

www.isir.fr

tel : 01 44 27 51 41  
fax : 01 44 27 51 45

Benoît Girard  
Chargé de Recherche CNRS, HDR  
ISIR - UPMC/CNRS  
Pyramide - Tour 55 Boîte courrier 173  
4 place Jussieu  
75252 Paris Cedex 05 - France  
e-mail : benoit.girard@isir.upmc.fr

Paris, le 23 août 2012

### Rapport sur le manuscrit de thèse de Mme Wahiba Taouali.

Wahiba Taouali présente, en vue de l'obtention d'un Doctorat de l'Université de Lorraine, une thèse intitulée « Modélisation de populations neuronales pour l'intégration visuo-motrice : Dynamiques et décisions ».

Le travail présenté concerne la modélisation computationnelle des régions du système nerveux impliquées dans l'oculomotricité. Il est composé de trois contributions distinctes, portant sur la convergence de champs de neurones dynamiques (*Dynamic Neural Fields* - DNF) mis à jour de manière asynchrone, la modélisation du colliculus supérieur (*superior colliculus* - SC) comme un DNF de géométrie complexe-logarithmique, et enfin la modélisation des ganglions de la base (*basal ganglia* - BG) et en particulier de leur capacité à sélectionner des comportements exploratoires.

Le manuscrit principal comporte 120 pages structurées en 6 chapitres, complété par deux annexes, la première reproduisant le texte des articles publiés et la seconde le paramétrage des modèles.

Un premier court chapitre introductif présente le cadre général de la thèse : l'étude du fonctionnement du système nerveux, sa modélisation dans un contexte incarné d'interactions continues avec l'environnement, le choix du système visuo-moteur comme objet d'étude et enfin l'objet des trois contributions proposées.

Le second chapitre comprend à la fois une présentation bibliographique des outils computationnels utilisés dans l'ensemble de la thèse et la première contribution sur la convergence des DNF asynchrones. Sont rapidement introduits les neurones et les différentes théories sur la manière dont l'information neuronale est encodée. S'ensuit une présentation des modèles canoniques de neurones artificiels (soit à potentiels d'action, soit à taux de décharge), l'évocation des réseaux de type perceptron puis DNF, un survol des modèles de plasticité neuronale, puis la présentation des questions de discrétisation spatiale et temporelle des DNF. Sur la base de la constatation que les simulations de DNF font toujours appel, de manière implicite ou explicite, à une horloge centralisée, la première

Sous la co-tutelle de

contribution du manuscrit s'attaque à la question de la convergence des DNF s'ils sont mis à jour de manière asynchrone.

La question est particulièrement pertinente et originale. Sur la base de travaux d'analyse numérique, les conditions et la vitesse de convergence théoriques des DNF asynchrones sont précisées, de sorte qu'il est conclu que les DNF peuvent être, en règle générale, simulés de manière asynchrone. Ce résultat est d'un grand intérêt, en particulier dans le contexte de l'implémentation totalement parallèle d'algorithmes autrement très coûteux à mettre en œuvre sur des machines strictement séquentielles. Enfin, une implémentation asynchrone de DNF est réalisée sur la base du modèle de Rougier et Vitay (2006). On pourra regretter que celle-ci ne soit qu'à peine esquissée : il est indiqué que la vraie solution a été approchée avec précision, que le biais est négligeable, mais aucune évaluation chiffrée de ces éléments n'est fournie ; il est également indiqué que les fonctions qualitatives d'entrée/sortie des DNF ont été reproduites, mais elle ne sont pas présentées. Enfin, il est mentionné qu'un mécanisme de mise à jour parcimonieux a été testé, mais il aurait mérité d'être présenté plus clairement.

Globalement, ce chapitre aurait gagné à être plus sélectif sur les points traités en bibliographie, en ne conservant que ceux qui relèvent directement des thématiques abordées, et en étant plus détaillé et précis sur ceux-ci. Ainsi, la présentation des théories du codage neuronale est très succincte, et sa nécessité n'apparaît pas clairement, les travaux proposés ne relevant pas particulièrement de ce domaine spécifique. De la même manière, le paragraphe sur le connexionisme et les réseaux de neurones n'évoque que le perceptron avant de passer aux DNF, ce qui fait l'impasse sur de nombreux autres travaux (par exemple, les modèles de Hopfield, de Grossberg, les *echo state networks*, etc.), néglige de parler de l'apprentissage supervisé qui est pourtant le cœur du perceptron, tout en ne parlant pas du perceptron lorsqu'est évoqué, ensuite, l'apprentissage supervisé. Enfin, la section introduisant la plasticité et l'apprentissage traite des apprentissages supervisé, non-supervisé et par renforcement en moins d'une page et de manière trop superficielle.

Le troisième chapitre présente les mouvements oculaires et l'ensemble des structures du système nerveux qui y participent : répertoire des mouvements oculaires, rétine, aires visuelles, voies ventrales et dorsales de traitement de l'information, lobe pariétal, lobe frontal, colliculus supérieur, ganglions de la base, formation réticulée et cervelet. Il se termine par des considérations sur les transformations spatio-temporelles des cartes des champs oculaires frontaux (FEF) et du SC vers la commande des générateurs de saccades. Il est clair et offre un panorama synthétique et assez complet de l'ensemble du système oculomoteur, considérant que les détails concernant l'anatomie, la physiologie et les modèles du SC et des BG sont reportés aux chapitres suivants.

Le quatrième chapitre est consacré à une seconde contribution, portant sur le colliculus supérieur. Il commence par une présentation plus approfondie du SC : organisation en couches de cartes du champ visuel, topologie complexe-logarithmique de ces cartes, connexions avec le reste du cerveau, classification fonctionnelle et électrophysiologique des neurones enregistrés. Les nombreux modèles du SC existants sont ensuite évoqués, sans être analysés en détails (avec un notable retour sur la question de la transformation spatio-temporelle), la méthodologie adoptée consistant en la construction d'un modèle minimal du SC,



puis en l'évaluation des propriétés qu'un tel modèle est suffisant pour expliquer. Dans le cadre d'un foisonnement de modèles complexes et très différents, une telle approche semble a priori salutaire.

La typologie neuronale du SC rapportée aurait gagnée à être complétée, par exemple au vu des catégories proposées par (McPeck & Keller 2002, J Neurophysiol). Par ailleurs, les explications complémentaires de ce chapitre sur la transformation spatio-temporelle ne considèrent que les hypothèses de *vecteur somme* et *vecteur moyenne*, alors qu'un autre schéma dit de *sommation-inhibition* proposé par (Groh, 2001, Biol Cybern) – qui a reçu des supports expérimentaux (Goossens & van Opstal, 2006, J Neurophysiol) et théoriques (Tabareau et al., 2007, Biol Cybern) – permet également d'expliquer les observations expérimentales de (Lee et al., 1988, Nature).

Le modèle proposé est essentiellement un DNF dont les entrées proviennent d'une entrée visuelle déformée par la géométrie particulière des cartes colliculaires. Le premier effet résultant de cette combinaison est la décroissance, avec l'excentricité, de l'activité maximale de la bulle générée dans le DNF. La cause de cette observation n'est malheureusement pas analysée, alors que, d'une part, cela n'est pas une conséquence évidente du modèle; d'autre part, qu'elle est probablement en partie dépendante de la paramétrisation choisie du noyau de connexions latérales du DNF; enfin, qu'elle est particulièrement intéressante au regard des données expérimentales qui existent sur le sujet. En effet, il a été noté en introduction de ce chapitre que certains modèles, comme (Godjin & Theeuwes, 2002, J Exp Psych), utilisent une sensibilité des neurones décroissante avec leur excentricité : il ne s'agit pas d'une « astuce » de modélisateur, mais d'une observation expérimentale déjà présente dans (Anderson et al., 1998, J Neurophysiol) et confirmée dans (van Opstal & Goossens, 2008, Biol Cybern), dont la cause n'est pas clairement établie. Nous aurions donc peut-être ici une proposition de mécanisme, qu'il reste à décortiquer. Le comportement du modèle est ensuite mesuré en terme de précision des saccades générées, des effets d'inactivations comparables à celles opérée dans (Lee et al., 1988, Nature), de latence des saccades en fonction des propriétés des cibles (excentricité, saillance, taille), et des effets de la compétition entre deux cibles sur la nature de la saccade générée (sélection ou fusion) et sa latence.

Il est cependant dommage que certaines conclusions n'aient pas été creusées plus avant. En particulier sur les prédictions spécifiques au modèle proposé, alors qu'il y aurait eu matière, en particulier sur la décroissance d'activité dépendante de l'excentricité ainsi que sur les tendances des temps de latence des saccades. D'autre part, un bilan des effets expliqués par ce modèle minimaliste par comparaison à la manière dont ils sont obtenus dans les modèles antérieurs potentiellement plus complexes aurait permis de mieux situer a posteriori les contributions du travail réalisé.

Le dernier chapitre de contributions porte sur la modélisation des ganglions de la base dans leur rôle de sélection. Les ganglions de la base sont d'abord présentés en terme de noyaux les composants, de leur connectivité, de leur organisation globale en boucles cortico-basales et de leurs fonctions supposées. Sur ce point, la distinction fonctions motrices/cognitives n'est pas claire : la structure des circuits moteurs est la même que celle des circuits cognitifs, et les facultés d'apprentissage, reportées à la partie sur les circuits cognitifs a pourtant toutes les raisons d'être aussi présente dans les circuits moteurs. Une distinction sé-

lection/apprentissage aurait peut-être été plus appropriée.

Sont ensuite présentés de manière claire et didactique le modèle d'interprétation de la connectivité proposés par Albin et al. en 1989, sa réinterprétation proposée dans le modèle computationnel GPR de Gurney et al. (2001, Biol Cybern), et enfin la modification du GPR proposée par Girard et al. (2008, Neur Ntwk), le modèle CBG. C'est sur ces bases et en comparaison avec ces deux modèles qu'est proposé un nouveau modèle des BG.

Il s'agit d'un modèle dont la connectivité est légèrement différente du GPR (la distinction D1/D2 dans le striatum, dont la pertinence est en effet très discutable, a été supprimée) et du CBG (disparition du TRN, des FSI, des projections du GPi vers le Striatum). Son originalité réside, d'une part, dans l'addition de bruit dans le striatum, afin de permettre occasionnellement la sélection d'actions jugées a priori non optimales (exploration), puis d'autre part, de la stabilisation de cette sélection par l'utilisation d'inhibitions latérales dans le cortex frontal et de la projection thalamus-striatum, ce qui répercute le choix final dans l'ensemble du circuit. La question de la mise en œuvre de l'exploration, dont on sait qu'elle est nécessaire pour la convergence des algorithmes d'apprentissage par renforcement couramment utilisés pour modéliser les BG, n'est pour autant en général pas explicitement modélisée (voir par exemple Humphries et al. 2012, Frontiers). Seul les travaux –peu convaincants– de l'équipe de Chakravarthy se sont aventurés sur ce terrain. La question abordée apparaît donc de première importance, et l'est ici avec des outils de modélisation adaptés.

Les premiers résultats obtenus montrent l'efficacité du système pour la fonctionnalité désirée, à savoir sa capacité à alterner entre deux options en compétition si elles ont des valeurs d'entrées proches. Par comparaison, le GPR et le CBG convergent toujours vers l'option la plus forte. Il aurait cependant été intéressant, pour une comparaison à armes égales, d'ajouter le même bruit striatal à ces deux modèles, pour mettre en évidence de manière claire leur incapacité à se fixer occasionnellement mais de manière durable sur l'option la moins forte. Il aurait aussi été utile d'effectuer des tests systématiques des ratios de choix *winner-takes-all* vs. exploratoire, en fonction de la valeur relative des options mises en compétition.

Un second résultat consiste en la comparaison du profil d'activation du GPi simulé et de sa comparaison avec les mesures présentées dans (Nambu et al., 2000, J Neurophysiol). Cette comparaison semble inadaptée, d'une part parce que les stimulations réalisées dans cette expérience sont très éloignées d'activations corticales physiologiques; ensuite parce que le profil d'activité du STN de Nambu et al. ne ressemble, lui, pas du tout à celui du modèle présenté; enfin parce que l'intégration étant réalisée avec la méthode d'Euler, il ne faut pas prêter une attention trop forte aux activités des phases transitoires.

Enfin, un mécanisme simple d'apprentissage est développé afin de pouvoir activer une sélection guidée par une information contextuelle, éventuellement capable de se passer ensuite des entrées sensorielles.

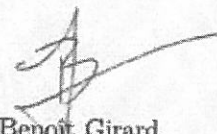
La conclusion du manuscrit reprend l'ensemble des résultats obtenus en mettant en avant les liens qui unissent les contributions variées réalisés durant cette thèse, et met en avant un certain nombre de développements futurs qui pourraient venir compléter les travaux présentés ici.

Au vu de cet ensemble, il apparaît que Wahiba Taouali présente des contribu-



tions intéressantes et sérieuses sur la modélisation du cerveau en général et celle du circuit oculomoteur en particulier.

Pour ces raisons, j'émet un avis favorable à ce que Wahiba Taouali soit autorisé à présenter publiquement ses travaux en vue de l'obtention d'un Doctorat de l'Université de Lorraine.



Benoît Girard  
Chargé de recherche au CNRS