chunk* est largement supérieur à celui d'une action au sein d'un *chunk*. Si l'essentiel de la littérature de ce domaine expose des données psychométriques, quelques études essaient de modéliser le mécanisme de formation des *chunks* [3, 4]. En revanche, elles ne rendent pas compte de l'effet de la récompense sur la formation des automatismes.

Notre travail unifie ces deux approches au sein d'un modèle qui explicite tant les aspects comportementaux que les caractéristiques temporelles de la formations des automatismes par *chunking* des séquences motrices. Dans cet article, nous présentons donc un modèle qui doit simultanément rendre compte de la signature temporelle caractéristique des *chunks* et de la difficulté d'adapter le comportement à un environnement changeant. Pour cela, nous le testons sur une tâche classique de réalisation de séquences discrètes.

Après avoir présenté nos hypothèses de départ, nous décrivons les points clefs de notre modèle avant de présenter les résultats principaux des simulations que nous avons menées. Enfin, nous discutons des implications de notre modèle par rapport à quelques autres études de la littérature apparentée.

2 Matériel et méthode

Notre modèle se fonde sur deux hypothèses centrales. La première consiste à considérer que le codage des représentations des séquences motrices planifiées se fait par gradient d'activité [3]. Nous supposons qu'une série motrice est maintenue en mémoire sous la forme d'un gradient qui peut ensuite être mis en œuvre grâce à un module de sélection par compétition qui libère successivement les actions dans l'ordre décroissant de l'activité de leurs représentations associées. Cette théorie, imaginée pour expliquer les observations psychométriques, semble confirmée par des travaux d'enregistrement unitaire montrant l'existence de tels gradients dans le cortex préfrontal (CPF) [5].

La deuxième hypothèse postule que les ganglions de la base (GB) effectuent une compression de l'information corticale modulée par un signal de récompense véhiculé par la dopamine [6, 7]. Si le rôle des GB dans l'apprentissage de séquences est reconnu, aucun modèle à notre connaissance ne tient compte des capacités de compression

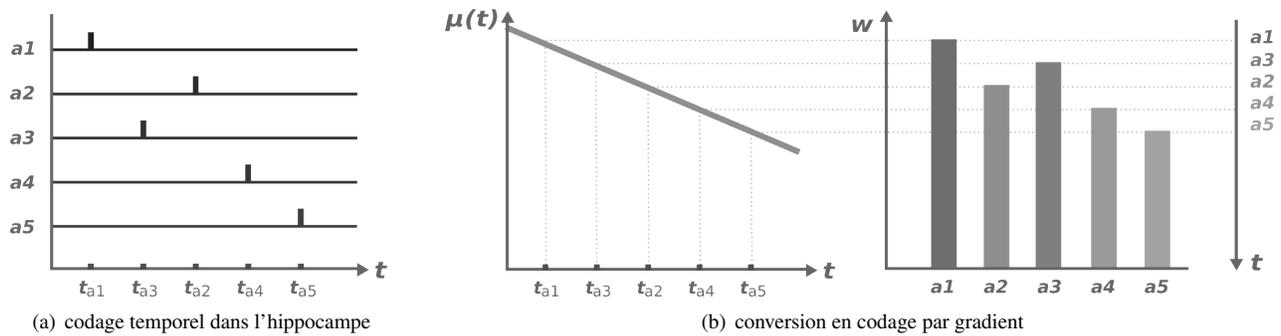


FIG. 1. Exemple de l'apprentissage d'une séquence grâce à l'hippocampe.

L'hippocampe possède des neurones sélectifs pour chacune des actions de la séquence. Ces neurones s'activent les uns après les autres dans l'ordre correspondant à la séquence réalisée, et ce de façon cyclique (au cours de chaque oscillation θ). Le graphe (a) représente l'activité des cinq neurones de l'hippocampe de notre modèle au cours d'un cycle de temps après la réalisation de la séquence $a_1 > a_3 > a_2 > a_4 > a_5$. La mémorisation de la séquence se fait sous la forme d'un vecteur de poids dépendant de la tâche qui permet de reconstituer un gradient d'activité dans le CPF. Ce vecteur est modifié lorsque l'agent reçoit une récompense en fin de séquence : pour chaque action, le nouveau poids est le résultat de la modulation par une fonction strictement décroissante $\mu(t)$ de l'activité du neurone correspondant à cette action dans l'hippocampe ; plus l'action est réalisée tôt dans la séquence, plus l'activité du neurone de l'hippocampe sélectif à cette action est prise en compte et plus le poids dans le vecteur mémoire est fort. Le gradient en (b) à droite représente l'exemple de la conversion par la fonction $\mu(t)$, au centre, de la séquence codée temporellement montrée en (a).

des GB pour expliquer la formation des *chunks* moteurs, même si quelques auteurs ont déjà envisagé cette idée de façon très générale [8]. Par ailleurs, il existe des modèles détaillés de l'implémentation de ce mécanisme de compression [9], mais ils sont peu connectés avec les problématiques de l'automatisation.

Partant de ces deux hypothèses, nous proposons que les GB contribuent à la formation des automatismes en permettant la compression en *chunks* moteurs des gradients d'activité présents dans le CPF. Dans ce contexte, le signal dopaminergique permet de restreindre cet apprentissage aux comportements qui mènent de façon stable à une récompense. Nous devons alors résoudre un certain nombre de problèmes pour obtenir un modèle complet de la constitution des automatismes fondé sur ces deux hypothèses.

2.1 Formation et maintien des gradients

Le premier problème concerne la formation et le maintien d'un gradient d'activité représentant une séquence. Sur la base d'une revue de la littérature, nous attribuons ces activités à une collaboration entre le CPF [10, 11] et l'hippocampe [12]. Dans notre modèle, l'hippocampe permet après chaque répétition d'une séquence de renforcer une mémoire à long terme qui peut facilement être chargée dans le CPF sous la forme d'un gradient. Une vue synoptique de cette partie de notre modèle est donnée par la figure 1.

2.2 Des gradients aux séries d'actions

Le deuxième problème, soulevé par la contrainte du codage par gradient, porte sur la conversion des gradients d'activité en séries temporelles motrices. Nous attribuons

cette fonction aux GB et, en particulier, au striatum, qui est fortement connecté au CPF et impliqué dans la fonction motrice. Le fonctionnement du modèle suit un principe classique de sélection de l'action par compétition présenté sur la figure 2a (voir légende). Cette architecture est inspirée de travaux sur les processus de décision de type modèle à accumulateur [13]. Selon cette approche, à chaque action est assigné un accumulateur qui intègre les indices corticaux associés à cette action. Dès qu'un accumulateur dépasse un certain seuil, l'action correspondante est sélectionnée. La vitesse d'accumulation est également régulée par un facteur proportionnel au degré de compétition : à stimulation égale, plus il y a d'actions en compétition, plus l'intégration est lente et plus il faut de temps pour choisir une action.

Une telle architecture permet une grande réactivité du système en cas de modification de l'environnement perceptif de l'agent. En effet, les informations maintenues dans le CPF relatives aux actions envisagées sont mises à jour dynamiquement à partir des afférences perceptives du système. Le contrôle visuo-spatial supporté par le CPF, lorsqu'il est relayé par le processus de sélection d'action des GB, permet une adaptation en temps réel à la tâche.

2.3 Formation des *chunks*

Le problème suivant est celui de la formation des *chunks*. Nous proposons l'hypothèse que les GB, en identifiant les gradients présents de façon répétée dans le CPF, permettent la création de ces groupes d'actions. Notre modèle est inspiré de celui de [9], qui reprennent les réseaux de neurones utilisés pour réaliser des analyses statistiques et développent un modèle des GB qui compresse l'information corticale en se fondant sur le renforcement. Dans la pratique, chaque neurone du striatum code soit pour une

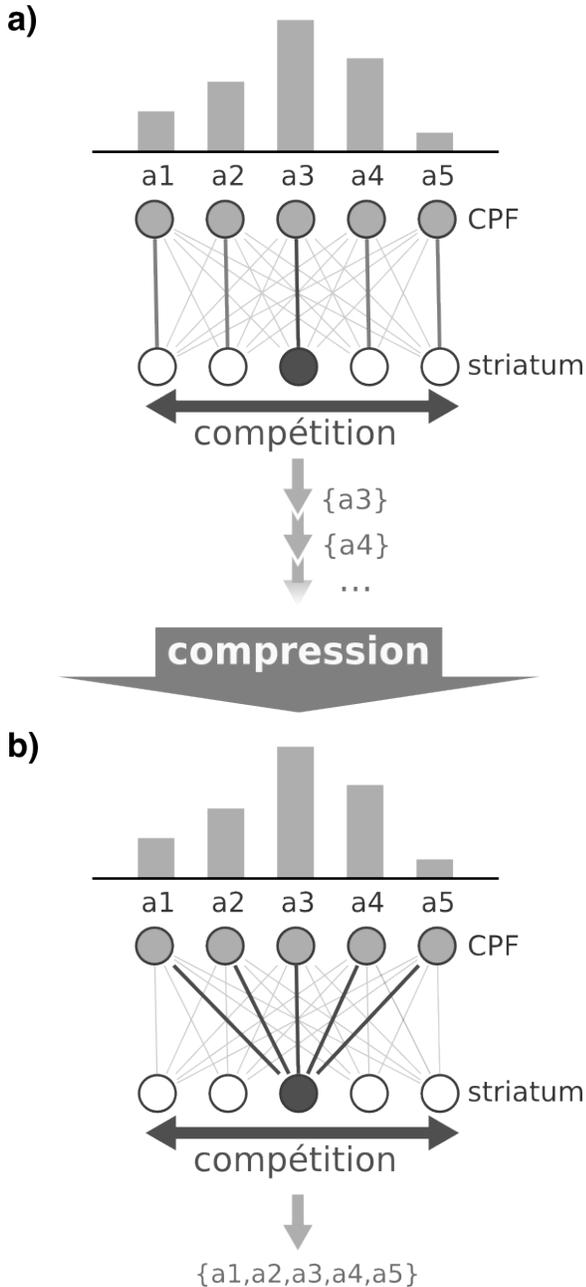


FIG. 2. Formation des *chunks* par compression basale.
 Avant automatisisation (a), chaque neurone du CPF projette sur un unique neurone du striatum. La représentation d'une séquence sous la forme d'un gradient d'activité peut être convertie en séquence temporelle par une succession de compétitions : le neurone striatal le plus actif gagne la compétition, l'action correspondante est exécutée et le neurone du CPF associé est réinitialisé. Par itération, toutes les actions représentées dans le CPF sont exécutées par ordre décroissant du gradient.
 Avec l'entraînement, l'algorithme de compression permet de rendre un unique neurone du striatum sélectif à l'ensemble du gradient. Ainsi, après automatisisation (b), lorsque le seul neurone striatal actif « gagne » la compétition, il valide l'ensemble des actions qui lui sont associées sous la forme d'un chunk. L'exécution de ce groupe d'actions est ensuite régulée par le cervelet. Ce dernier filtre les actions sélectionnées par les GB pour éviter qu'elles ne soient toutes réalisées simultanément.

action individuelle, soit pour plusieurs actions dès lors que ces dernières se retrouvent régulièrement associées dans un même schéma d'activation dans le CPF (voir fig. 2b). Ainsi, avec l'entraînement, les GB reçoivent comme candidats à la compétition des *chunks* plutôt que des actions simples et la sélection des actions gagnantes se fait par groupe. Par conséquent, le système actif de résolution de buts de l'agent passe progressivement d'un mode où il exécute son plan étape par étape en suivant un guidage visuo-spatial à un mode où le plan d'action est validé d'un seul coup et transféré en bloc au système moteur qui doit alors gérer le guidage en interne.

2.4 Exécution des *chunks*

Un nouveau problème se pose alors, celui de la réalisation temporellement ordonnée des actions qui sont validées en groupe par les GB. Autrement dit, il faut expliquer comment se fait le contrôle moteur après compression, lorsque toutes les actions constituant le *chunk* sont libérées d'un seul coup et que les GB ne jouent plus leur rôle de convertisseur de gradient en séquence temporelle.

Une solution est de faire intervenir un module capable d'apprendre une représentation de la séquence dans des coordonnées musculo-squelettique sous la forme d'associations du type « après l'action a_1 , faire l'action a_2 ». Ce module, après compression, peut réguler l'exécution motrice en filtrant le *chunk* sélectionné par les GB, libérant progressivement les actions par un chaînage de proche en proche. Pour poursuivre l'exemple de la figure 2 : si les GB sélectionnent le *chunk* composé des actions a_1 à a_5 , et si ce module a appris les associations « au début de la tâche faire l'action a_3 », « après a_3 laisser faire l'action a_4 », etc. le système exécute la séquence motrice telle qu'avant automatisisation. Nous proposons d'attribuer au cervelet (Cb) l'apprentissage de ces associations à l'instar de [3, 4]. En effet, des données empiriques suggèrent que le Cb est impliqué dans la rétention des séquences automatisées. De plus, cette structure cérébrale semble posséder une architecture adaptée à cette fonction et notre travail s'inspire fortement des données anatomiques connues pour modéliser ce module.

Après automatisisation, le système n'est plus capable de réagir à des changements environnementaux et continue à réaliser la séquence à l'identique, même si les informations maintenues dans le CPF sont modifiées au cours de la tâche. Ceci est dû à la perte du guidage visuo-spatial initial qui est remplacé par un contrôle intrinsèque de type musculo-squelettique soutenu par le Cb. L'adéquation de la performance motrice avec le but visé devient alors dépendante de la stabilité de l'environnement qui a permis l'apparition du comportement automatique.

2.5 Gestion du renforcement

Les GB effectuent une compression de l'information corticale qui est modulée par un signal dopaminer-

gique. Concernant le comportement des neurones dopaminergiques, nous nous appuyons sur les modèles classiques de « différence temporelle » — populaires en apprentissage par renforcement [14] — pour simuler le signal dopaminergique dans notre architecture en restant à un niveau de description fonctionnel, sans perdre de vue pour autant les contraintes biologiques existantes. Ce mécanisme vient dans notre modèle réguler l'apprentissage de manière à ce que le système s'oriente vers la réalisation de séquences récompensées, *i.e.* dirigées vers les buts artificiels que nous lui imposons.

3 Résultats

Nous testons notre modèle sur une expérience inspirée des tâches de Réalisation de Séquences Discrètes largement utilisées pour tester les effets de la formation des *chunks* sur la modification des temps de réponse. L'agent a accès à cinq boutons lumineux sur lesquels il peut appuyer et dont l'état « allumé » ou « éteint » détermine l'entrée du système.

L'expérience se déroule en deux phases. Dans un premier temps, l'agent doit simplement apprendre à associer l'allumage d'un bouton à l'action consistant à appuyer dessus. On constate qu'il atteint une performance parfaite au bout de quelques essais. Nous laissons volontairement le système réaliser un nombre d'essais largement supérieur à celui strictement nécessaire à l'apprentissage des associations sensori-motrices en elles-mêmes. En effet, la pondération cortico-striatale doit se stabiliser afin que chaque action soit identifiée comme une action élémentaire, *i.e.* que chaque neurone du striatum ne soit sélectif que pour un seul neurone du CPF. Ainsi, chaque action peut être sélectionnée indépendamment au sein d'un gradient lors de la phase suivante.

Dans une seconde phase, l'allumage des boutons se fait en série et le système doit réaliser la séquence correspondante le plus vite possible jusqu'à automatisation complète. Les temps de réponses sont calculés en tenant compte de la compétition au sein des GB ainsi que de l'action du Cb. Les boutons sont allumés séquentiellement, un à un. Une unique séquence est répétée en boucle et une récompense est délivrée à la fin de chaque exécution réussie. Le rythme d'affichage dépend de la vitesse de l'agent : dès qu'un bouton est enfoncé, le suivant s'allume. Ceci permet à l'agent d'accélérer au fur et à mesure qu'il mémorise et automatise la série d'actions. Les résultats sont donnés sur la figure 3.

Nous avons testé avec succès deux des principaux critères chronométriques identifiés par [3] dans le cadre de la réalisation de séquences motrices discrètes, l'augmentation du rapport latence/IRI et la disparition de l'effet de longueur sur la latence avec l'apprentissage, comme précisé sur les figures 4 et 5. Notre modèle semble donc rendre compte de l'automatisation au point de vue de la psychophysique en répondant aux critères habituels de la définition des *chunks*.

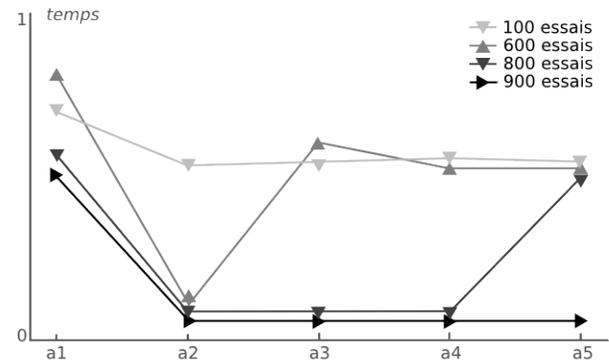


FIG. 3. Temps de réponses pour une séquences de cinq actions
L'ordonnée donne les temps inter-réponses (IRI), *i.e.* pour chaque action le temps écoulé depuis l'action précédente. Pour la première action cette valeur correspond au temps de latence, c'est-à-dire le temps qui la sépare de l'instant d'affichage du premier bouton. Au bout de 600 essais, un chunk comprenant les actions a_1 et a_2 se forme. A l'essai 800, ce sont les quatre premières actions qui sont groupées. Après 900 essais, toute la séquence est compressée.

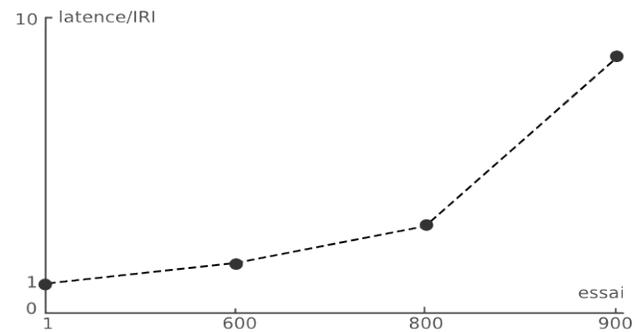


FIG. 4. Augmentation du ratio latence/IRI
Ce graphique présente l'évolution au cours de l'apprentissage du rapport entre la latence et le temps de réponse intra-séquence. La valeur élevée en fin de simulation de ce ratio, qui est un indicateur du degré du chunking, indique que toutes les actions font alors partie d'un seul et unique chunk.

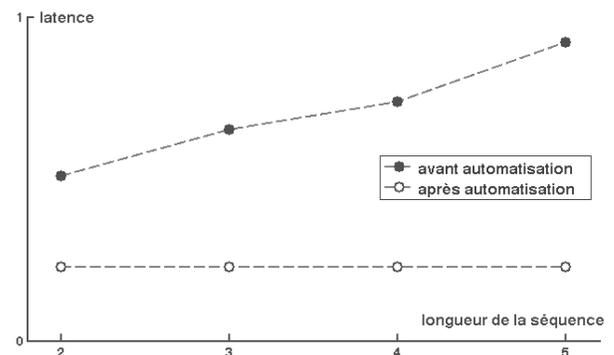


FIG. 5. Effet de longueur sur la latence
Avant automatisation, la latence augmente linéairement avec la longueur de la séquence. Selon la littérature empirique, cet effet connu sous le nom de Sequence Length Effect on Latency disparaît après automatisation. Ce graphique montre la latence avant et après automatisation pour des simulations sur des séquences de différentes longueurs : les résultats obtenus sont conformes aux données expérimentales.

Par ailleurs, pour vérifier que le modèle reproduit fidèlement les caractéristiques comportementales des mouvements automatiques, nous introduisons une séquence alternative apparaissant tous les cent essais pour tester la capacité du système à garder un contrôle actif sur son comportement, c'est-à-dire à s'adapter à un affichage lumineux modifié. La colonne 3 du tableau 1 donne les résultats. En début d'apprentissage, le système peut s'adapter immédiatement au changement de séquence bien qu'il utilise sa mémoire hippocampale. Comme attendu, l'agent ne peut s'empêcher de réaliser une série d'actions si celle-ci a été automatisée, y compris en présence d'indices visuels contradictoires. Le modèle semble donc valide par rapport à la définition de l'automatisation que donnent les théories de l'apprentissage par renforcement.

Notre modèle prédit que le plan d'action maintenu dans le CPF n'est plus nécessaire à la réalisation d'une séquence automatisée après le début de celle-ci. Par conséquent, tout comportement réalisé de façon automatique laisse les ressources cognitives de l'agent disponibles pour la réalisation d'un autre travail. Pour vérifier que notre modèle répond également à cette contrainte, nous simulons la perturbation occasionnée par une tâche supplémentaire en bruitant le CPF juste après le début de l'exécution de la séquence apprise. La colonne 4 du tableau 1 donne les résultats. Ce test simpliste permet d'envisager une extension du modèle utilisant le CPF pour organiser le comportement simultanément à la réalisation des automatismes. Ce phénomène peut aussi être compris en terme d'activité cérébrale. Avant automatisation, l'information relative à une séquence planifiée est relayée étape par étape au bloc moteur sous la forme d'actions individuelles. Ce processus se déroule au même rythme que la séquence ce qui permet un contrôle en ligne par le CPF. Par contre, après automatisation, le CPF se désengage rapidement en laissant au bloc moteur le soin de gérer l'exécution motrice avec l'aide du Cb.

TAB. 1. Performances sur les séquences alternatives et capacité à résister aux perturbations du CPF

Au cours de l'apprentissage, de plus en plus d'actions sont compressées dans le *chunk*. Une action ainsi automatisée devient insensible aux perturbations du CPF mais, en contrepartie, n'est plus sous contrôle actif en cas de changement de séquence.

essai	<i>chunk</i> formé	perturbation séquence	perturbation CPF
1	{a1}	succès	échec à partir de a2
600	{a1, a2}	erreur en a2	échec à partir de a3
800	{a1, a2, a3}	erreur en a2 et a3	échec à partir de a4
900	séquence entière	échec	succès

4 Discussion

4.1 Représentation du temps

Tous les calculs de temps de notre modèle sont basés sur l'hypothèse que les GB implémentent un processus de décision de type modèle à accumulateur (cf 2.2). Les temps

de réponses dépendent d'une intégration sur le temps dont la vitesse dépend d'un grand nombre de paramètres, tels que la force de stimulation ou l'attention, que nous ne prenons pas en compte dans notre modèle. La ligne de temps utilisée dans ce type de modélisation est relative, ce qui explique l'absence d'unité sur les axes temporels de nos graphes. Cette faiblesse pourrait être compensée par une modélisation plus fine des processus de sélection qui se déroulent dans les GB. Cependant, une formalisation à un tel niveau de précision des mécanismes neurophysiologiques en jeu sort du cadre de notre étude.

4.2 Relation aux autres modèles de l'automatisation

Certains auteurs [3] proposent un modèle qui rend compte des propriétés psychométriques de l'automatisation en utilisant un codage des séquences par gradient d'activité et un module de sélection par compétition implémenté par les GB. Pour ces auteurs, l'automatisation résulte d'une participation du Cb qui accélère le mécanisme de sélection des GB. Si ce modèle simule avec succès les propriétés chronométriques de la formation des *chunks*, il ne permet pas d'expliquer l'impact du renforcement sur l'apprentissage et plus particulièrement la perte du contrôle actif lors de la réalisation d'une séquence automatisée.

Par opposition au modèle précédent, de nombreuses études se focalisent sur l'influence de la récompense sur le comportement. Ces travaux, issus des théories de l'apprentissage de tradition béhavioriste, postulent une dichotomie entre deux systèmes relativement indépendants : l'un à l'origine des automatismes, l'autre en charge du comportement dirigé vers un but.

Le système automatique serait régi par un processus d'apprentissage par renforcement. Les caractéristiques comportementales que l'on attend de ce système sont très bien simulées par les architectures dites « acteur/critique ». La popularité de ces modèles tient principalement dans leur capacité à formaliser le signal dopaminergique associé à la récompense. Leur ancrage anatomique est néanmoins encore sujet à de vifs débats [7].

Le système de contrôle actif, quant à lui, connaît plusieurs variantes selon les auteurs mais son rôle est généralement attribué au CPF. Selon les versions, ses capacités d'adaptation aux changements de l'environnement sont dues au type de représentations qu'il manipule [2], plus abstraites que celles utilisées par le système automatique, ou bien à son mécanisme d'apprentissage plus flexible [1]. Dans tous les cas, ce système de contrôle se trouve en compétition avec les GB pour le choix de l'action.

Un problème épineux que rencontrent les modèles de cette approche est l'arbitrage de l'attribution du contrôle entre les deux systèmes. Une solution proposée par [1] est basée sur une compétition régulée par la qualité des prédictions respectives des deux sous-systèmes. Ce résultat n'est pas spécifique à une implémentation particulière et n'est

pas exclusif avec notre modèle. Il est possible de trouver une interprétation de ces données dans notre architecture. En effet, il peut s'agir d'une interaction entre la qualité des gradients représentés dans le CPF et le type de projection sur le *striatum* qui en découle. Aussi, l'algorithme de compression que nous utilisons est proche (à quelques contraintes près sur la non-linéarité des neurones) des algorithmes d'analyse en composantes indépendantes. Ces algorithmes bénéficient d'une interprétation en terme de théorie de l'information [15]. Celle-ci nous donne donc un point d'attaque pour des recherches ultérieures sur la relation entre notre modèle et les travaux de [1].

Pour résumer, la différence fondamentale entre notre approche et les autres tient dans le rôle attribué aux GB. Pour les modèles qui s'appuient sur des architectures acteur/critique, ils apprennent à contrôler le comportement de façon autonome et, pour pallier la lenteur de ce système, un module de planification est ajouté avec un but redondant. Dans notre cas, les GB ne font que valider, de manière plus ou moins laxiste, les actions planifiées par le CPF, laissant le contrôle automatique au Cb. Notre approche présente deux avantages. D'une part, elle donne une explication précise du mécanisme permettant la libération du CPF lors de l'exécution des automatismes. D'autre part, elle éclaire le rôle du Cb dans la rétention des séquences automatisées, hypothèse que soutiennent les études en lésion et sur la gestion du temps mais qu'ignorent les modèles de l'apprentissage par renforcement que nous connaissons.

Nous ne pouvons confronter directement notre modèle aux architectures acteur/critique uniquement sur le fait que ces derniers ne rendent pas compte pour l'instant des données chronométriques ni de l'implication du Cb dans l'automatisation. Toutefois, il y a une prédiction de notre modèle qui diffère de celles que peuvent faire les modèles exclusivement fondés sur l'apprentissage par renforcement. Cette prédiction porte sur l'ordre dans lequel les actions sont progressivement automatisées. Dans les architectures acteur/critique, l'apprentissage se fait par un chaînage arrière depuis la dernière action récompensée alors que, dans notre cas, la progression se fait dans l'autre sens, en partant de la première action de la séquence. Cette propriété peut être testée empiriquement dans un protocole qui reste à déterminer.

4.3 Conséquences du codage par gradient

Le codage par gradient présente quelques défauts. Le premier concerne la gestion des répétitions. En effet, il n'est *a priori* pas possible de coder plusieurs fois la même action au sein d'un unique gradient. Toutefois, il est possible d'imaginer diverses solutions, telle qu'une redondance dans le codage, le recours à des oscillateurs ou encore à un mécanisme permettant la répétition du chargement du gradient dans le système. Ce problème est malheureusement souvent éludé dans les travaux empiriques.

Une autre contrainte porte sur la limitation du nombre

d'actions qui peuvent être codées simultanément et de façon robuste dans un gradient. Dans le cas de séquences trop longues pour être représentées par un seul gradient, il est possible de faire appel à des processus hiérarchique qui devraient permettre de scinder et d'organiser un comportement complexe en plusieurs gradients exécutés en série. Une conséquence de cette hypothèse est que le découpage des séquences en *chunks* n'est sûrement pas entièrement dépendant des GB.

5 Conclusion

Nous avons présenté un modèle qui rend compte simultanément des caractéristiques comportementales de l'automatisation et de l'accélération de l'exécution motrice due à la formation de *chunks*. Dans nos travaux futurs, nous étudierons notamment l'extinction d'un automatisme lorsque son exécution devient inappropriée. Nous améliorerons également notre modèle pour permettre l'exécution de doubles tâches.

Références

- [1] Daw N.D., Niv Y., and Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*, 8(12) :1704–11, Dec 2005.
- [2] Haruno M. and Kawato M. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops : fMRI examination in stimulus-action-reward association learning. *Neural Netw*, 19(8) :1242–54, Oct 2006.
- [3] Rhodes B.J., Bullock D., Verwey W.B., Averbeck B.B., and Page M.P.A. Learning and production of movement sequences : behavioral, neurophysiological, and modeling perspectives. *Human movement science*, 23(5) :699–746, Nov 2004.
- [4] Bullock B. Adaptive neural models of queuing and timing in fluent action. *Trends Cogn Sci*, 8(9) :426–33, Sep 2004.
- [5] Averbeck B.B., Chafee M.V., Crowe D.A., and Georgopoulos A.P. Parallel processing of serial movements in prefrontal cortex. *PNAS*, 99(20) :13172–7, Oct 2002.
- [6] Bar-Gad I. and Bergman H. Stepping out of the box : information processing in the neural networks of the basal ganglia. *Current Opinion in Neurobiology*, 11(6) :689–95, Dec 2001.
- [7] Joel D., Niv Y., and Ruppin E. Actor-critic models of the basal ganglia : new anatomical and computational perspectives. *Neural Netw*, 15(4-6) :535–47, Jan 2002.
- [8] Graybiel A.M. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem*, 70(1-2) :119–36, Jan 1998.
- [9] Bar-Gad I., Morris G., and Bergman H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog Neurobiol*, 71(6) :439–73, Dec 2003.
- [10] Marshuetz C. Order information in working memory : an integrative review of evidence from brain and behavior. *Psychological Bulletin*, 131(3) :323–39, May 2005.
- [11] Mushiaki H., Saito N., Sakamoto K., Itoyama Y., and Tanji J. Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron*, 50(4) :631–41, May 2006.
- [12] Schendan H.E., Searl M.M., Melrose R.J., and Stern C.E. An fmri study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron*, 37(6) :1013–25, Mar 2003.
- [13] Bogacz R. and Gurney K. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural computation*, 19(2) :442–77, Feb 2007.
- [14] Niv Y., Duff M.O., and Dayan P. Dopamine, uncertainty and TD learning. *Behavioral and brain functions : BBF*, 1 :6, May 2005.
- [15] Nadal J.P. and Parga N. Analyse en composantes indépendantes et séparation de sources : approches basées sur la théorie de l'information et recherche de conditions minimales. *16^e Colloque sur le traitement du signal et des images*, 1997.