

De la nécessité de l'intégration d'un modèle d'hippocampe dans une approche corticale de la sélection de l'action

Hervé Frezza-Buet, Nicolas P. Rougier

► **To cite this version:**

Hervé Frezza-Buet, Nicolas P. Rougier. De la nécessité de l'intégration d'un modèle d'hippocampe dans une approche corticale de la sélection de l'action. Frédéric Alexandre, Jean-Daniel Kant. Neurosciences et Sciences pour l'Ingénieur, 1998, Munster, France. 4 p, 1998. <inria-00098694>

HAL Id: inria-00098694

<https://hal.inria.fr/inria-00098694>

Submitted on 26 Sep 2006

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.

De la nécessité de l'intégration d'un modèle d'hippocampe dans une approche corticale de la sélection de l'action

Hervé Frezza-Buet (frezza@loria.fr) **Nicolas Rougier** (rougier@loria.fr)
LORIA, B.P. 239, 54506 Vandœuvre-lès-Nancy CEDEX

Résumé

Cet article présente dans un premier temps une architecture inspirée par les données neurobiologiques concernant le cortex permettant à un agent autonome d'évoluer dans un environnement ouvert (démuni d'obstacles). Cet animat est alors capable d'agir en accord avec l'état de ses différentes motivations. Cependant, cette mémoire associative procédurale (le savoir comment) ne semble pas suffisante à la survie de l'animat. Il semble donc qu'il faille adjoindre à cette architecture l'équivalent d'une mémoire explicite (le savoir quoi) afin de fournir à l'animat la possibilité de manipuler des représentations à caractère global et épisodique pour l'élaboration d'un comportement. L'introduction d'un modèle d'hippocampe semble constituer une solution adéquate.

1 Introduction

Le problème de déterminer les séquences d'actions qui permettent à un robot de maximiser une récompense, alors qu'il est immergé dans un environnement dont la configuration et les lois lui sont inconnues, est usuellement abordé suivant le paradigme skinnérien du conditionnement opérant, dont un des modèles algorithmiques parmi les plus connus est le Q-Learning [4]. Nous nous proposons dans cet article de montrer comment un modèle de cortex associatif, capable d'extraire puis d'exploiter des régularités spatio-temporelles multimodales, peut contribuer à une sélection de l'action. Cette sélection est consécutive à un apprentissage par essai-erreur, comme le fait le Q-Learning, mais fournit au robot un jeu de représentations procédurales plus explicites que les Q-values, généralisables et utilisables dans une certaine mesure pour des transferts d'apprentissage. Toutefois, face aux limites du modèle en terme de complexité des représentations qu'il peut manipuler, ils nous est apparu nécessaire de conférer au robot la capacité de construire représentations hautement multimodales qui joueraient le rôle de mémoire épisodique. C'est dans cette optique que nous détaillons les fonctionnalités qu'un tel modèle peut attendre d'une modélisation de l'hippocampe, en nous attachant à définir la nature des relations entre le modèle cortical et le modèle hippocampique.

2 Un modèle cortical pour la sélection de l'action

2.1 Maxicolonnes et cartes

Le cortex est une structure du cerveau constituée d'une couche en deux dimensions, pavée d'éléments similaires : les maxicolonnes corticales [3]. Celles-ci reçoivent des informations provenant des capteurs sensoriels (via divers traitements non détaillés ici) ou d'autres maxicolonnes du cortex. Les maxicolonnes recevant des informations de même modalité sont regroupées et forment des cartes corticales. Une carte est dite primaire lorsque ses maxicolonnes reçoivent des informations mono-modales, chaque maxicolonne détectant un événement de cette modalité. Afin de représenter des événements bi-modaux, il existe des cartes dites associatives. Une carte associative associant deux cartes primaires possède des maxicolonnes connectées à une maxicolonne de chacune des deux cartes primaires. Les maxicolonnes des cartes associatives reçoivent donc des informations concernant la survenue de deux événements de modalités différentes, ce qui leur permet de détecter leur conjonction temporelle¹. De même, en associant deux cartes associatives à l'aide d'une carte associative, on peut représenter des événements de forte multimodalité [2].

2.2 Gestion de flux d'information

Dans notre modèle, chaque maxicolonne a un conditionnement autonome, non-supervisé, qui repose sur la nécessité pour une maxicolonne d'intégrer trois principaux flux d'information. Les flux **feed-forward** correspondent aux flux perceptifs à partir desquels la maxicolonne, qu'elle appartienne à une carte primaire ou à une carte associative, exerce son activité de détection. Il s'agit des flux d'informations que nous avons déjà évoqués. Les flux **feedback** sont des flux qui proviennent de cartes associatives supérieures. Les flux **latéraux** sont enfin des flux qui proviennent des autres maxicolonnes de la carte. Ils permettent de réaliser certains contrôles présentés plus loin.

¹ Les données biologiques font état de relations temporelles plus complexes que nous n'abordons pas dans notre modèle.

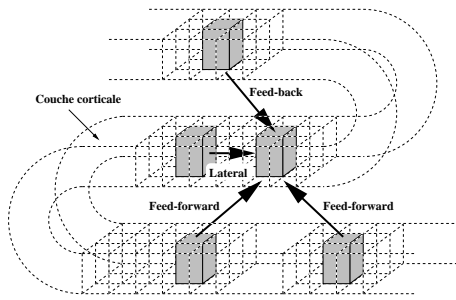


FIG. 1 – Intégration de trois types de flux d'informations

2.3 Différents types d'activité

Nous inspirant des données présentées dans [3], nous attribuons aux maxicolumnes du modèle trois principaux types d'activité correspondant aux cas suivants :

1. L'**excitation** d'une maxicolonne survient lorsque l'événement perceptif qu'elle est censée détecter est présent.
2. Une maxicolonne a une activité d'**appel** lorsque le système, suite aux buts visés par l'animat et aux effets d'apprentissages antérieurs, estime que l'occurrence de l'événement que détecte la maxicolonne est un état souhaitable, autrement dit un but.
3. Une maxicolonne est dite **satisfaite** lorsqu'elle est à la fois appelée et excitée ; ce que demandait le système est arrivé.

Les activités d'excitation se transmettent le long des voies feed-forward. Les activités d'appel se transmettent le long des voies feed-back (un but bi-modal exerce une demande suivant ses deux modalités) et le long des voies feed-forward (un but peut avoir besoin d'une association supérieure, avec une autre modalité) pour être réalisé. La transmission de l'activité d'appel suivant les voies latérales n'a lieu que lorsque des relations séquentielles entre deux buts d'une même carte ont été apprises, ce qui sera détaillé plus loin. Comme la propagation de l'activité d'appel est très large, il nous est nécessaire de mettre en compétition les maxicolumnes appelantes d'une même carte, suivant une estimation de leur aptitude à effectivement être satisfaite² de sorte à inhiber l'appel des maxicolumnes peu prometteuses. Une carte agit ainsi comme un **filtre d'appel vertical**, puisqu'après compétition, une seule maxicolonne de la carte transmet encore une activité d'appel via les voies feed-back et feed-forward.

Sans entrer dans le détail des calculs d'activité, on peut déjà entrevoir le fait que l'excitation est liée à la perception de l'état de l'environnement, que l'appel est lié à la répercussion sur les différentes cartes modalitaires des « besoins » de l'animat. Il s'agit d'une activité de commande. Dès lors, il est naturel

2. Cette estimation tient compte de l'état actuel des perceptions et des régularités temporelles apprises, elle n'est pas détaillée ici.

lorsqu'on s'inscrit dans un paradigme skinnérien de considérer l'activité de satisfaction comme étant au cœur des mécanismes d'apprentissage.

2.4 La construction et l'exploitation de séquences temporelles

La mémoire corticale apparaît dans notre modèle essentiellement procédurale, fonctionnelle, du fait de mécanismes d'apprentissage temporels prédominants. Ces apprentissages temporels sont en fait des détections de causalité, estimée à partir de corrélations temporelles. **Ce type d'apprentissage n'a lieu qu'entre maxicolumnes d'une même carte.**

Le raisonnement que nous avons suivi est le suivant : si la survenue d'un événement est souvent suivie de la survenue d'un événement « désiré » par le système, le premier événement est certainement nécessaire à la survenue du second. En terme de maxicolonne, on cherche la maxicolonne d'une carte dont l'excitation prédit au mieux la satisfaction d'une autre maxicolonne, un but, de la carte. La maxicolonne excitée sera alors considérée comme un sous-but du but, et l'appel du but, lorsqu'il surviendra à nouveau, sera transmis « latéralement » au sous-but. De façon analogue à [10], nous utilisons une adaptation de la règle de Rescorla-Wagner présentée dans [1] pour juger de la pertinence avec laquelle on peut attribuer les relations de cause à effet.

2.5 Résultats expérimentaux

Nous avons testé la validité de l'approche sur une expérience représentée en figure 2. Le modèle est utilisé pour guider un robot simulé dans un environnement inconnu à partir de perceptions provenant d'une caméra.

Après une phase d'apprentissage, on observe que le robot cherche à satisfaire le besoin le plus pressant, en réagissant exclusivement au contenu de l'image en cours. Des propriétés de généralisation et de transfert d'apprentissages sont mises en évidence lorsque le robot se sert d'une régularité apprise dans une carte pour guider la construction de nouvelles séquences au sein de cartes associatives de plus haut niveau.

3 Les limitations du modèle

3.1 Différentes mémoires, différentes fonctions

Les limitations de l'animat sont directement liées au type de mémoire dont il dispose. En effet, on distingue généralement deux grands types de mémoires. La mémoire explicite est liée aux connaissances explicites (accessibles à une récupération consciente) et à la représentation des faits et des événe-

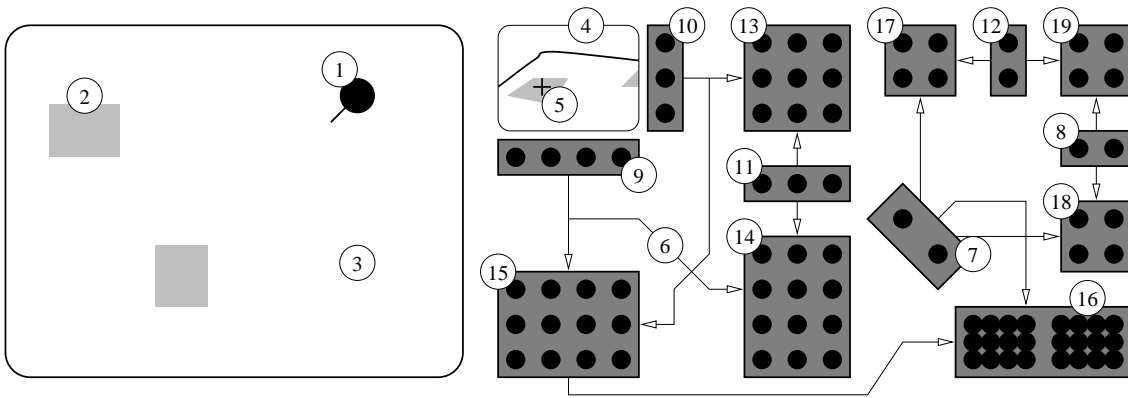


FIG. 2 – *Expérimentation* : (1) Robot, (2) Zone de « nourriture » ou de « boisson » suivant la couleur, (3) Environnement, (4) Image de caméra, (5) Point de focalisation, (6) Un flux feed-forward, (7) Carte des besoins : un appel est généré sur une des deux maxicolonnes suivant que l’animat a « faim » ou « soif », les maxicolonnes sont excitées quand le besoin est satisfait, (8) Carte des couleurs : Un appel provoque le centrage du point de focalisation sur une tache de la couleur correspondante, l’excitation traduit la présence effective de cette couleur sur l’image, (9-10) L’excitation des maxicolonnes est fonction de la position du point de focalisation sur tache de la couleur demandée, (11) Carte motrice des déplacements, l’appel sur une maxicolonne provoque l’exécution de l’action « gauche », « droite », « en avant » correspondante, (12) Carte motrice des actions d’« ingestion » pour « manger » et « boire », (13-19) Cartes associatives, on y stocke des régularités comme le fait que manger satisfait la faim, qu’il faille d’abord centrer la tache sur la caméra avant d’avancer vers elle, etc...

ments. La mémoire implicite est quant à elle liée au savoir-faire et aux représentations d’aptitudes cognitives et motrices. Chez l’homme, comme chez l’animal, la dissociation de ces deux mémoires a été mise en évidence suite à des amnésies antérogrades globales. Il est à noter que chez l’animal, on préfère distinguer la mémoire (mémoire explicite) et les habitudes (mémoire implicite).

3.2 Amnésie antérograde

L’animat ne possédant que des connaissances de type associatif (faim \leftrightarrow manger), il n’est pas à même de pouvoir bénéficier de l’expérience passée, se condamnant ainsi à adopter un comportement purement réactif en fonction de ses motivations internes. Or, ce comportement réactif, s’il est viable pour des environnements peu hostiles (présence homogène de sources d’eau et de nourriture peu éloignées), est un comportement « suicidaire » dans des environnements modérément hostiles. Ainsi, un environnement où chaque source serait suffisamment éloignée des autres sources³ constituerait un problème difficilement résolvable par l’animat. En effet, ne possédant pas l’équivalent d’une mémoire explicite, il ne peut prétendre mémoriser des informations spatiales. L’histoire passée de l’animat l’autorise donc à extraire les régularités du monde et à les mémoriser sous forme associative, mais celui-ci ne peut se souvenir des faits récents sous une forme explicite. *L’animat souffre d’amnésie antérograde.*

3. On peut imaginer un environnement où les sources sont éloignées de telle façon que l’animat ne puisse voir les autres sources lorsqu’il est lui-même sur une source.

4 Introduction de l’hippocampe

Au vu des limitations du modèle existant, il semble donc nécessaire d’introduire une forme de mémoire déclarative autorisant une mémorisation des faits récents du type « je viens de passer par là » ou « je sais où il y a de la nourriture ». Si l’animat est capable de retenir ce type d’information (conduisant éventuellement à la conception d’une carte cognitive interne), il sera alors capable d’évoluer de façon plus autonome dans un environnement complexe.

4.1 Rôle de l’hippocampe

L’hippocampe, partie intégrante du système limbique, est fortement lié à la mémoire explicite tant chez l’homme que chez l’animal et la lésion, provoquée ou accidentelle, de cette structure entraîne une incapacité de mémorisation des faits récents. Chez le rat, une lésion de l’hippocampe entraîne une incapacité à résoudre les tâches de discrimination négative (faire une action en présence d’un stimulus, ne rien faire en l’absence de ce même stimulus). L’introduction d’une structure hippocampique semble donc nécessaire à l’architecture globale afin de doter l’animat de capacité de mémorisation explicite.

4.2 Différentes approches

La modélisation de la structure hippocampique peut se faire selon différentes approches. Ainsi, chez le rat, les neurobiologistes s’intéressent essentiellement aux fonctions caractéristiques de l’hippocampe (cellules de lieux, système directif de la tête,

intégration de chemin) et tentent de rendre compte de leur possible fonctionnement au travers d'un modèle en accord avec les différentes données anatomiques ([7], [8]). La modélisation chez l'homme est quant à elle plus difficile puisque l'on ne dispose pas d'une base de données comparable, en conséquence de quoi les modèles résultants sont généralement plus globaux et tentent de rendre compte du fonctionnement global de l'hippocampe et de son interaction avec le cortex. Ces modèles reposent essentiellement sur l'utilisation de matrices auto-associatives et hétéro-associatives ([5]) et produisent un système capable de prototyper les entrées multimodales reçues via le cortex entorhinal et de guider le cortex dans l'apprentissage d'événements multimodaux ([9], [6]).

4.3 Architecture globale

L'architecture globale doit donc comporter un modèle de cortex et un modèle d'hippocampe capables d'interagir efficacement lors de la navigation. On peut voir figure 3 une vue générale de ce que devra être l'architecture. A un instant t , les différents stimuli externes (principalement la rétine, mais éventuellement des capteurs mesurant les distances aux

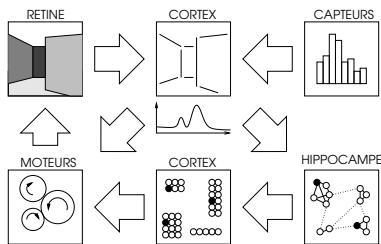


FIG. 3 – Architecture globale

obstacles) et internes (niveaux des différentes variables vitales, motivations) arrivent au niveau du cortex où un pré-traitement est effectué selon les différentes modalités. Cette information peut alors être utilisée de deux façons. D'une part, elle peut permettre au cortex de réagir immédiatement et d'envoyer des séquences de contrôle vers les moteurs, d'autre part, elle peut être acheminée en direction de l'hippocampe. On dispose donc au niveau de celui-ci d'une information multimodale rendant compte de l'ensemble de la situation de l'animat. Cette information est alors prototypée (ce qui permettra une mémorisation économique), mémorisée localement et transmise au cortex. Le modèle d'hippocampe va donc jouer un double rôle. Dans un premier temps il servira de mémoire de travail, et dans un deuxième temps, il servira de superviseur pour l'apprentissage au niveau du cortex en détectant ce qu'il est "pertinent" d'apprendre.

5 Conclusion

Nous disposons donc à l'heure actuelle d'un animat simulé capable d'évoluer de façon autonome

et cohérente dans un monde peu hostile. Cependant, l'absence de mémoire explicite handicape fortement son comportement puisqu'il ne peut mémoriser qu'un savoir faire (si j'ai soif, je dois aller vers du bleu). Il faut donc lui adjoindre un système autorisant la rétention de faits explicites et c'est là le rôle que nous souhaitons donner à l'hippocampe. En se projetant plus avant dans le futur, il serait souhaitable que cet animat, désormais pourvu de mémoire explicite, soit capable de planifier ses actions et donc d'émettre des hypothèses sur le comportement à adopter vis-à-vis d'un environnement et de motivations internes. Ceci supprimerait la capacité d'excitations et d'appels virtuels (les connexions motrices devant être inhibées), ce qui semblerait être l'un des rôles du cortex frontal.

Références

- [1] Lorraine G. Allan. Humain contingency judgments : Rule based or associative. *Psychological Bulletin*, 114(3):435–448, 1993.
- [2] Dana H. Ballard. Cortical connections and parallel processing : Structure and function. *The Behavioral and Brain Sciences*, (9):67–129, 1986.
- [3] Yves Burnod. *An adaptive neural network : the cerebral cortex*. Masson, 1989.
- [4] Tommi Jaakkola, Michael I. Jordan, and Satinder P. Singh. On the convergence of stochastic iterative dynamic programming algorithms. *Neural Network*, 6:1185–1201, 1994.
- [5] B. L. McNaughton and Lynn Nadel. Hebb-marr networks and the neurobiological representation of action in space. *Neuroscience and Connectionist Theory*, Laurence Erlbaum Associates, 1990.
- [6] Jaap M.J. Murre. *TraceLink : A Model of Amnesia and Consolidation of Memory*. Wiley-Liss, Inc., 1996.
- [7] J. O'Keefe and J. Dostrovsky. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Res.*, 34:171–175, 1971.
- [8] A. David Redish. *Contributions to a Computational Neuroscience Theory of Rodent avigation*. PhD thesis, 1997.
- [9] Edmund T. Rolls. Functions of neuronal networks in the hippocampus and neocortex in memory. *Neural Networks and Memory*, 13:241–265, 1990.
- [10] Richard S. Sutton and Andrew G. Barto. Toward a modern theory of adaptive network : Expectation and prediction. *Psychological Review*, 88(1):137–170, 1981.