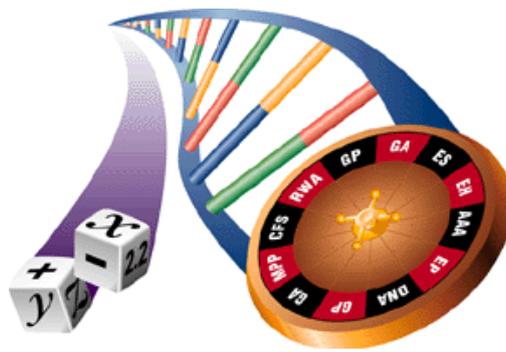


La formation réticulée médiane : un substrat pour la sélection de l'action ? Modélisation via réseaux de neurones et algorithmes évolutionnistes.



Franck Deroncourt
franck.deroncourt@gmail.com

Supervisé par Stéphane Doncieux et Benoît Girard



Paris, Juin 2011

Table des matières

Table des matières	i
Table des figures	iii
1 Introduction	3
1.1 Remarques liminaires	3
1.2 Sélection de l'action	5
1.3 Données anatomiques internes	7
1.4 Démonstration de $P(l) > 45 \times P(p)$	9
1.5 Modèles existants	10
1.5.1 Le modèle de Kilmer-McCulloch - 1969	10
1.5.2 Le modèle de Humphries - 2006	13
1.6 Démarche et objectifs de ce rapport	14
2 Matériel théorique	15
2.1 Les réseaux de neurones	15
2.2 Les algorithmes évolutionnistes	18
2.2.1 Définitions	18
2.2.2 Application	20
3 Tâche désincarnée : la tâche vectorielle abstraite	23
3.1 Expérience	23
3.2 Résultats	26
4 Tâche incarnée : la tâche de survie minimale	31
4.1 Expérience	31
4.2 Résultats	34
5 Discussions et perspectives	40
6 Conclusion	44
Bibliographie	46

A	Notes techniques sur l'implémentation	51
B	Exemple d'une mRF à 4 clusters	53
C	Exemple d'un cluster de la mRF	55

Table des figures

1.1	Tronc cérébral chez un cerveau humain (partie en rouge)	4
1.2	Localisation de la mRF	4
1.3	Projections sur et depuis la mRF	4
1.4	Coupes du cerveau au niveau postérieur du tronc cérébral.	6
1.5	Coupe sagittale de la mRF	8
1.6	mRF et structures smallworld	11
1.7	Modèle de Humphries	13
2.1	Exemple d'un neurone avec 2 entrées	16
2.2	Exemple d'une mRF à 4 clusters	17
2.3	Exemple d'un cluster de la mRF	17
2.4	Fonctionnement d'un algorithme évolutionniste	19
2.5	Exemple de front de Pareto de dimension 2	21
2.6	Mono vs multi-objectif	21
3.1	Tâche désincarnée	24
3.2	Front de Pareto de dimension 2	27
3.3	Front de Pareto de dimension 3	28
3.4	Front de Pareto 2D sans contraintes lors des mutations	29
3.5	Front de Pareto 2D de l'évolution sans aucune contrainte	30
4.1	Environnement de la tâche de survie minimale	32
4.2	Comparaison des contrôleurs aléatoire et WTA dans la tâche de survie	35
4.3	Front de Pareto 2D de l'évolution d'un contrôleur mRF	36
4.4	Comparaison des contrôleurs dans la tâche de survie	37
4.5	Comparaison des contrôleurs dans la tâche de survie	37
4.6	Front de Pareto 2D de l'évolution d'un contrôleur mRF avec modulation	38
4.7	Front de Pareto 2D de l'évolution d'un contrôleur mRF sans saliences	39
A.1	Statistiques sur le code source	52
B.1	Exemple d'une mRF à 4 clusters	54
C.1	Exemple d'un cluster de la mRF	56

Résumé

La formation réticulée médiane (mRF) est une région du tronc cérébral recevant des entrées sensorielles très riches, en mesure de contrôler les actions motrices via ses connexions vers la moelle spinale et les nerfs crâniens. Phylogénétiquement, la mRF est l'une des plus anciennes structures nerveuses du tronc cérébral, lui-même étant ce qu'on peut considérer comme un des centres les plus anciens du système nerveux central. Elle semble donc constituer un proto système de sélection de l'action.

Le premier modèle de la mRF a été proposé par [Kilmer et al., 1969], il interprète déjà le rôle de cette région comme un sélecteur de mode de fonctionnement. [Humphries et al., 2005] ont voulu tester l'efficacité de ce modèle en tant que mécanisme de sélection de l'action de bas niveau en le confrontant à la tâche de survie minimale proposée dans [Girard et al., 2003]. Le modèle original s'est avéré peu efficace, ce qui a conduit [Humphries et al., 2006] à proposer un second modèle exprimé dans un formalisme de réseau de neurones et prenant quelques données anatomiques et physiologiques supplémentaires. Néanmoins, ce second modèle n'a qu'un succès mitigé sur la tâche de survie minimale et est assez peu plausible anatomiquement en raison de sa forte abstraction.

Nous proposons dans ce rapport un nouveau modèle de la mRF :

- contraint par les informations anatomiques quant à sa structuration ;
- construit sur la base de réseaux de neurones générés par évolution artificielle ;
- évalué sur des tâches de sélection de l'action.

Le modèle obtenu réussit avec succès les tâches de sélection, ce qui montre que la mRF peut servir de système de sélection de l'action. Nous démontrerons également une propriété anatomique de la mRF qui couplée aux résultats de l'article [Humphries et al., 2006] montre qu'il est très probable que la mRF réseau ait une structure de type small-world.

Ce projet est financé par l'ANR (ANR-09-EMER-005-01) dans le cadre du projet EvoNeuro (<http://pages.isir.upmc.fr/EvoNeuro>).

Remerciements

Je souhaite tout d'abord exprimer ma plus profonde gratitude envers mes deux encadrants, Stéphane Doncieux et Benoît Girard : outre leurs précieux conseils indispensables pour mener à bien ce projet, ils ont suivi avec minutie l'avancée de mes travaux tout au long du stage et leurs exigences élevées m'ont poussé à donner le meilleur de moi-même. La complémentarité de leurs compétences, Stéphane étant spécialiste de l'évolution artificielle et Benoît expert en modèles neuro-computationnels de sélection de l'action, revêtait une importance cruciale dans la réalisation des travaux que nous allons présenter dans ce rapport et m'a permis personnellement de gagner considérablement en compétences sur ces deux problématiques en quelques mois seulement.

Je tiens également à remercier très chaleureusement Jean Liénard, Tony Pinville et Paul Tonelli, doctorants à l'ISIR : leur aide généreuse et leurs avis spontanés ont orienté mes réflexions vers des chemins plus fructueux, notamment en partageant leurs diverses expériences de thèse.

Enfin, mes pensées de remerciements vont à toutes les autres personnes du CogMaster et de l'ISIR avec qui j'ai pu échanger sur mon projet de stage et qui m'ont permis par leurs réflexions de mûrir ce projet et d'avancer dans mes recherches, et en particulier à mon tuteur du CogMaster et d'HEC, Philippe Mongin, dont la gentillesse, la perspicacité et l'ouverture d'esprit pour ne citer que quelques-unes de ses qualités sont sans limite. Beaucoup de choses n'eurent pas été possibles sans lui.

Chapitre 1

Introduction

Ce chapitre présente une vue d'ensemble de la formation réticulée médiane : après quelques remarques préliminaires, nous résumerons des études qui tendent à montrer que la mRF est un proto-système de sélection de l'action. Dans un second temps, nous présenterons de façon synoptique l'ensemble des données anatomiques connues de la mRF permettant de la modéliser. Enfin, nous étudierons les deux et uniques modèles existants de la formation réticulée médiane afin de voir leurs lacunes, tout en tirant parti de leurs points forts pour construire le nôtre.

1.1 Remarques liminaires

La formation réticulée, dont fait partie la formation réticulée médiane, est un composant du système nerveux central qui doit son nom (tiré du latin *reticulum* signifiant filet) à sa présentation anatomique en réseau dense et intriqué. Située sur le plancher du tronc cérébral entre la moelle allongée et le mésencéphale, elle projette et reçoit des fibres nerveuses de toute la moelle spinale ainsi que de nombreux autres composants du névraxe tels que le cortex cérébral, l'hypothalamus ou le cervelet (cf figures 1.1, 1.2 et 1.3).

Nous nous intéresserons dans ce rapport à une partie de la formation réticulée appelée formation réticulée médiane (ou médiale), que nous abrègerons par mRF dans la suite d'après la traduction anglaise *medial Reticular Formation*.

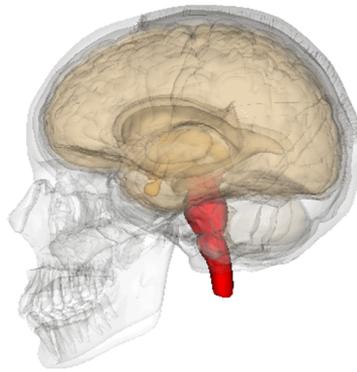


FIGURE 1.1: Tronc cérébral chez un cerveau humain (partie en rouge)

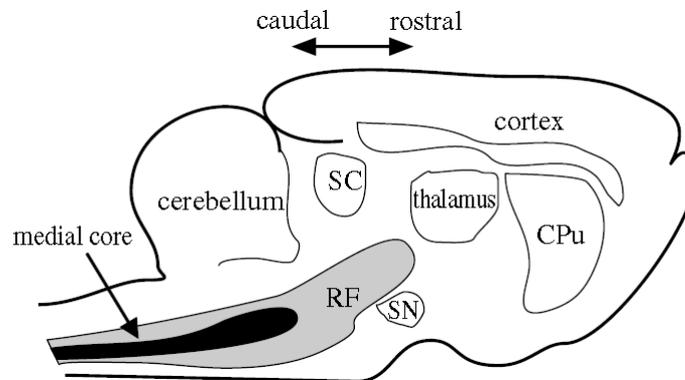


FIGURE 1.2: Localisation de la mRF (zone en noir) dans le cerveau d'un chat. RF : reticular formation. CPu : caudate-putamen. SC : superior colliculus. SN : substantia nigra. Source : [Humphries et al., 2006]

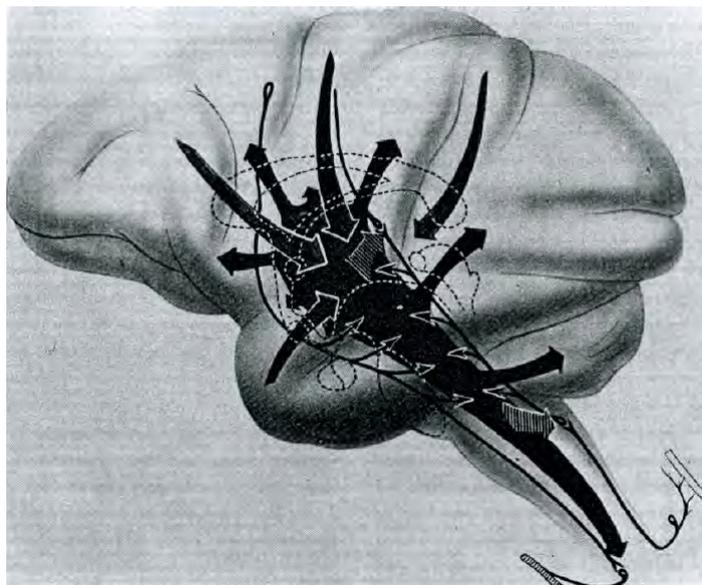


FIGURE 1.3: Projections sur et depuis la mRF. Source : [Humphries et al., 2005]

La phylogénèse du système nerveux montre que la mRF est l'une des plus anciennes structures nerveuses du tronc cérébral, lui-même étant ce qu'on peut considérer comme un des centres les plus anciens du système nerveux central. Ainsi, parmi les différentes espèces animales, la mRF est très comparable, comme le montre par exemple [Ramón-Moliner and Nauta, 1966] entre les requins et les humains, ce qui nous permettra d'agréger directement les données issues d'études sur des animaux différents.

1.2 Sélection de l'action

La sélection de l'action désigne le problème crucial que rencontre tout agent autonome, qu'il soit de nature animale ou robotique, qui consiste à choisir et coordonner continuellement ses comportements dans l'optique de mener à bien ses objectifs au long terme, tels que la survie, la reproduction, ou tout autre tâche définie par le concepteur dans le cas d'un robot. Les animaux doivent donc nécessairement mettre en œuvre des solutions efficaces de sélection de l'action, d'où notre recherche au sein du système nerveux d'un tel mécanisme.

Plusieurs données tendent à montrer que la mRF est impliquée dans la sélection de l'action, nous allons les résumer succinctement dans cette section.

Tout d'abord, la mRF semble disposer de toutes les informations auxquelles ont accès les systèmes de monitoring et sensoriels aussi bien externe qu'interne d'un animal : elle reçoit ainsi une quantité considérable d'entrées sensorielles, comme le synthétise [Humphries et al., 2007], provenant notamment des systèmes sensoriel, respiratoire, viscéral, vestibulaire, proprioceptif, nociceptif ou encore cardio-vasculaire. Ces données sont corroborées par les divers enregistrements effectués sur la mRF montrant que cette dernière réagit à des stimuli très variés [Segundo et al., 1967, Bowsher, 1970, Langhorst et al., 1983].

De façon complémentaire, la mRF projette en masse sur tous les niveaux de la moelle spinale ainsi que sur les nerfs crâniens [Torvik and Brodal, 1957, Eccles et al., 1976, Jones, 1995], ce qui lui donne la capacité de contrôler aussi bien la musculature axiale que la face. Par conséquent, la mRF dispose des entrées et des sorties nécessaires à tout candidat de système de sélection de l'action.

Expérimentalement, il a été montré que des rats qui avaient subi une coupe complète au niveau postérieur du tronc cérébral, plus précisément postérieur au thalamus et à l'hypothalamus, en enlevant toute la partie rostrale du cerveau par rapport à cet axe de coupe (cf les 3 axes de coupe sur la figure 1.4), présentaient un comportement

surprenamment cohérent [Woods, 1964], hormis les erreurs engendrées par la perte de la vision, de l'odorat ainsi que des troubles de la régulation hormonale. Les rats de l'expérience étaient toujours capable d'effectuer des sélections de l'action de bas niveau, tel manger, se mouvoir, boire ou dormir en fonction des stimuli. Les résultats très importants de cette expérience furent plus tard confirmés par [Lovick, 1972, Berntson and Micco, 1976, Berridge, 1989], mettant l'accent sur le fait que ces rats peuvent effectuer des combinaisons d'actions coordonnées, tel tenir, ronger et mâcher de la nourriture, combinaisons plus complexes que de simples réflexes qui pourraient émaner de la moelle spinale.

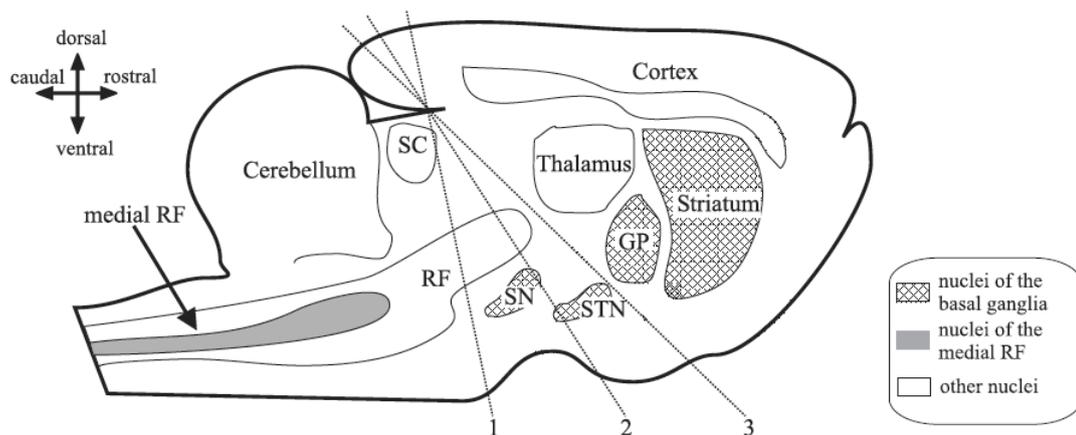


FIGURE 1.4: Coupes du cerveau au niveau postérieur du tronc cérébral. Les lignes en pointillés montrent 3 coupes, toute la partie du cerveau située du côté rostral des lignes a été enlevée. GP : globus pallidus. RF : reticular formation. SN : substantia nigra. STN : subthalamic nucleus. SC : superior colliculus. Source : [Humphries et al., 2007]

Pour vérifier le rôle de la mRF, des stimulations locales électriques et chimiques ont été faites sur des animaux normaux, stimulations qui eurent pour réaction de les faire changer de comportement, par exemple manger, dormir, boire, s'enfuir, chercher ou encore se laver [Magoun and Rhines, 1946, Glickman and Schiff, 1967]. Ces résultats montrent l'ampleur des différents comportements que contrôle au moins en partie la mRF.

A contrario, des études ont porté sur l'impact de lésions à la mRF. Elles témoignent de troubles sévères du comportement, notamment des troubles du sommeil, les sujets étudiés montrant d'un déphasage entre le sommeil cérébral et le reste du corps [Birkmayer and Pilleri, 1966], ou aussi des alternances fréquentes entre sommeil profond et rage extrême [Jouvet, 1967]. [Parvizi and Damasio, 2003] ont même montré

que des lésions de certaines parties de la mRF peuvent provoquer le coma, voire la mort chez un être humain.

Enfin, contrairement à la plupart des structures neuronales, les cellules de la mRF existent dès la naissance [Hammer Jr et al., 1981], ce qui peut représenter une indication de l'importance de leur présence pour la survie de l'individu.

À la lueur de ces différentes études, il semble raisonnable de supposer que la mRF constitue un proto-système de sélection de l'action, proto car les actions en question semblent être d'un niveau assez rudimentaire et en raison de l'ancienneté phylogénétique.

1.3 Données anatomiques internes

Dans cette section, nous allons rassembler toutes les données anatomiques qui seront utiles pour construire un modèle de la mRF.

Première remarque d'ordre général, la littérature neuroscientifique montre que les mécanismes de sélection de l'action sont implémentés suivant deux architectures différentes dans le cerveau des mammifères :

- une architecture de type centralisée, où chaque module neuronal projette sur un système de décision central. Les ganglions de la base auraient ce type d'architecture [Redgrave et al., 1999, Prescott et al., 1999].
- une architecture de type modulaire, où chaque module neuronal est en compétition, inhibant ou excitant les autres modules. La mRF aurait ce type d'architecture [Humphries et al., 2007].

Une des particularités de la mRF est son organisation en clusters (également appelés stacks ou chips selon les articles) [Scheibel and Scheibel, 1967], comme le montre la coupe sagittale de la mRF présentée sur la figure 1.5.

La mRF d'un rat mesure environ 7 mm sur l'axe rostral-caudal et contient entre 35 et 75 clusters [Humphries et al., 2006]. La mRF d'une grenouille contient approximativement 0.75 million de neurones, tandis que celle d'un être humain en contient 2 millions [Kilmer et al., 1969]. Les clusters sont de la même taille en terme de dimension et de nombre de neurones [Humphries et al., 2007]. Ils reçoivent tous les mêmes entrées et projettent vers les mêmes zones [Humphries and Prescott, 2006].

Il existe deux catégories principales de neurones dans la mRF :

- Les **interneurones** : de taille petite ou moyenne, ils sont tous **inhibiteurs** (quelques exceptions existent cependant) et projettent quasi exclusivement dans leur propre

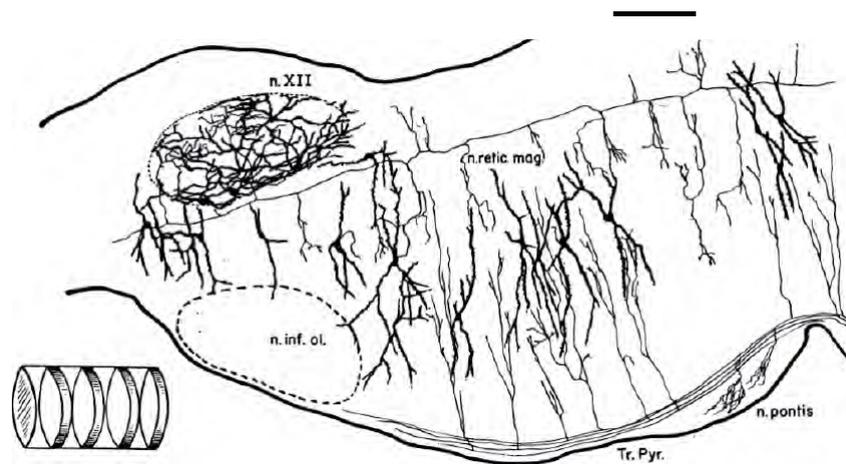


FIGURE 1.5: Coupe sagittale de la mRF sur un rat, montrant son organisation en clusters. Source : [Scheibel and Scheibel, 1967]

cluster, aussi bien sur des neurones de projection que sur des interneurones. Ils représentent environ 20% des neurones de la mRF [Humphries et al., 2006].

- Les **neurones de projection** : de taille moyenne à très grande, ils sont tous **excitateurs** (quelques exceptions existent cependant) et ils projettent quasi exclusivement en dehors de leur propre cluster, vers d'autres clusters ou bien en dehors de la mRF [Humphries et al., 2006]. Environ 45% des synapses en entrée des neurones de projection sont GABAergiques, c'est-à-dire inhibiteurs, donc provenant d'interneurones [Humphries et al., 2007]. Les neurones de projection représentent environ 80% des neurones de la mRF.

La probabilité notée $P(c)$ qu'un neurone de projection projette sur un cluster donné est l'objet de deux modèles différents dans la littérature. Selon [Grantyn et al., 1987], $P(c) = 0.25$ quel que soit le cluster source (où se trouve le noyau du neurone de projection) et le cluster destination (vers lequel projette le neurone de projection). Selon [Kilmer et al., 1969], la probabilité $P(c)$ dépend de la distance entre le cluster source et le cluster destination : $P(c) = d^{-a}$, où d est la distance et a un paramètre supérieur ou égal à 1.

La probabilité notée $P(p)$ qu'un neurone de projection projette sur un neurone donné, sachant qu'il projette dans le cluster de ce dernier, ainsi que la probabilité notée $P(l)$ qu'un interneurone forme une connexion avec un neurone donné de son cluster sont toutes les deux faibles, probablement inférieures à 0.1 d'après [Schuz, 1998, Albert and Barabási, 2002].

En résumé, étant donné les contraintes anatomiques, six paramètres décrivent

entièrement la structure du réseau :

- c : le nombre de clusters (entre 35 et 75) ;
- n : le nombre de neurones dans un cluster (environ $1500000/55 \approx 30000$) ;
- p : le pourcentage de neurones de projection (environ 80%). Le pourcentage d'interneurones est donc de $1 - p$;
- $P(c)$: la probabilité qu'un neurone de projection projette sur un cluster donné ($P(c) = 0.25$ ou bien $P(c) = d^{-a}$ avec d la distance et a un paramètre) ;
- $P(p)$: la probabilité qu'un neurone de projection projette sur neurone donné, sachant qu'il projette dans le cluster de ce dernier ($P(p) < 0.1$) ;
- $P(l)$: la probabilité qu'un interneurone forme une connexion avec un neurone donné de son cluster ($P(l) < 0.1$).

Après avoir présenté les données anatomiques connues sur la mRF, nous allons maintenant présenter une démonstration que nous avons réalisé au cours des recherches anatomiques montrant que $P(l) > 45 \times P(p)$.

1.4 Démonstration de $P(l) > 45 \times P(p)$

Soient, en plus des notations vues dans la section précédente :

- $P(I \rightarrow I)$ la probabilité qu'un interneurone forme une connexion avec un interneurone donné du même cluster (le signe \rightarrow symbolise une connexion dans le graphe) ;
- $P(I \rightarrow P)$ la probabilité qu'un interneurone forme une connexion avec un neurone de projection donné du même cluster ;
- nb_I le nombre d'interneurones dans un cluster de la mRF ;
- nb_P le nombre de neurones de projection dans un cluster de la mRF.

Supposons :

- qu'un interneurone ne projette que dans son propre cluster (le nombre de projections d'interneurones en dehors de leur cluster est négligeable) ;
- que $P(c) = 0.25$. C'est un des deux modèles anatomiques connus pour $P(c)$. Le second modèle, faisant dépendre $P(c)$ de la distance entre le cluster source et le cluster destination, favorise une structure de type small-world par rapport à 0.25 comme le montre la figure 1.6. Par conséquent, l'hypothèse prise ici permettra de généraliser le résultat de la démonstration au second modèle, étant donné que pour obtenir une structure small-world, intuitivement et comme le montre aussi la figure 1.6, il faut un nombre élevé de connexions dans un même cluster (défini par $P(l)$) par comparaison au nombre de connexions inter-cluster (défini par $P(p)$), ce que favorise le second modèle par rapport au premier où $P(c) = 0.25$.

Par construction :

- nous avons $P(l) = P(I \rightarrow I) + P(I \rightarrow P)$ (car nous considérons qu'un interneurone ne projette que dans son propre cluster) ;
- nous avons $nb_P/(nb_P + nb_I) = 80\%$ (c'est la proportion de neurones de projection, connue dans la littérature), donc $nb_P = 4 \times nb_I$.

Comme nous l'avons vu précédemment, 45% des synapses de neurones de projection sont GABAergiques. Autrement dit, 45% des connexions entrantes sur un neurone de projection proviennent d'un interneurone, car les interneurones sont tous inhibiteurs et GABAergiques (il se peut qu'il y ait des exceptions, mais négligeables).

Formellement, cela signifie que $\frac{nb_I \times P(I \rightarrow P)}{nb_I \times P(I \rightarrow P) + nb_P \times P(p) \times P(c) \times c} = 45\%$
ce qui est équivalent à $nb_I \times P(I \rightarrow P) = (45\%/55\%) \times (nb_P \times P(p) \times P(c) \times c)$
ce qui peut s'écrire également $P(I \rightarrow P) = 4 \times (45\%/55\%) \times (P(p) \times P(c) \times c)$
(car $nb_P = 4 \times nb_I$).

Remplaçons par les valeurs par $P(c) = 0.25$ et $c = 55$ (moyenne entre 35 et 75) :
Cela donne donc $P(I \rightarrow P) = 45 \times P(p)$
Or $P(l) = P(I \rightarrow I) + P(I \rightarrow P)$
D'où $P(l) > 45 \times P(p)$

Ce résultat était fortement la thèse selon laquelle la mRF a une structure de type small-world [Humphries et al., 2006], comme le montre la figure 1.6.

Après avoir présenté la structure de la mRF, nous allons maintenant analyser les deux uniques modèles publiés de la mRF.

1.5 Modèles existants

1.5.1 Le modèle de Kilmer-McCulloch - 1969

Au cours des années 1960, W.L. Kilmer, W.S. McCulloch, and J. Blum ont publié plusieurs articles proposant un premier modèle de la mRF en se basant sur les études anatomiques de Scheibel & Scheibel. En 1969, ils ont synthétisé l'ensemble de leurs recherches dans un seul article, qui devint un article de référence sur le sujet [Kilmer et al., 1969].

Le fil conducteur de leur modèle est le concept de mode de fonctionnement : un animal à un instant donné suit un et un seul mode de fonctionnement, comme par exemple manger ou dormir. La mRF a pour fonction de passer d'un mode à un autre. Pour appuyer cette hypothèse, lorsque la RF est endommagée il est constaté des changements de mode de fonctionnement pathologique [Jouvet, 1967]. Il est posé comme hypothèse qu'à un cluster est associé à un et un seul mode de fonctionnement.

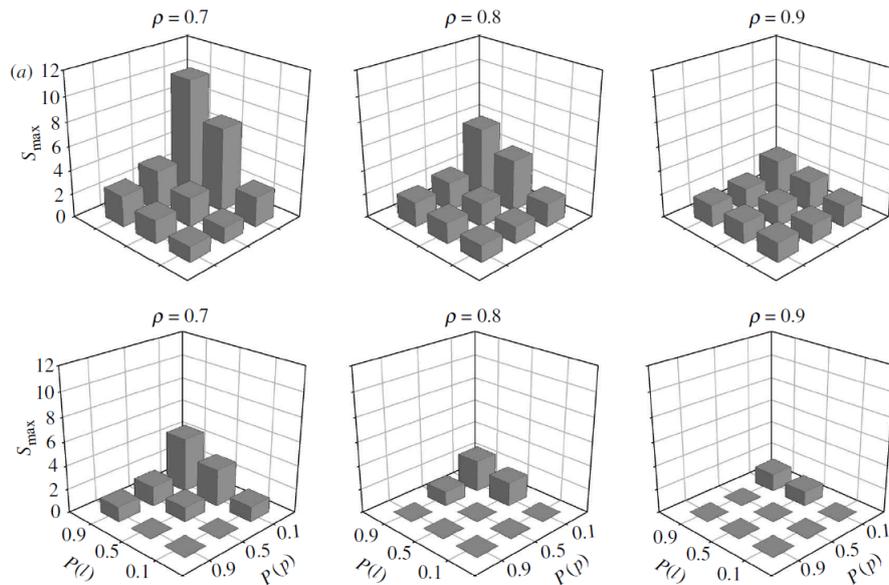


FIGURE 1.6: Cette figure montre à quel degré la mRF peut être considérée comme étant un réseau de type small-world selon les valeurs de $P(l)$, ($P(p)$) et p . En hauteur, la valeur S_{max} montre le degré de small-world : si elle est supérieure à 1, alors le réseau est considéré comme étant small-world. Nous voyons clairement que la démonstration de $P(l) > 45 \times P(p)$ accroît fortement les chances que la mRF ait une structure de type small-world. Source : [Humphries et al., 2006]

Trois variantes du modèle sont proposées :

1. S-RETIC (S pour Simple, Retic pour Reticular) : Comme le nom l'indique, ce premier modèle se veut simple et directement issu des données anatomiques. Ce modèle est composé d'une douzaine de modules recevant des stimuli et devant déterminer le mode à choisir, un module correspondant à un cluster de la mRF. Ces modules reçoivent des informations leur permettant d'être à la fois les généralistes, afin de pallier toute défaillance de modules, et spécialisés afin d'être à même de prendre une décision. Les modules sont plus ou moins liés entre eux en fonction de la distance qui les sépare. Chaque module prend une décision et lui donne une probabilité, et la décision finale prend en compte l'ensemble des modules. Si le consensus pour un mode est assez grand, il y a alors convergence vers ce mode. Ce modèle peut facilement être étendu à davantage de modes et modules. Ce modèle donne des résultats corrects, cependant il souffre de plusieurs manques, que la seconde variante va essayer de combler.
2. STC-RETIC (STC pour Spatio-Temporal Conditioning) : c'est un modèle enrichissant S-RETIC en introduisant les notions de développement, généralisation,

discrimination, habitude et conditionnement. Pour introduire ses propriétés, il faut fournir des informations supplémentaires permettant d'indiquer si un stimuli ou un choix de mode de fonctionnement est bon, neutre ou mauvais, à l'instar de l'apprentissage par renforcement. Cet apprentissage se fait au niveau de chaque module et de façon coopérative entre les différents modules. Les résultats pour ce modèle ne sont pas détaillés. L'article souligne que STC-RETIC a plusieurs caractéristiques peu attrayantes : ses connexions ne respectent pas strictement le modèle anatomique de la RF et il ne peut changer de mode que si un nouveau stimulus arrive.

3. H-RETIC (H pour Hardware) : c'est une version hardware du modèle, mais en cours d'élaboration au moment de l'écriture de l'article, et jamais achevée par la suite, STC-RETIC étant implémenté en un software devenu trop compliqué à maintenir et développer efficacement avec les outils de l'époque.

Néanmoins, les résultats obtenus par les auteurs avec ce modèle ne sont jamais vraiment exposés, et les imprécisions concernant sa description font qu'il est difficile d'implémenter le modèle pour l'analyser concrètement. En outre, beaucoup de considérations de l'article sont certes intéressantes, mais nullement explorées : ainsi exposé, le modèle ne satisferait pas les exigences de publications scientifiques actuelles. Il en résulte que ce premier modèle permet davantage d'avoir une vue d'ensemble sur le mRF tout en donnant des pistes intéressantes, à défaut d'être précisément décrit et réellement utilisable.

Pendant presque 40 ans, aucun nouveau modèle de la mRF ne fut proposé. En 2005, Mark Humphries [Humphries et al., 2005] tenta de reproduire le modèle de Kilmer-McCulloch et, afin d'évaluer ses performances, il implémenta le modèle dans un robot simulé ainsi que dans un robot réel. Ce dernier fut placé dans le cadre une tâche de survie minimale inspirée de [Girard et al., 2003] que nous détaillerons plus tard dans le chapitre 4.1 : dans cette expérience, le robot a accès à 4 variables (inputs de la mRF), à partir desquelles il doit choisir une action (output de la mRF) parmi 5 à sa disposition, étant donné que le but pour le robot est de survivre le plus longtemps possible et que le temps de survie reflète la qualité du choix des actions.

Les résultats obtenus furent plutôt décevants : certes le modèle de la mRF obtient des résultats souvent meilleurs qu'un modèle purement aléatoire (c'est-à-dire dans lequel les décisions sont prises aléatoirement sans prendre en compte les variables en entrée), néanmoins le modèle de la mRF est bien inférieur à un simple modèle de type Winner-Takes-All (WTA) dont la décision prise correspond simplement à la plus élevée des variables en entrée.

1.5.2 Le modèle de Humphries - 2006

En 2006, Mark Humphries et ses collègues présentèrent leur propre modèle de la mRF dans [Humphries and Prescott, 2006], le second donc dans la littérature de la mRF. Ils adoptèrent le formalisme classique des réseaux de neurones et choisirent d'utiliser un modèle de population, où chaque neurone du modèle représente un ensemble de neurones réels de la mRF, comme le montre le schéma 1.7. Ils reprirent l'hypothèse de Kilmer-McCulloch qui suppose qu'à un cluster est associée une et une seule action.

Chaque cluster est composé de 2 neurones, modélisant chacun l'activité moyenne des neurones de chaque type de ce cluster :

- 1 neurone excitateur, projetant sur tous les autres neurones hormis ceux de son cluster ;
- 1 neurone inhibiteur, ne projetant que sur lui-même et sur le neurone excitateur de son cluster.

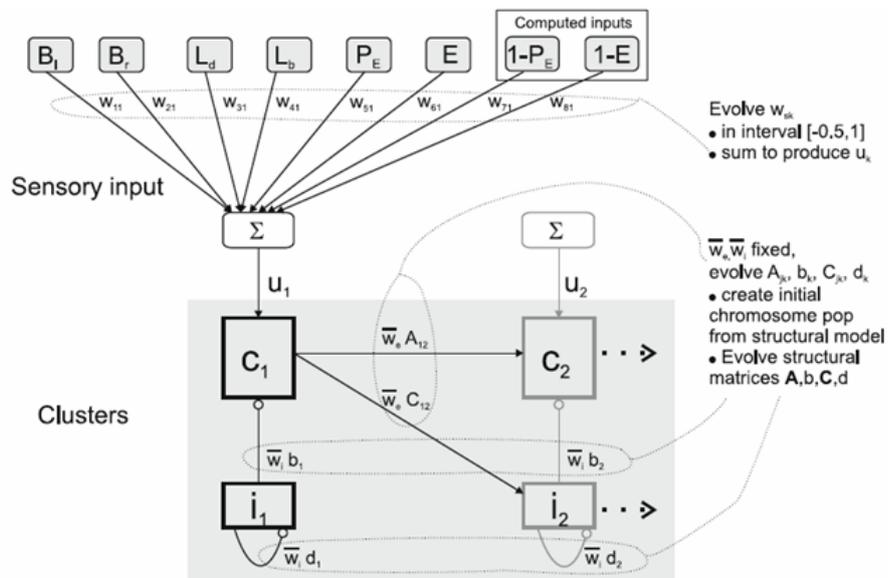


FIGURE 1.7: Modèle de Humphries. Ce schéma montre deux clusters, les entrées qu'il reçoit ainsi que les connexions des neurones du cluster de gauche. Chaque cluster contient un neurone inhibiteur noté i et un neurone excitateur noté c . Source : [Humphries and Prescott, 2006]

Les poids des connexions reflètent d'une part les probabilités des connexions issues des données anatomiques et d'autre part le poids des connexions réelles dans la mRF, inconnues dans la littérature.

Ce modèle a également été évalué par une tâche de survie minimale inspirée de

[Girard et al., 2003] et certains paramètres ont été optimisés avec des algorithmes évolutionnistes, dont nous parlerons plus en détail dans le chapitre 2.2. Les résultats s'avèrent également plutôt décevants en terme de durée de survie, la plupart des simulations de ce modèle ne furent que légèrement meilleures par rapport au modèle purement aléatoire.

1.6 Démarche et objectifs de ce rapport

Bien que leurs résultats n'eurent pas de très bonnes performances sur la tâche de survie en comparaison avec un contrôleur aléatoire, les deux modèles de la mRF ont eu le mérite de proposer des pistes de réflexion intéressantes ainsi que, dans le cas du second, des méthodes d'évaluation du modèle.

Notre démarche sera de construire un nouveau modèle de la mRF, en observant strictement le formalisme des réseaux de neurones et en respectant les données anatomiques, tout en diminuant le niveau d'abstraction en augmentant le nombre de neurones par cluster. Nous avons décidé de supprimer l'hypothèse posée par les concepteurs des deux précédents modèles qu'un cluster correspond à une et une seule action, car aucune donnée anatomique n'appuie cette hypothèse et au vu de la diversité des actions, il semble raisonnable de ne pas garder une telle contrainte.

Afin de ne pas avoir à définir tous les paramètres du réseau à la main et trouver les meilleurs réseaux, meilleur en terme de tâches de sélection de l'action et de respect des contraintes anatomiques, nous utiliserons des algorithmes évolutionnistes à l'instar de Humphries. À la différence de ce dernier qui ne fixait qu'un seul objectif dans ses évolutions, nous utiliserons un algorithme évolutionniste multi-objectifs, ce qui nous permettra de mieux optimiser et étudier nos réseaux.

Nous évaluerons notre modèle sur deux tâches de sélection de l'action : une tâche *désincarnée* montrant de façon statique la capacité d'un réseau à sélectionner des actions, et une tâche *incarnée* mettant le modèle de la mRF en situation réelle en s'inspirant de la tâche de survie minimale qu'utilise également Humphries pour ses évaluations.

Le chapitre suivant expliquera en détail notre démarche et les outils que nous avons utilisés pour mener à bien ce travail.

Chapitre 2

Matériel théorique

Pour commencer, nous allons expliquer comment la mRF dans notre modèle sera formalisée en un réseau de neurones. Puis, comme nous la ferons évoluer par algorithmes évolutionnistes, nous allons présenter le fonctionnement de ces derniers et voir ce qu'ils peuvent nous apporter. Les aspects techniques de l'implémentation de ces outils théoriques seront évoqués dans l'annexe A.

2.1 Les réseaux de neurones

Un réseau de neurones est composé d'un ensemble de neurones et d'un ensemble de connexions orientées liant certains neurones d'entre eux. Formellement, nous pouvons le considérer comme étant un graphe orienté et pondéré, chaque noeud correspondant à un neurone.

Il existe différents types de neurones : dans notre modèle, nous utiliserons une variante des neurones à taux de décharge de type intégrateurs à fuite, appelés IPDS (locally Projected Dynamical Systems) car ils permettent de modéliser une population de neurones. Nous avons choisi les IPDS en raison de leur propriété de stabilité intéressante, puisqu'il a été montré, par exemple, que la stabilité (au sens de la contraction) d'un système non linéaire composé de IPDS découle directement de la stabilité du même système sans IPDS, ce qui n'a pas été montré pour les intégrateurs à fuite standard [Girard et al., 2008]. Comme nous cherchons ici à construire un système permettant la sélection de l'action, la stabilité est préférable à l'instabilité.

Un neurone IPDS est caractérisé par 2 paramètres :

- τ , correspondant à la constante de temps ;
- *threshold*, correspondant au seuil d'activation.

Par simplicité, afin de ne pas multiplier les paramètres libres, nous avons fixé τ à 5ms et mis le *threshold* à 0. Nous avons également fixé le pas d'itération dt , qui doit être par construction toujours inférieur à τ , à 1ms.

La première opération réalisée par le neurone consiste en une somme des grandeurs reçues en entrées, pondérées par les coefficients synaptiques, c'est-à-dire la somme $w_1x_1 + \dots + w_mx_m = \sum_{j=1}^m w_jx_j$, où les x_i étant les entrées et w_i les coefficients synaptiques.

Nous devons ajouter le seuil *threshold* à cette formule : $\left(threshold + \sum_{j=1}^m w_jx_j \right)$

Nous utiliserons l'intégration des IPDS par la méthode approchée d'Euler. La fonction d'activation sera donc la suivante, a_n étant la valeur interne actuelle du neurone, a_{n+1} sa future valeur interne, égale à la valeur de sortie :

$$a_{n+1}(x) = \pm \max \left(1, \min \left(0, \left(a_n + (x - a_n) * \frac{dt}{\tau} \right) \right) \right)$$

Comme x correspond à la somme pondérée des grandeurs reçues en entrées, cela nous donne au final :

$$a_{n+1} = \pm \max \left(1, \min \left(0, \left(a_n + \left(threshold + \sum_{j=1}^m w_jx_j - a_n \right) \times \frac{dt}{\tau} \right) \right) \right)$$

Le \pm présent dans la formule traduit le fait qu'un neurone IPDS peut être soit excitateur, soit inhibiteur.

La figure 2.1 montre un exemple de neurone, et le graphe B.1 représente un réseau de neurones issu de notre modèle correspondant à une mRF avec 4 clusters.

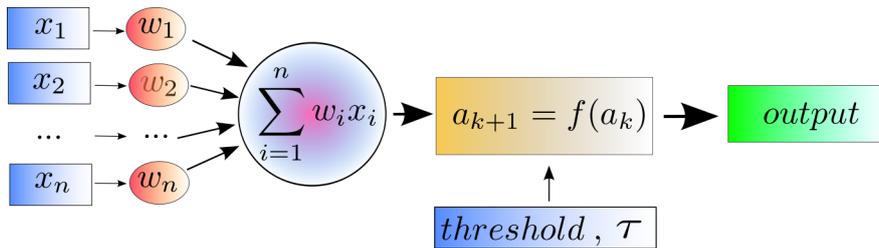


FIGURE 2.1: Exemple d'un neurone avec 2 entrées et une fonction d'activation à seuil.

Chaque cluster de la mRF a le même nombre d'entrées et reçoit des valeurs identiques.

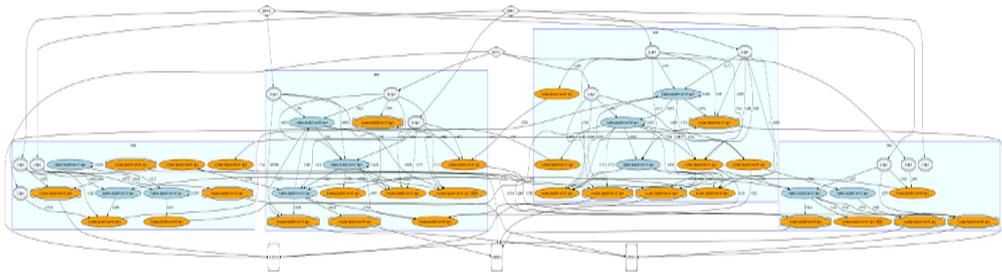


FIGURE 2.2: Exemple d'une mRF à 4 clusters. Les neurones oranges sont excitateurs, les neurones bleus foncé sont inhibiteurs. Un cluster correspond à un rectangle bleu. Les neurones situés en dehors des rectangles bleus représentent les entrées que reçoit la mRF ainsi que les neurones vers lesquels elle projette. Cette figure se trouve également dans l'annexe B en version agrandie.

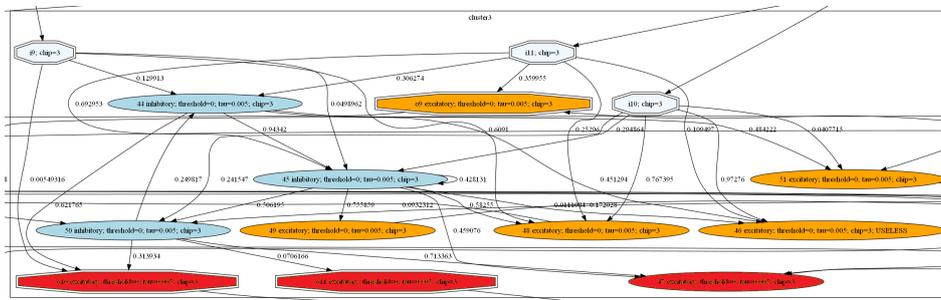


FIGURE 2.3: Exemple d'un cluster de la mRF. Les neurones oranges sont excitateurs, les neurones bleus foncé sont inhibiteurs. Chaque connexion synaptique à un poids entre 0 et 1. Les 3 neurones en bleu clair sont les entrées (neurones d'entrée), les 3 neurones en rouge sont les sorties de la mRF (neurones de sortie). Cette figure se trouve également dans l'annexe C en version agrandie.

De même, chaque cluster de la mRF a le même nombre de sorties. Au niveau global de la mRF, les valeurs de sorties correspondent à la moyenne des valeurs des sorties de chaque cluster. La figure C.1 montre un cluster. Un cluster a un nombre de neurones et de connexions variables.

Ces réseaux comportant une quantité importante de neurones, de connexions et de paramètres, il serait fastidieux de les optimiser à la main pour étudier en quelle mesure leur structure permet la sélection de l'action. Par conséquent, nous avons choisi d'utiliser les algorithmes évolutionnistes pour trouver des solutions par cette méthode d'optimisation qui a des propriétés particulièrement intéressantes pour notre problème comme nous allons le voir dans la section suivante.

2.2 Les algorithmes évolutionnistes

2.2.1 Définitions

Les algorithmes évolutionnistes, également appelés algorithmes évolutionnaires, sont une famille d'algorithmes d'optimisation s'inspirant du principe de **sélection naturelle** de la théorie darwinienne. Dans le cadre de la sélection naturelle, un environnement donné contient une population d'individus qui sont en concurrence pour la survie et la reproduction. L'aptitude de chaque individu à réaliser ces deux objectifs détermine leur chance d'avoir des enfants, autrement dit de transmettre leurs gènes à des individus de la génération suivante, lesquels auront pour des raisons génétiques une chance accrue de bien réussir à leur tour, voire mieux, ces deux objectifs.

Ce principe d'amélioration constante au cours des générations est repris par les algorithmes évolutionnistes pour optimiser des solutions à un problème. À la **génération** initiale, une **population** composée d'**individus** différents est générée, aléatoirement ou bien selon d'autres méthodes. Un individu correspond à une solution au problème, plus ou moins bonne : la qualité de l'individu par rapport au problème est appelée **fitness**, le terme anglais traduisant le degré d'adéquation de la solution par rapport au problème à résoudre. Plus la fitness d'un individu est élevée, plus ce dernier a des chances de transmettre une partie ou la totalité de son génotype dans des individus de la génération suivante.

Un individu est codé sous la forme d'un **génotype**, qui peut avoir n'importe quelle forme, telle une chaîne de caractères (algorithmes génétique) ou bien un vecteur de réels (stratégies d'évolution). Chaque génotype est transformé en un **phénotype** au moment de l'évaluation de l'individu, autrement dit lorsque sa fitness est calculée. Dans certains cas, le phénotype est identique au génotype : on parle alors de **codage direct**. Sinon, le codage est dit indirect. Par exemple, imaginons que l'on souhaite optimiser la taille d'un parallépipède rectangle défini par sa longueur, sa hauteur et sa largeur. Pour simplifier l'exemple, supposons que ces trois quantités soient des nombres entiers compris entre 0 et 15. On peut alors décrire chacune d'elles en utilisant un nombre binaire de 4 bits. Un exemple de solution potentielle peut avoir pour génotype 0001 0111 01010. Le phénotype correspondant serait un parallépipède de 1 de long, 7 de haut et 10 de large.

Dernière définition avant d'appliquer ces théories à notre modèle de la mRF, au moment du passage de l'ancienne à la nouvelle génération, sont appliquées des **opérateurs de variation** dont le but est de manipuler les individus. Il existe deux types d'opérateurs de variation distincts :

- les **opérateurs de mutation**, qui servent à introduire des variations au sein d'un même individu, à l'instar des mutations génétiques ;

- les **opérateurs de croisement**, qui servent à se faire croiser au moins deux génotypes différents, à l’instar des croisements génétiques issus de la reproduction.

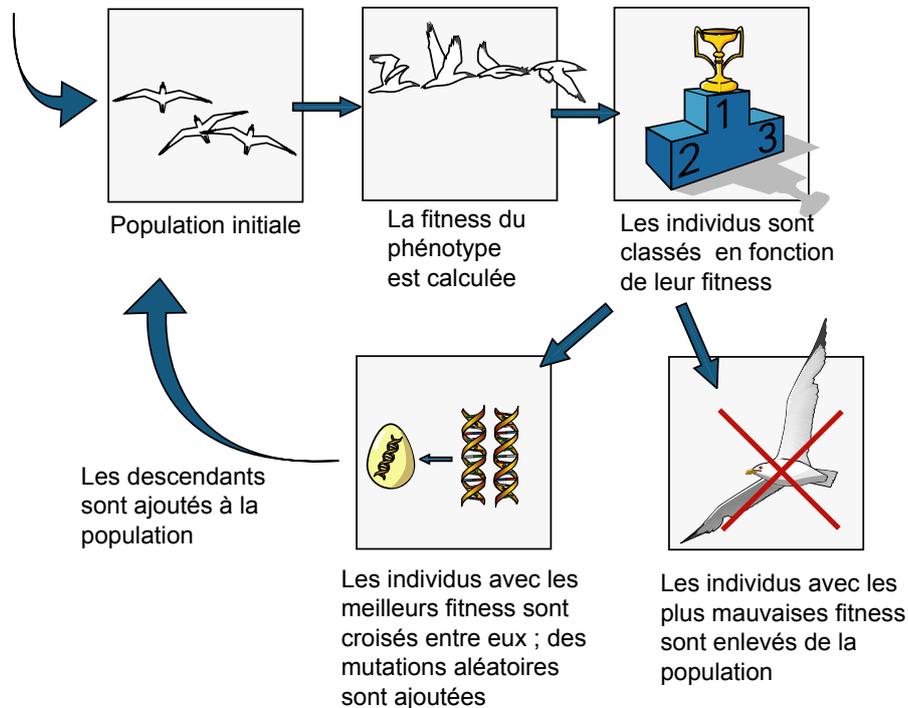


FIGURE 2.4: Fonctionnement d’un algorithme évolutionniste : à partir d’une population initiale de solutions, ces dernières sont classées selon leur fitness, les moins bonnes sont éliminées et les meilleures sont utilisées pour produire de nouvelles solutions. Source : [Doncieux et al., 2004]

Nous avons choisi les algorithmes évolutionnistes car ils ont fait leurs preuves dans des domaines divers tels la recherche opérationnelle, la robotique, la biologie, la finance ou encore la cryptographie. De plus, ils permettent d’optimiser plusieurs objectifs en parallèle et nous pouvons les utiliser comme des boîtes noires car ils ne présupposent aucune propriété mathématique sur le modèle à optimiser, permettant ainsi dans notre cas d’optimiser un système dynamique et non linéaire tel un modèle neuronal. Leur seule réelle limite est la complexité computationnelle, d’où la décision de coder notre programme dans un langage rapide (C++), multi-threadé, et de l’exécuter sur une grappe de serveurs. L’annexe A expose en détail les aspects techniques de l’implémentation.

2.2.2 Application

Dans notre modèle, la mRF est modélisée sous forme d'un réseau de neurones. Le génotype choisi lors de l'implémentation est un ensemble de réseaux de neurones correspondant chacun à un cluster de la mRF ainsi qu'un vecteur contenant l'ensemble des connexions entre les clusters, que nous appellerons **interconnexions**. Le phénotype est obtenu à partir du génotype en copiant chacun de ces réseaux dont un grand réseau, la mRF, sans oublier d'y rajouter les interconnexions.

Nos opérateurs de mutation sont :

- Ajout/suppression d'un neurone ;
- Ajout/suppression/modification d'une connexion (intra-réseau) ou d'une interconnexion (inter-réseau).

Nous aurions pu au cours des mutations modifier d'autres paramètres, par exemple certaines propriétés des neurones (e.g. inhibiteur/excitateur), néanmoins nous avons préféré limiter le degré de liberté de l'évolution. Également, nous n'avons pas choisi d'opérateurs de croisement : bien qu'intuitivement nous pourrions penser qu'il serait intéressant de croiser des mRF en leur permettant de mélanger leurs clusters, une telle opération est d'une part très délicate à implémenter car les interconnexions sont propres à chaque cluster et chaque mRF, et d'autre part de tels croisements ne seraient pas vraiment interprétables au niveau de l'évolution étant donné que le rôle de chacun des clusters n'est pas défini a priori.

Une partie très délicate fut l'implémentation des contraintes anatomiques de la mRF afin que l'évolution produise des réseaux de neurones cohérents avec les connaissances anatomiques. Nous l'avons implémenté à deux niveaux complémentaires :

- en amont, au niveau des opérateurs de mutation : à chaque mutation, nous veillons à rester aux alentours des données anatomiques ;
- en aval, au niveau du calcul de la fitness : nous avons utilisé un algorithme évolutionniste multi-objectif, ce qui nous permet de définir un objectif de plausibilité anatomique, poussant ainsi les réseaux à respecter les contraintes anatomiques.

La définition des objectifs impacte considérablement les résultats. Nous avons mis en place un objectif de plausibilité anatomique, en plus des objectifs propres tâches de sélection de l'action que nous détaillerons dans la section suivante.

Enfin, nous avons choisi d'utiliser l'algorithme **NSGA-II** [Deb, 2001, Deb et al., 2002], qui est à ce jour un des plus performants algorithmes évolutionnistes multi-objectifs et de loin le plus utilisé. Contrairement à un algorithme mono-objectif où il n'y a qu'un seul meilleur individu (avec possiblement des individus *ex æquo*), les meilleurs individus issus d'une évolution multi-objectif formeront un front appelé front de Pa-

reto, d'une dimension égale au nombre d'objectifs fixés. La figure 2.5 montre un front de Pareto de dimension 2 et la figure 2.6 compare l'ensemble de résultats obtenus par un algorithme mono-objectif par rapport à l'ensemble de résultats obtenus par un algorithme multi-objectif.

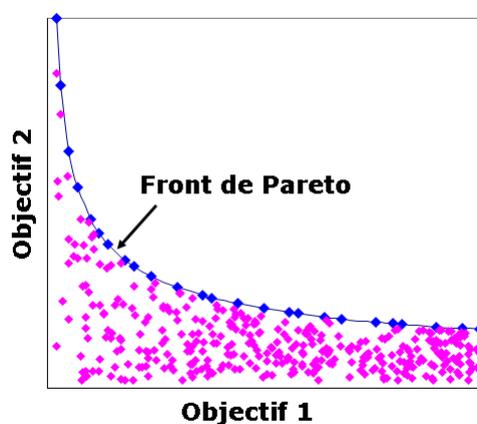


FIGURE 2.5: Exemple de front de Pareto de dimension 2 : sauf mention contraire, les algorithmes évolutionnistes maximisent les objectifs contrairement à la majorité des algorithmes d'optimisation dont le but est de les minimiser.

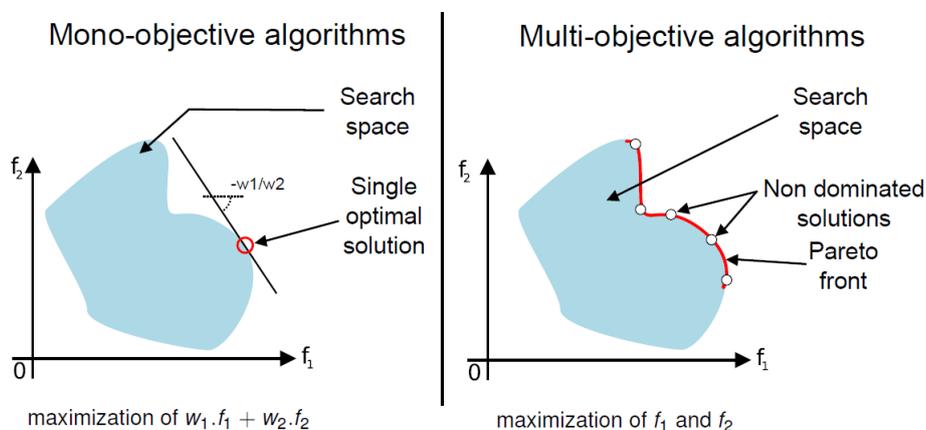


FIGURE 2.6: Mono vs multi-objectif. Un algorithme mono-objectif donnera 1 résultat, tandis qu'un algorithme multi-objectif donne un ensemble de résultat. Sur la figure de gauche, f_1 et f_2 correspondent respectivement aux scores obtenus pour l'objectif 1 et 2. w_1 et w_2 sont des poids affectés aux deux scores, la combinaison linéaire correspondant à un objectif unique afin de pouvoir utiliser un algorithme mono-objectif. Source : *Stéphane Doncieux et Jean-Baptiste Mouret.*

À présent que nous avons présenté d'une part la mRF et d'autre part les outils théoriques que nous avons utilisés pour le projet, tout en expliquant leur rôle dans la modélisation de la mRF, nous allons dans le chapitre suivant détailler les expériences de sélection de l'action réalisées et analyser les résultats obtenus.

Chapitre 3

Tâche désincarnée : la tâche vectorielle abstraite

Dans ce chapitre, nous allons présenter la première expérience pour évaluer notre modèle de la mRF. Dans un premier temps, nous détaillerons l'expérience réalisée, et dans un second temps nous analyserons les résultats.

3.1 Expérience

La tâche désincarnée correspond à une tâche abstraite de sélection. Le principe est simple : nous donnons en entrée de la mRF un vecteur de dimension 3, toutes les valeurs étant comprises entre 0 et 1. Ce vecteur d'entrée représente un ensemble de valeurs issues de diverses sources de données.

L'objectif de la mRF est de sélectionner une action. Pour cela, nous avons décidé en nous inspirant de [Humphries et al., 2007] que pour que la mRF choisisse la bonne action, son vecteur de sortie, également de dimension 3, doit avoir sa valeur maximum dans la même dimension que celle contenant la valeur maximum du vecteur d'entrée. Par exemple, si le vecteur d'entrée est (0.3, 0.5, 0.1), la mRF sélectionnera la bonne action si et seulement si son vecteur de sortie est de la forme (a, b, c), où $b > a$ et $b > c$, comme le montre la figure 3.1.

Afin que l'évolution ne sur-apprenne pas un ensemble restreint de vecteurs, ce qui empêcherait une bonne généralisation, la mRF sera évaluée sur sa capacité à sélectionner une action sur l'ensemble de vecteurs de dimension 3 où nous faisons varier chaque élément d'un entre 0 et 1 d'un pas de 0.1, et en supprimant les vecteurs ayant plusieurs composantes au maximum, afin d'éliminer les cas ambigus. Voici

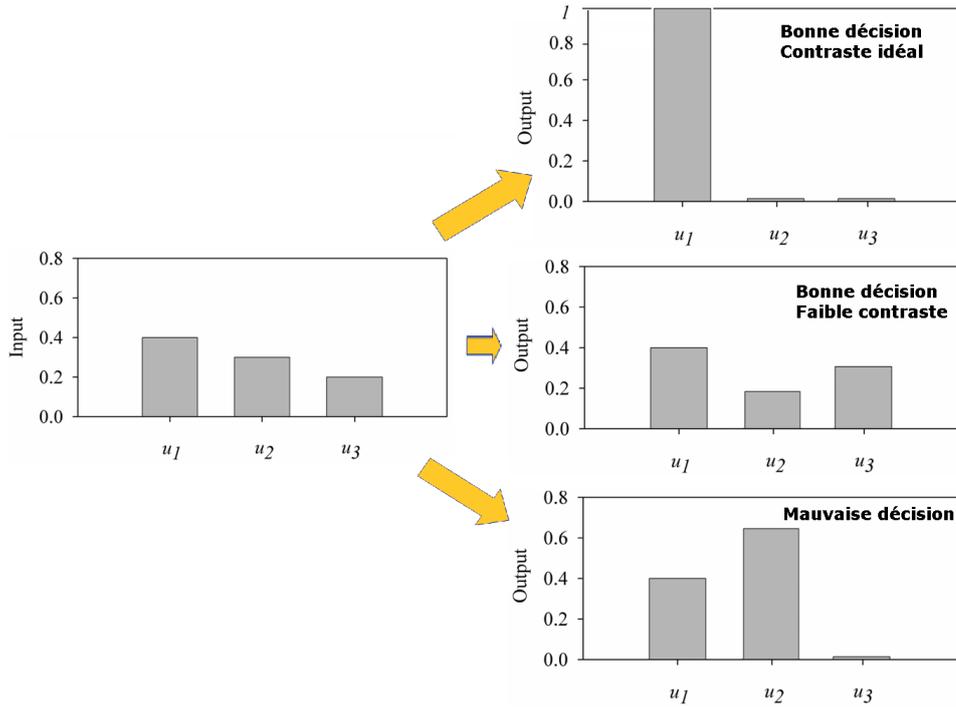


FIGURE 3.1: Ce schéma montre les deux premiers objectifs de la tâche désincarnée : le réseau reçoit en entrée un vecteur de dimensions 3 dont les valeurs sont entre 0 et 1, et il doit à la fois prendre la bonne décision (c'est-à-dire sélectionner la bonne dimension) et maximiser le contraste.

cet ensemble de vecteurs écrit formellement, ensemble qui contient en tout 1155 éléments :

$$\left\{ X = (x_1, x_2, x_3) \mid (x_1, x_2, x_3) \in S^3 \wedge (k = \arg \max_i x_i, \rightarrow ((i \neq k \wedge i \in \llbracket 1, 3 \rrbracket)) \rightarrow x_k > x_i) \right\}$$

où $S = \{0.1 \times i \mid i \in \llbracket 0, 10 \rrbracket\}$

Néanmoins, ce premier objectif admet des solutions inintéressantes, comme par exemple un réseau de neurones qui copieraient simplement les entrées en sortie. Afin d'opérer une réelle sélection de l'action, nous avons défini un second objectif forçant la mRF à maximiser le contraste entre la valeur maximum du vecteur de sortie et ses autres valeurs. La formule générale pour calculer le contraste est la suivante, en notant x_i le $i^{\text{ème}}$ élément du vecteur X et n sa dimension :

$$\text{contrast}(X) = \sqrt{\frac{\left(\sum_{i=1}^n (x_i - x_k)^2 \right)}{n-1}} \quad \text{où } k = \arg \max_i x_i.$$

Prenons par exemple le vecteur (0.6, 0.3, 0.5), le contraste est calculé ainsi :

$contrast = \sqrt{\frac{((0.6-0.5)^2 + (0.6-0.3)^2)}{2}} \approx 0.22$. Plus le contraste est élevé, plus l'action est clairement sélectionnée.

Le troisième et dernier objectif sera de respecter les contraintes anatomiques, pour que l'évolution produise des réseaux se rapprochant le plus possible d'une structure de réseau de type mRF. Le score de cet objectif de plausibilité anatomiques correspond à une combinaison linéaire de scores de sous-objectifs, ces derniers représentant chacun une des contraintes anatomique suivante (nous reprenons les notations du chapitre 1.3) :

- p doit être égal à 80%, la formule du score est $score_p = (-1) \times (0.8 - p)^2$;
- P(c) doit être égal à 25%, la formule du score est $score_{pc} = (-1) \times (0.25 - pc)^2$;
- les synapses des neurones de projection sont à 45% GABAergic, la formule du score est $score_{gabaergic} = (-1) \times (0.45 - percent_{gabaergic})^2$;
- les interneurons ne peuvent pas projeter en dehors de leur cluster, nous notons $nb_interneuron_not_within_chip$ le nombre d'interneurons ne respectant pas cette contrainte ;
- les neurones de projection ne peuvent pas projeter dans leur cluster, nous notons $nb_projection_within_chip$ le nombre de projection ne respectant pas cette contrainte.

Nous n'avons pas mis de contraintes sur P(p) et P(l) car les données anatomiques connues à leur sujet sont assez imprécises. Comme nous l'avons souligné dans le chapitre 2.2, les algorithmes évolutionnistes maximisent les objectifs, d'où les scores négatifs puisque se rapprocher des contraintes équivaut à un score se rapprochant de la valeur 0.

La formule globale pour calculer le score du troisième objectif est :

$$score_{anat} = 8 \times (score_p + score_{pc} + score_{gabaergic}) - 1 \times nb_interneuron_not_within_chip - 0.5 \times nb_projection_within_chip$$

Les pondérations 8, 1 et 0.5 ont été choisies arbitrairement, le seul critère étant que chacun des scores pondérés ait une valeur très approximativement du même ordre de grandeur, afin de maximiser les chances que l'évolution optimise chacun d'entre eux, et non juste de certains potentiellement aux dépens des autres.

En résumé, voici les trois objectifs que nous avons retenus :

- Objectif 1 : nombre de bonnes décisions. Minimum : 0, maximum : nombre de vecteurs testés ;
- Objectif 2 : importance du contraste. Minimum : 0, maximum : 1 (car toutes les valeurs des vecteurs sont comprises entre 0 et 1) ;
- Objectif 3 : plausibilité anatomique. Minimum : $-\infty$, maximum : 0.

En plus de ces trois objectifs situés en aval, nous avons également posé des contraintes anatomiques en amont, c'est-à-dire au coeur de la définition des opérateurs de mutation, afin de pousser l'évolution à respecter le plus possible les données anatomiques. Nous avons déjà mentionné dans le chapitre 2.2.2 cette double application des contraintes. Voici les contraintes que nous avons mises dans les opérateurs de mutation :

- impossibilité pour un neurone de projection de projeter dans leur cluster ;
- impossibilité pour un interneurone de projeter en dehors de leur cluster ;
- $P(c)$ doit être proche 25% ;
- p doit être proche 80%.

Comme nous le voyons, ces contraintes ne contiennent qu'une partie des données anatomiques connues : l'intérêt de les ajouter dans les opérateurs de mutation est que l'évolution génère moins d'individus manifestement non plausibles d'un point de vue anatomique. Mettre trop de contraintes dans les opérateurs des mutations serait néanmoins risqué car, outre son coût computationnel, restreindre excessivement l'espace de recherche peut rendre plus difficile l'obtention d'une bonne solution.

Nonobstant sa simplicité apparente, cette tâche désincarnée permet de valider ou d'invalider la possibilité pour une structure anatomiquement proche de la mRF d'effectuer des sélections. Nous allons analyser les résultats obtenus dans la section suivante.

3.2 Résultats

Tout d'abord, voici les paramètres que nous utiliserons tout au long des expériences, sauf indication contraire :

- taille de la population : 500 individus ;
- nombre de générations : 500 ;
- nombre de clusters : 4 ;
- nombre de neurones initial par cluster (en plus des neurones d'entrée et de sortie) : tirage aléatoire uniforme entre 3 et 10, les clusters pouvant avoir un nombre de neurones différents au sein de la même mRF ;
- probabilité d'ajouter/supprimer un neurone/une connexion (inter ou intra-cluster) : 0.05 ;
- probabilité de modifier le poids d'une connexion (inter ou intra-cluster) : 0.1 ;
- nombre d'itérations pour propager une entrée dans le réseau : 100.

Les évolutions calculées nous donnent des individus présentant au bout de 300 générations plus de 95% de bonnes décisions, voire dans certains cas 100%, respectant presque parfaitement les contraintes anatomiques et dont les vecteurs de sortie ont

un contraste supérieur à 0.6. Le graphique 3.2 montre un front de Pareto 2D obtenu typiquement lorsque l'on représente les scores des objectifs de bonnes décisions et de contraste obtenus par chaque individu. Le graphique 3.3 montre un front de Pareto 3D montrant les scores des 3 objectifs.

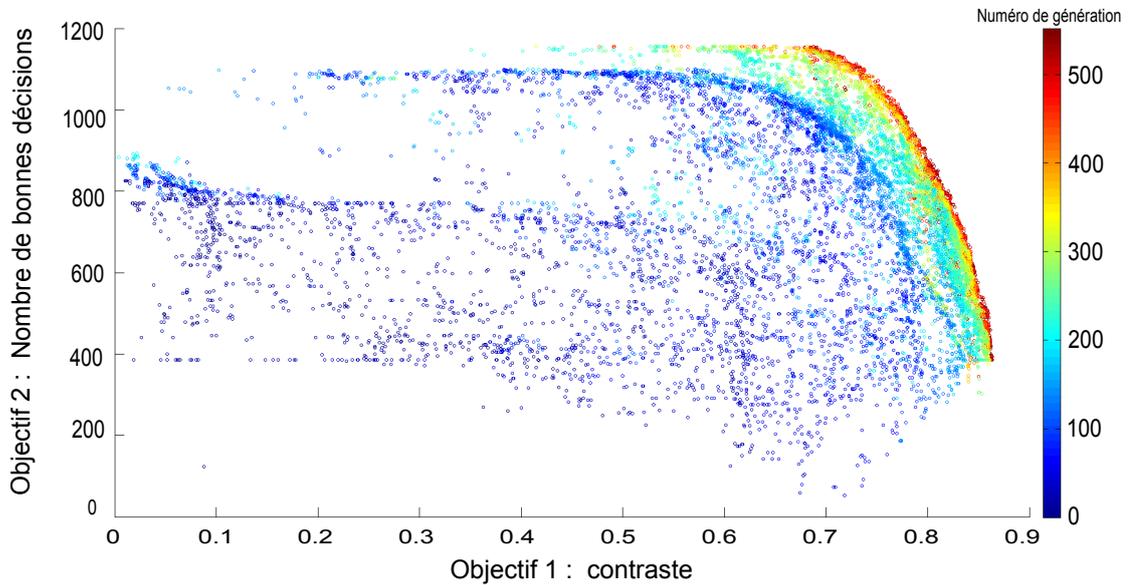


FIGURE 3.2: Front de Pareto de dimension 2 représentant les scores des objectifs de contraste en abscisse et de bonnes décisions en ordonnée obtenus par chaque individu de chaque génération au cours de l'évolution.

Prenons les 5 meilleurs individus présentant 1155 de bonnes décisions et regardons la moyenne de leurs résultats :

- score de nombre de bonnes décisions : 1155 (ce qui correspond au maximum théorique) ;
- score de contraste : 0.68761 (le maximum théorique est 1) ;
- score de plausibilité anatomique : $-1.20792e-15$ (le maximum théorique est 0).

Le tableau suivant résumé les statistiques anatomiques moyennes de ces 5 individus.

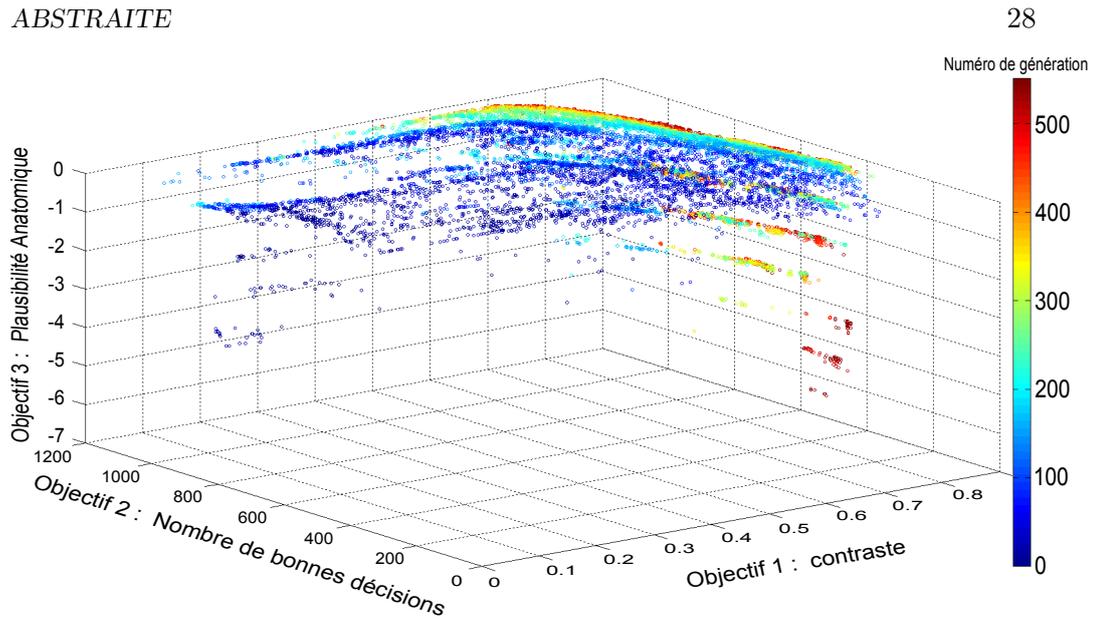


FIGURE 3.3: Front de Pareto de dimension 3 représentant les scores des objectifs de bonnes décisions, de contaste ainsi que de plausibilité anatomique obtenus par chaque individu de chaque génération au cours de l'évolution.

	Statistiques sur les 5 meilleurs individus	Contrainte dans opérateurs de mutation	Contrainte dans objectif	Données anatomiques connues
nombre total de neurones	57	Aucune	Aucune	NA
nombre total de connexions	117	Aucune	Aucune	NA
p	80%	80%	80%	80%
$percent_{gabaergic}$	45%	Aucune	45%	45%
$P(c)$	25%	25%	25%	25%
$P(p)$	8.73626%	Aucune	Aucune	< 10%
$P(l)$	8.61552%	Aucune	Aucune	< 10%
Nombre de interneurones projetant en dehors de leur cluster	0	0	0	0
Nombre de neurones de projection projetant dans leur cluster	0	0	0	0

Ce résultat montre qu'en respectant les données anatomiques connues de la mRF, cette dernière peut effectuer une tâche de sélection de l'action. Pour comparaison, [Humphries et al., 2007] dont cette tâche est inspirée obtient environ 75% de bonnes décisions, sans se pencher sur le contraste, et considère que cela suffit pour montrer la possibilité d'une sélection de l'action.

Nous avons également fait tourner le programme en enlevant les contraintes situées dans les opérateurs de mutation. Comme le montre le graphique 3.4, cela a pour effet d'augmenter le nombre d'individus ayant une faible plausibilité anatomique, car nous voyons que le front 2D représentant les scores des objectifs de bonnes décisions et de contaste obtenus par chaque individu est beaucoup moins nette que celui obtenu dans l'expérience initiale dont nous venons de présenter les résultats. Cette observation correspond à l'intuition que nous avons. Néanmoins l'évolution présente toujours des individus, certes plus rares que dans l'expérience initiale, presque parfaitement plausibles anatomiquement, de contraste supérieur à 0.5 et dont le taux de bonnes décisions est supérieur à 95%.

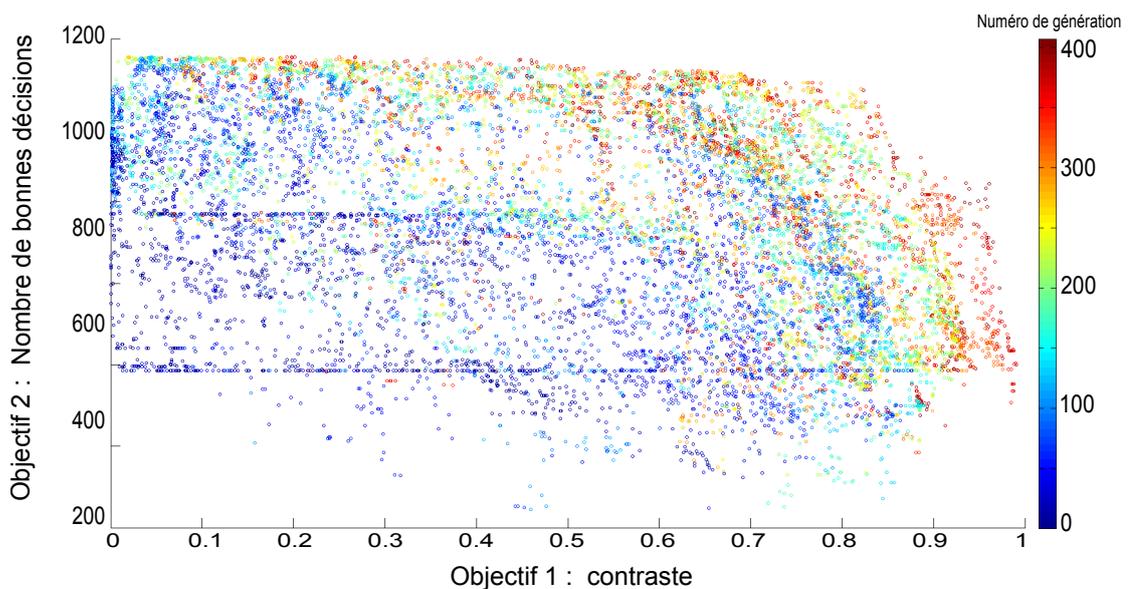


FIGURE 3.4: Front de Pareto 2D sans contraintes anatomiques lors des mutations. Le front représente les scores des objectifs de bonnes décisions et de contaste obtenus par chaque individu.

Une autre variante que nous avons essayée est d'enlever les contraintes des opérateurs des mutations et d'éliminer en plus l'objectif de plausibilité anatomique. Autrement dit, nous avons cherché à voir comment évoluerait un réseau sans contrainte anato-

mique vis-à-vis de la tâche de sélection.

Les résultats montrent que d'une part les réseaux obtenus ont des performances similaires à celles obtenues avec les réseaux contraints, voire légèrement meilleures (cf graphique 3.5) si nous considérons le contraste, et que d'autre part en faisant des analyses statistiques sur la structure de ces réseaux, ils ne tendent pas vers des structures de type mRF. Cela signifie que les données anatomiques connues sur la mRF ne représentent ni un avantage (car il existe d'autres structures de réseau tout aussi performantes) ni un inconvénient pour la sélection.

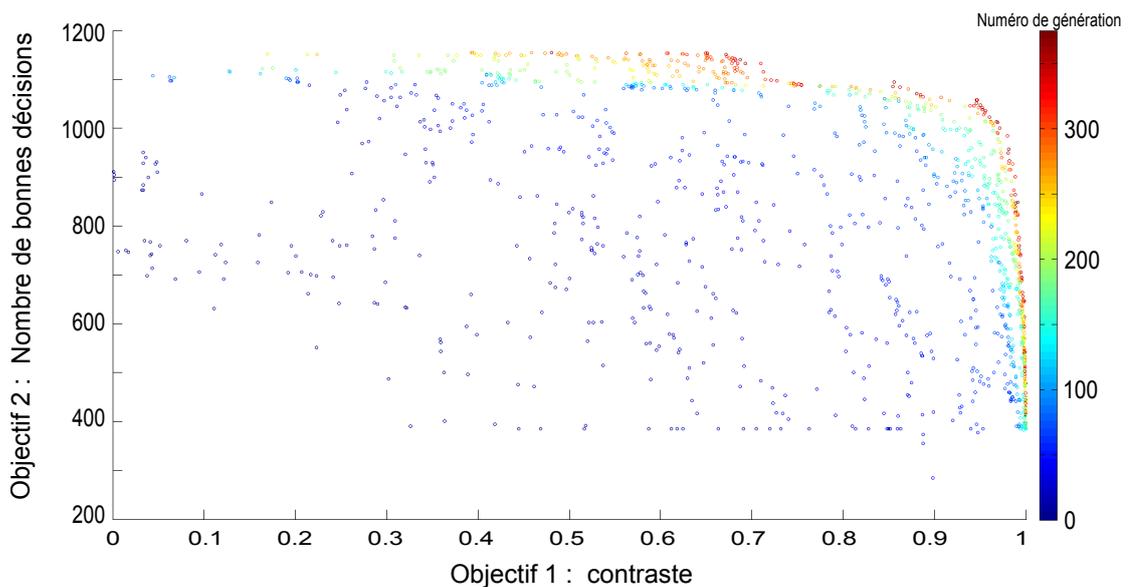


FIGURE 3.5: Front de Pareto 2D de l'évolution sans aucune contrainte anatomique (ni dans les mutations, ni dans un objectif). Le front représente les scores des objectifs de contraste en abscisse et de bonnes décisions en ordonnée obtenus par chaque individu de chaque génération au cours de l'évolution.

Cette première série d'expériences basées sur une tâche désincarnée a montré la capacité computationnelle de la mRF à effectuer une tâche de sélection. Notre modèle est plus performant que celui de Humphries dans cette tâche et nous y avons ajouté comme contrainte supplémentaire le contraste. Nous allons à présent effectuer une seconde série d'expériences pour analyser les performances de la mRF sur une tâche incarnée de simulation robotique.

Chapitre 4

Tâche incarnée : la tâche de survie minimale

À l'instar du chapitre précédent, nous allons dans ce chapitre tout d'abord présenter l'expérience dans un premier temps, et dans un second nous analyserons les résultats.

4.1 Expérience

La tâche de survie minimale que nous utiliserons ici est fortement inspirée de [Girard et al., 2003]. Il s'agit pour un robot de sélectionner efficacement ses comportements afin d'assurer sa survie, en maintenant ses variables d'état interne dans des intervalles tolérables, sa zone de viabilité [Ashby, 1952]. Cette survie dépend directement de la capacité du robot à se ravitailler auprès de deux types de ressources différents, en un temps limité par son niveau de recharge. L'utilisation de deux ressources différentes force le robot à se déplacer dans l'environnement pour accéder à l'une puis à l'autre et le met en situation de conflit pour déterminer quelle ressource est prioritaire à un instant donné, susceptible de générer des oscillations comportementales. Cette tâche sera simulée sur ordinateur.

Le robot sera placé dans un environnement où il pourra trouver deux types de ressources : des zones d'*ingestion* qui lui permettront de faire des réserves et des zones de *digestion* où il pourra *assimiler* ses réserves et les transformer en énergie utilisable. Sachant que tous les comportements du robot consomment de l'énergie, il va donc devoir alterner phases d'*ingestion* et de *digestion* pour survivre.

L'environnement expérimental est une surface plane de 400x400 unités entourée de murs. Elle est recouverte de 25 carreaux de 80x80 unités, de trois types différents : 21

carreaux gris (zones neutres), 2 carreaux noirs (zones d'ingestion), dont les ressources sont inépuisables, et 2 carreaux blancs (zones de digestion). La figure 4.1 montre l'environnement.

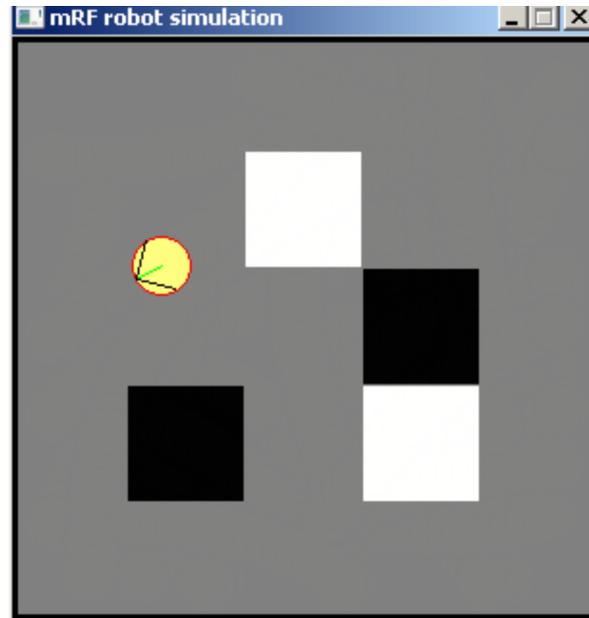


FIGURE 4.1: Environnement de la tâche de survie minimale. Nous y voyons 21 carreaux gris, 2 carreaux noirs et 2 carreaux blancs, tous les carreaux étant de la même taille. Le cercle jaune représente le robot.

Le robot est un disque de rayon 20 unités et possède 2 variables internes :

- l'**Energie Potentielle** (EP) correspond aux réserves puisées sur les zones d'ingestion, entre 0 et 1 ;
- l'**Energie** (E) est l'énergie réellement utilisable pour survivre dans l'environnement, elle est obtenue par digestion de l' EP sur les zones claires. Pour survivre, le robot doit maintenir E au dessus de 0 et le maximum de E est 1.

Le robot a accès à 4 capteurs externes :

- un **pare-choc gauche** (B_L pour left bumper), valeur binaire, à 1 si contacts avec le mur, à 0 sinon ;
- un **pare-choc droit** (B_R pour left bumper), valeur binaire, à 1 si contacts avec le mur, à 0 sinon ;
- un capteur de **carreau noir** (L_D pour light darkness), valeur binaire, à 1 si sur un carreau noir, à 0 sinon ;
- un capteur de **carreau blanc** (L_B pour light brightness), valeur binaire, à 1 si sur un carreau blanc, à 0 sinon.

Le robot a 5 actions à sa disposition :

- **Explorer aléatoirement** (*wander*) : le robot se déplace aléatoirement (rotation aléatoire entre 0 et 9° suivie d'un déplacement vers l'avant de 5 unités). À noter qu'en l'absence de capacités de navigation et de mémoire sur l'environnement, seul ce comportement permet de trouver des zones de recharges (carreaux blancs ou noirs). Cette action dure 2 unités de temps.
- **Éviter un obstacle** (*avoid*) : le robot effectue une marche arrière de 60 unités suivie d'une rotation de 180°. Cette action dure 2 unités de temps.
- **Recharger sur zone noire** (*reload on dark*) : le robot s'arrête et recharge son EP : $\delta EP = 0.027 \times L_D$. Nous remarquons que le robot ne recharge réellement son EP que s'il se trouve sur une zone noire. Cette action dure 1 unité de temps.
- **Recharger sur zone blanche** (*reload on light*) : le robot s'arrête et recharge son E : $\delta E = 0.027 \times L_B$ et $\delta EP = -0.027 \times L_B$. Nous remarquons que le robot ne recharge réellement son E que s'il se trouve sur une zone blanche. Cette action dure 1 unité de temps.
- **Se reposer** (*rest*) : le robot ne fait rien. Cette action dure 1 unité de temps.

À chaque unité de temps, le robot consomme 0.002 de son énergie, même lorsqu'il choisit de se reposer. Si l'énergie devient négative ou nulle, alors le robot meurt.

Afin de choisir une action, la mRF reçoit en entrée la **salience** de chacune des actions calculée à partir des variables internes et externes. La salience correspond au degré d'urgence ou de motivation à effectuer une action. Les formules utilisées ici pour le calcul des saliences sont les mêmes que dans l'évaluation du modèle de la mRF par [Humphries et al., 2005] ainsi que dans l'évaluation d'un modèle des ganglions de la base par [Girard et al., 2003].

- $S_{wander} = -B_L - B_R + 0.8(1 - P_E) + 0.9(1 - E)$
- $S_{avoid} = 3B_L + 3B_R$
- $S_{reload_on_dark} = -2L_B - B_L - B_R + 3L_D(1 - P_E)$
- $S_{reload_on_light} = -2L_D - B_L - B_R + 3L_B(1 - E)\sqrt{1 - (1 - P_E)^2}$

Si la mRF ne réussit pas à converger avec les saliences données en entrée, alors l'action de repos est sélectionnée. Dans notre expérience, un nouveau vecteur de salience est propagé pendant 100 itérations (1 itération durant 1ms) dans la mRF et nous considérons qu'il y a convergence dès lors que sur les 50 dernières itérations la variation de chacune des valeurs du vecteur de sortie de la mRF est inférieure à 0.001. Lorsque la mRF converge, alors nous considérons que l'action sélectionnée correspond à la sortie la plus élevée de la mRF. Nous testerons également une variante où l'action sélectionnée est modulée en fonction du contraste du vecteur de sortie de la mRF.

4.2 Résultats

Pour évaluer chaque contrôleur du robot, nous avons simulé 5 tâches de survie minimale dans la fonction fitness, en plaçant à chaque fois le robot à un endroit aléatoire sur la carte et initialisé avec comme valeurs 0.5 en énergie et 1 en énergie potentielle. Comme à chaque unité de temps le robot consomme 0.002 de son énergie, sa durée de vie minimale est de 500 unités de temps. Afin que l'évaluation prenne un temps raisonnable de calcul, nous avons limité les simulations à 3000 unités de temps chacune.

Tout d'abord, afin de vérifier une suffisante complexité de la tâche, nous avons évalué un contrôleur aléatoire, qui décidait au hasard une action parmi les cinq possibles. Les résultats de ce contrôleur montre une survie moyenne entre 500 et 600 unités de temps, ce qui confirme que la tâche ne peut être résolue par un contrôleur aléatoire et fournit une base de comparaison.

Une seconde vérification préalable que nous avons effectuée fut de tester un contrôleur de type Winner-Takes-All (WTA), ce dernier choisissant l'action en se basant sur la plus forte salience. Nos premiers résultats ont montré que ces contrôleurs avaient une durée de vie approchant souvent les 3000 unités de temps, ce qui signifiait que la tâche était trop simple pour évaluer notre modèle de la mRF correctement. Nous avons subséquemment essayé de trouver un facteur complexifiant la tâche et nous avons trouvé que la vitesse de déplacement du robot, non spécifiée dans [Humphries et al., 2005], influe grandement sur les résultats. Initialement, dans notre expérience, lorsque le contrôleur du robot choisit l'action à explorer aléatoirement, il se déplace en avant de 10 unités après avoir effectué une rotation aléatoire. Lorsque nous diminuons la vitesse déplacement à 5 unités, la tâche devient plus difficile et un contrôleur WTA n'a plus qu'une durée de vie moyenne d'environ 1250 unités de temps, ce qui est loin de la durée de vie maximum (3000 unités de temps). Par conséquent, la tâche semble non triviale à réaliser. La figure 4.2 compare les temps de survie des contrôleurs aléatoire et WTA sur 1000 tâches de survie.

Les résultats du contrôleur mRF montrent qu'en quelques générations seulement le robot réussit à vivre plus de 2500 unités de temps, voire le maximum 3000. Néanmoins, il faut plusieurs centaines de générations pour que le contraste devienne important. La figure 4.3 présente l'évolution du front de Pareto 2D d'un contrôleur mRF au bout de 500 générations : le temps de survie est proche ou égal à la valeur maximum, le contraste est d'environ la moitié du maximum théorique pour les meilleurs individus. Les scores de l'objectif de plausibilité anatomique montrent que la mRF respecte presque exactement les données connues sur la mRF. Concernant le temps passé sur chacune des actions en moyenne pendant une tâche de survie, la figure 4.5 montre qu'il n'y a pas de différence notable entre les différents contrôleurs, à l'exception

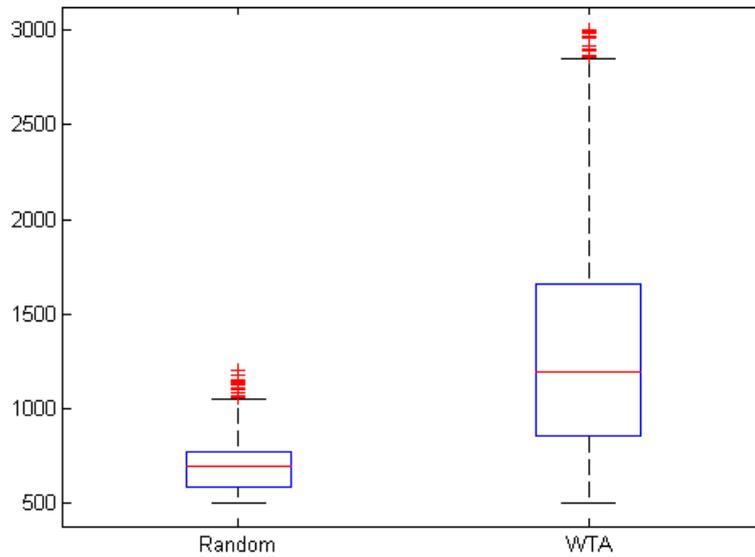


FIGURE 4.2: Comparaison des temps de survie des contrôleurs aléatoire et WTA sur 1000 tâches de survie.

triviale du contrôleur aléatoire.

Nous avons essayé de forcer la mRF à avoir un contraste élevé en modulant les actions en fonction du contraste. À cet effet, nous avons redéfini chacune des 5 actions en incluant la valeur du contraste, f étant la fonction de modulation du contraste :

- **Explorer aléatoirement** (*wander*) : le robot se déplace aléatoirement (rotation aléatoire entre 0 et $f(\text{contrast}) \times 9^\circ$ d'un déplacement vers l'avant de $f(\text{contrast}) \times 5$ unités).
- **Éviter un obstacle** (*avoid*) : le robot effectue une marche arrière de $f(\text{contrast}) \times 60$ unités suivie d'une rotation de $f(\text{contrast}) \times 180^\circ$. Cette action dure 2 unités de temps.
- **Recharger sur zone noire** (*reload on dark*) : $\delta EP = f(\text{contrast}) \times 0.027 \times L_D$.
- **Recharger sur zone blanche** (*reload on light*) : $\delta E = f(\text{contrast}) \times 0.027 \times L_B$ et $\delta EP = -f(\text{contrast}) \times 0.027 \times L_B$.
- **Se reposer** (*rest*) : le robot ne fait rien.

En prenant comme fonction de modulation $f(x) = \sqrt{x}$ et en évaluant toujours chaque réseau sur 5 tâches de survie, les réseaux mRF obtenus présentent des durées de survie similaires voire un peu inférieure aux contrôleurs WTA. Cependant, le contraste est un peu meilleur que lorsque nous introduisons une fonction de modulation, comme le

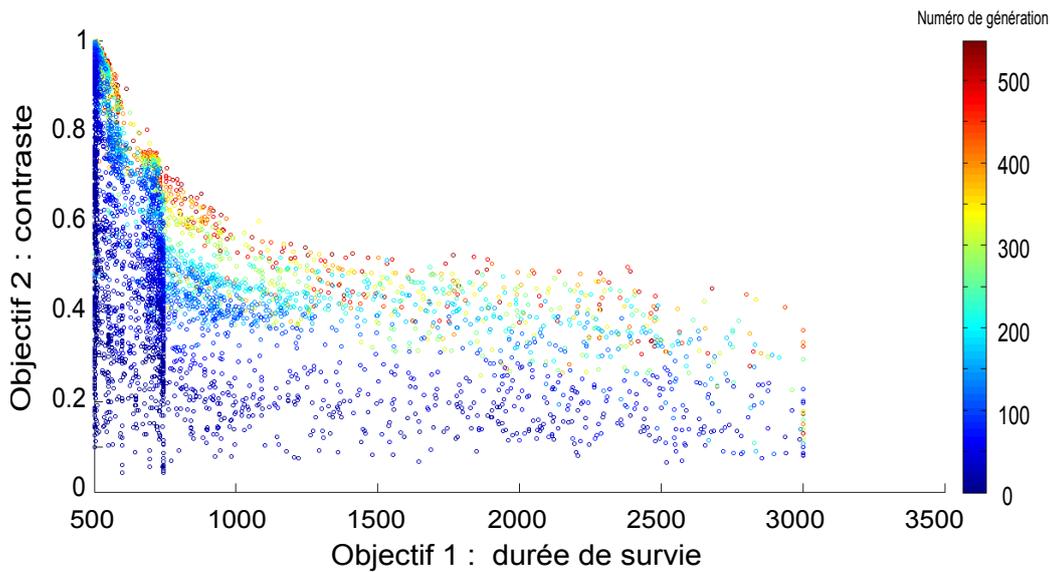


FIGURE 4.3: Front de Pareto 2D de l'évolution d'un contrôleur mRF représentant les scores des objectifs de durée de survie (abscisse) et de contraste (ordonnée) obtenus par chaque individu de chaque génération au cours de l'évolution. Le contrôleur mRF a été évalué sur 5 tâches de survie.

montre la figure 4.6. Par conséquent, la fonction de modulation introduit une pression de sélection favorisant le contraste aux dépens de la durée de survie. L'objectif de plausibilité anatomique a toujours un score entre -1 et 0, ce qui signifie que les réseaux ont bien une structure de type mRF.

Autre variante de l'expérience initiale, nous avons essayé de rendre la tâche plus réaliste en supprimant le calcul des saliences et donnant directement en entrée des réseaux les 4 variables externes (B_L , B_R , L_D et L_B) ainsi que les 2 variables internes (E et EP), ce qui a pour effet de complexifier la tâche. Nous avons également donné en entrée $1 - E$ et $1 - EP$ afin d'éviter que la tâche soit trop complexe, en reprenant ainsi la configuration de l'expérience de [Humphries and Prescott, 2006]. Les réseaux ont donc dans cette variante 8 entrées et 4 sorties.

Le graphique 4.7 montre les fronts de Pareto obtenus au bout de 1000 générations : nous voyons que certains réseaux ont une durée de vie dépassant 2000 unités de temps, néanmoins le contraste est très faible (inférieur à 0.1). Par comparaison avec les résultats obtenus avec le modèle de Humphries [Humphries and Prescott, 2006], ce dernier indique que la plupart des réseaux obtenus par évolution réussissent à faire à peine mieux qu'un contrôleur aléatoire, mais les actions sont modulées en fonction du contraste à l'instar de ce que nous avons fait dans l'expérience précédente. Ici,

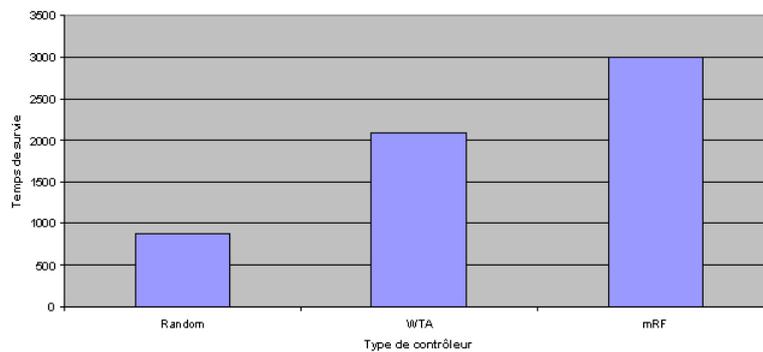


FIGURE 4.4: Comparaison des meilleures moyennes de temps de survie sur 5 tâches obtenues par des contrôleurs aléatoire, WTA et mRF. Pour chaque type de contrôleur, nous avons effectué 20 000 évaluations, chacune d'entre elles consistant à faire 5 tâches de survie et calculer la moyenne de la durée de survie. Nous n'avons gardé ici que la meilleure moyenne trouvée pour chaque type de contrôleur.

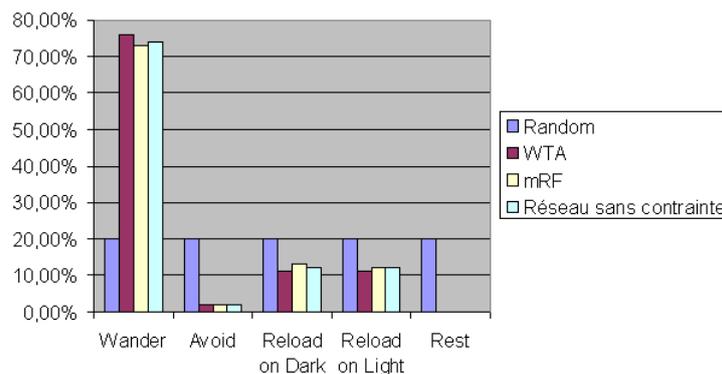


FIGURE 4.5: Comparaison du temps passé sur chacune des actions par les contrôleurs aléatoire, WTA et mRF. Ces statistiques agrègent les données 500 tâches de survie pour chacun des contrôleurs ayant les durées de survie les plus élevées parmi 50 000 tâches de survie simulées.

nos réseaux réussissent clairement mieux qu'un contrôleur aléatoire, cependant les actions ne sont pas modulées.

Ces résultats montrent que la mRF est globalement plus efficace qu'un simple réseau WTA et encore plus qu'un contrôleur aléatoire, comme le résume la figure 4.4, contrairement aux expériences de [Humphries et al., 2005] qui n'avaient pas réussi à évoluer la mRF de façon suffisamment optimale pour dépasser le WTA. Cela signifie que la mRF est non seulement apte à faire de la sélection de l'action, mais que celle-ci peut faire face à des situations complexes où un réseau WTA ne suffirait pas.

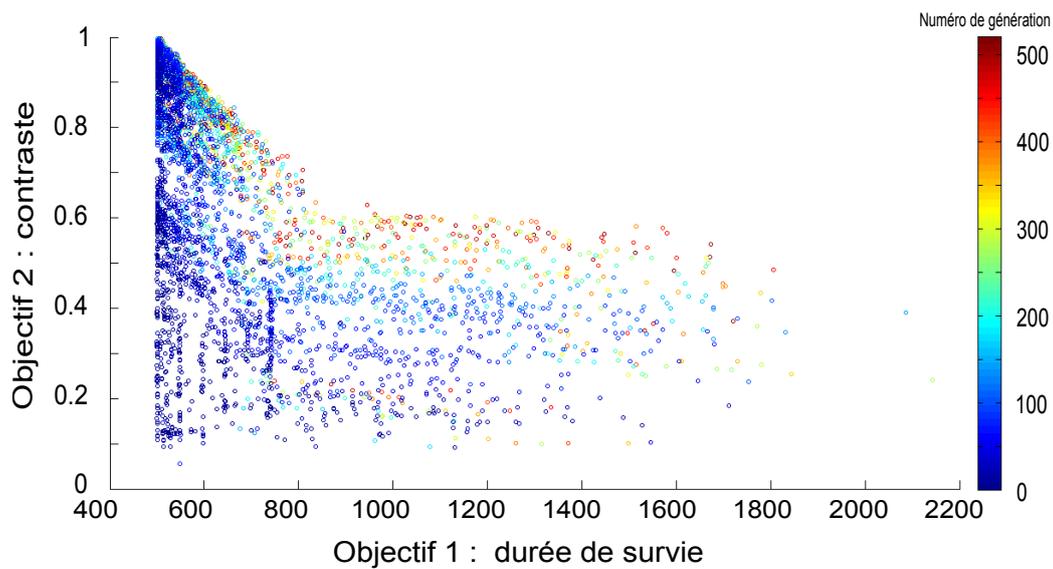


FIGURE 4.6: Front de Pareto 2D de l'évolution d'un contrôleur mRF représentant les scores des objectifs de durée de survie (abscisse) et de contraste (ordonnée) obtenus par chaque individu de chaque génération au cours de l'évolution, avec modulation des actions en fonction du contraste. La fonction de modulation est $f(x) = \sqrt{x}$ et le contrôleur mRF a été évalué sur 5 tâches de survie.

Néanmoins, dans les variantes où nous essayons de rendre la tâche de survie plus réaliste, en modulant les actions en fonction du contraste des vecteurs de sortie de la mRF ou en donnant à cette dernière directement les variables internes et externes sans calcul préalable de saliences, les résultats que nous avons obtenus sont moins probants : il faudrait approfondir davantage ces variantes. Autre point à creuser, il serait intéressant de quantifier la propension de nos réseaux mRF à généraliser leurs performances en les évaluant sur un nombre plus important de tâches de survie.

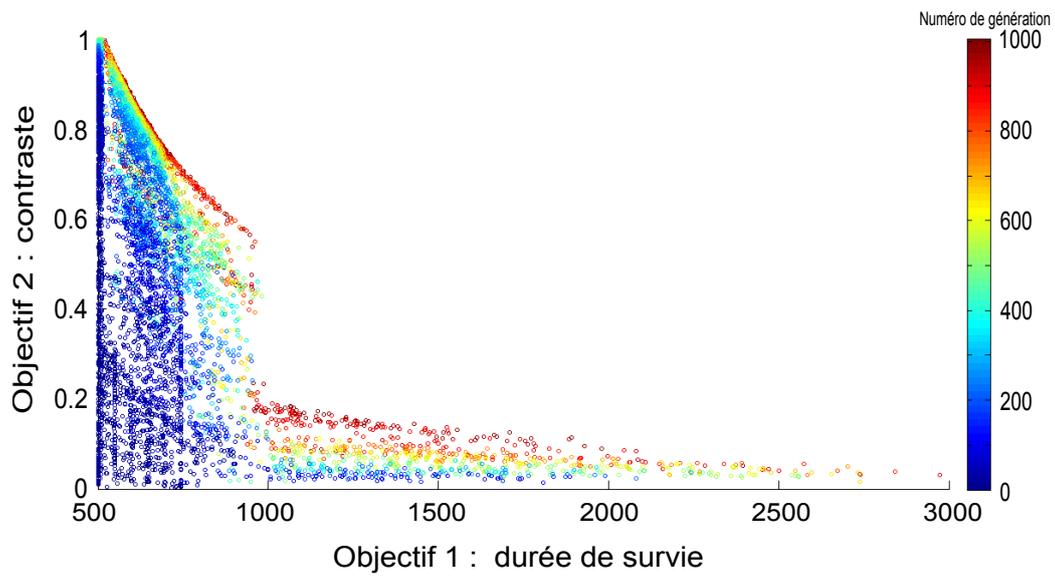


FIGURE 4.7: Front de Pareto 2D de l'évolution d'un contrôleur mRF représentant les scores des objectifs de durée de survie (abscisse) et de contraste (ordonnée) obtenus par chaque individu de chaque génération au cours de l'évolution. Le réseau reçoit directement en entrée toutes les variables externes et internes, et le contrôleur mRF a été évalué sur 5 tâches de survie.

Chapitre 5

Discussions et perspectives

La nature de ce travail de modélisation de la mRF était avant tout exploratoire. Seuls deux modèles existaient et le nombre d'articles portant sur cette zone du cerveau est assez faible. Nous avons donc forcément dû faire un certain nombre de choix et focaliser nos recherches sur certains points qui nous semblaient les plus essentiels pour initier un tel travail de modélisation. Cependant de nombreux axes d'études restent à être explorés, nous allons les évoquer brièvement dans ce chapitre.

Pour commencer, nous n'avons pas approfondi les paramètres $P(p)$ et $P(l)$ du réseau : nous avons vu dans la première partie de ce rapport que la littérature laissait supposer que ces valeurs sont inférieures à 0.1. Au cours de nos expériences, nous n'avons mis ni $P(p)$ ni $P(l)$ dans les contraintes anatomiques, néanmoins les structures des meilleurs individus que nous avons obtenus montrent qu'en moyenne les valeurs de $P(p)$ et $P(l)$ tournent toutes les deux autour de 0.09. Cela est en cohérence avec les données anatomiques connues de la mRF, cependant nous avons démontré à partir des données anatomiques connues que $P(l) > 45 \times P(p)$, relation que nous ne retrouvons pas dans nos résultats. Par conséquent, il serait intéressant de rajouter cette contrainte afin de voir les performances des mRF ainsi évoluées.

Nous avons également laissé de côté l'analyse et l'impact du modèle de population de neurones choisi, le IPDS, que nous avons décrit dans le chapitre 2.1 : en premier lieu, nous pourrions rendre libres les paramètres que nous avons fixés (τ à 5ms et *threshold* à 0). En second lieu, il existe d'autres modèles de population de neurones dont nous pourrions comparer l'efficacité dans notre modèle de la mRF par rapport aux IPDS. Enfin, les réseaux que nous avons fait évoluer sont de taille assez importante pour une modélisation, d'autant plus que nous les faisons évoluer par algorithmes évolutionnistes, avec lesquels il est préférable de manipuler des réseaux de

faible taille pour des raisons de vitesse de calcul ainsi que d'aisance d'interprétation : par conséquent, il serait intéressant de quantifier l'impact du nombre maximum de neurones que peut contenir un cluster.

Concernant la tâche incarnée, nous avons comparé indirectement nos résultats avec ceux de Humphries via la comparaison avec les performances de contrôleurs WTA et aléatoires. Il serait intéressant de comparer directement nos résultats avec ceux de Humphries en implémentant le modèle de ce dernier dans notre programme.

En outre, la tâche incarnée comporte un certain niveau d'abstraction : une évaluation encore plus poussée de la mRF serait par exemple de l'incorporer en tant que contrôleur d'une activité de métabolisme connue d'un animal, tels les comportements alimentaires et dipsiques de la souris [Guillot, 1988], et comparer les sélections opérées par la mRF avec le comportement réel de l'animal.

Autre voie d'exploration, il serait intéressant d'analyser les réseaux de neurones obtenus suite à l'évolution afin d'une part de mieux comprendre comment ils réussissent à résoudre une tâche de sélection, et d'autre part d'essayer d'extraire les similarités entre les réseaux obtenus. Dans ce même esprit, nous pourrions étudier les propriétés de ces réseaux, comme par exemple la présence de structures de type small-world, scale-free ou autre.

Du côté neuro-évolution, nous avons fait un usage intensif des algorithmes évolutionnistes sans évaluer l'utilité de ces derniers ainsi que leur impact sur les réseaux obtenus après évolution. Parmi les différents aspects qu'il serait utile d'approfondir, essayer d'introduire de nouveaux objectifs pourrait donner des résultats intéressants, en particulier :

1. un **objectif de généralisation**, permettant de s'assurer que les décisions prises par la mRF dans le cadre des tâches originales peuvent se généraliser à d'autres tâches semblables. Plus précisément, dans la tâche désincarnée, cela consisterait à s'assurer que si nous donnons en entrée de la mRF un vecteur non présent dans l'ensemble de vecteurs d'entrée que nous avons utilisé, par exemple (0.3, 0.5, 0.68), la mRF donne également la bonne sortie (0, 0, 1). Dans la tâche désincarnée, la généralisation pourrait se traduire par évaluer la mRF dans des cartes dont la disposition est différente que celle que nous avons utilisée [Pinville et al., 2011].
2. un **objectif de diversité**, grâce auquel l'évolution vérifierait si au sein de chaque population les individus (i.e. les mRF) ne sont pas tous semblables. Cet objectif permettrait de faire face au *genetic drift* (convergence vers un optimum local) souvent constaté, à l'instar de la sélection naturelle qui a vu se former des sous-groupes grâce aux contraintes géographiques. On

peut ainsi trouver autant d'optima locaux que de sous-groupes, permettant ainsi d'obtenir de meilleurs résultats en se rapprochant ainsi de l'optimum global [Mouret and Doncieux, 2009a, Mouret and Doncieux, 2009b, Doncieux and Mouret, 2009].

Une autre technique que nous pourrions utiliser pour essayer d'améliorer les résultats obtenus par algorithmes évolutionnistes serait d'effectuer l'évolution en plusieurs étapes. En effet, lorsque nous essayons de faire réaliser au réseau une tâche difficile, l'espace de recherche étant très important l'évolution peut rencontrer des difficultés à franchir certains seuils. Par exemple, nous avons vu que la variante de la tâche incarnée où nous nous donnions à la mRF directement les variables internes et externes sans calculer de salience donnait des réseaux dont le contraste des vecteurs de sortie était très faible, même après plus de 1000 générations. Une possibilité pour aider l'évolution à optimiser davantage le contraste aurait pu être dans un premier temps de faire apprendre le calcul des saliences au réseau, et dans un second temps résoudre la tâche de survie. Étant donné que nous avons montré dans la tâche désincarnée que des réseaux de type mRF pouvaient agir comme un WTA, et qu'un contrôleur WTA peut effectuer la tâche de survie à partir du calcul des saliences de façon moyennement efficace (environ 1500 unités de temps de survie sur un maximum de 3000) mais avec un contraste maximum, il est possible que cette introduction d'un objectif temporaire (apprentissage le calcul des saliences) dans l'évolution permette de franchir plus facilement des seuils dans l'optimisation des objectifs de durée de vie et de contraste.

Du côté neurophysiologique, la littérature donne un certain nombre de pistes pour élaborer des modèles plus fins de la mRF, en particulier pour mieux prendre en compte la diversité des noyaux et des neurones, diviser la mRF selon ses 3 régions (mésencéphalique, pontique, bulbaire) et considérer davantage les fonctions de la mRF. Pour cela, les différentes études de l'activité neurophysiologique de la formation réticulée (RF) représentent une source d'inspiration utile, en particulier la série d'études physiologiques sur la RF réalisées par JM. Siegel dans les années 1970 et 1980. Ainsi,

- [Siegel and McGinty, 1977] a trouvé que les neurones de la formation réticulée pontine (PRF) avaient un taux d'activité important en phase de forte activité oculaire constatée par électro-oculographie, ce qui est cohérent avec les études antérieures. Les auteurs montrent également l'existence d'un lien entre les décharges des neurones de la PRF et l'activité motrice. Les données trouvées suggèrent un rôle majeur des neurones de la PRF dans la régulation des activités motrices. L'étude donne également une indication très intéressante concernant les projections émanant de la PRF : *"The PRF's medial zone [...] is the principal source of pontine reticular projections to the spinal cord; more than half of its neurons send their axons*

directly into the ventral, motor areas of the cord."

- [Siegel, 1979] montre l'existence chez les chats de trois classes de cellules dans la RF bulbaire (*medullary reticular formation cells*), classification basée sur la corrélation entre leur décharge et le type de mouvements effectués : une classe décharge lorsque le mouvement est latéralement asymétrique (54% des cellules), l'autre décharge lorsque le mouvement est latéralement symétrique (38% des cellules). Les 8% de cellules restantes correspondent à la troisième classe et ne sont pas liés à un comportement moteur. Il n'existe pas de corrélation entre la RF bulbaire et le mouvement oculaire, contrairement à la RF pontine.
- [Siegel et al., 1979] signale la présence chez les chats d'une corrélation entre les décharges des neurones de la mRF au niveau moelle allongée (appelée *medial medullary reticular formation*, en français *mRF bulbaire*) et de certaines activités motrices durant les phases d'activité ainsi que pendant le sommeil paradoxal (REM sleep). L'étude remarque que les 3 types de cellules précédemment évoquées peuvent également être retrouvés dans le pont de Varole. Elle donne également une estimation de leurs proportions ainsi que leur localisation.

À plus long terme, un de nos objectifs est de connecter notre modèle de la mRF avec les modèles de ganglions de la base existants, connexion qui soulèvera à son tour un nombre important de problématiques : comment interagissent les réseaux, quel est le rôle précis de chacun, etc. Pour anecdote historique, l'article original présentant le premier modèle de la mRF s'achevait sur un dernier chapitre de considérations cybernétiques évoquant précisément ce genre de connexions entre zones du cerveau et en imaginant leurs équivalents dans un contrôleur robotique [Kilmer et al., 1969].

Chapitre 6

Conclusion

L'objectif de cette étude était de proposer un nouveau modèle de la mRF plus proche des données anatomiques que les deux modèles précédents existants dans la littérature et d'évaluer sa capacité à faire de la sélection de l'action. Notre démarche basée sur les algorithmes évolutionnistes nous a permis de montrer qu'un réseau présentant une structure de type mRF peut être capable d'une part de prendre les décisions attendues et d'autre part de les sélectionner de façon franche :

- la première série d'expériences basées sur une tâche désincarnée a montré la capacité computationnelle de la mRF à effectuer une tâche de sélection (chapitre 3) ;
- la seconde série d'expériences basées sur une tâche incarnée montre que la mRF est capable d'effectuer une tâche de sélection de l'action en condition simulée (chapitre 4).

Les résultats que nous avons obtenus sont meilleurs que ceux du modèle de Humphries dans les deux tâches que nous avons analysées : affiner le modèle en y ajoutant davantage de neurones et supprimer l'hypothèse du modèle de Kilmer-McCulloch reprise par modèle de Humphries qu'à chaque cluster correspond une action nous a ainsi permis d'améliorer les performances en termes de sélection tout en respectant davantage les données anatomiques connues sur la mRF.

Néanmoins, la structure de type mRF ne semble pas représenter un avantage particulier par rapport à un réseau de neurones sans contrainte. Ainsi, pour répondre à la question originale, la mRF peut constituer un substrat pour la sélection de l'action, mais il ne semble pas que sa structure représente un atout en particulier.

En outre, la démonstration de $P(l) > 45 \times P(p)$ dans le chapitre 1.4 nous permet avec une certitude accrue d'affirmer que la mRF a une structure de réseau de type

small-world, résultat pouvant s'avérer très utile dans les analyses futures.

Étant donné la nature exploratoire de ce travail de modélisation de la mRF, beaucoup d'axes de recherche que nous avons évoqués dans la discussion restent toutefois à être explorés pour affiner le modèle et approfondir les résultats. Cependant, les résultats de ce travail sont encourageants et outre leurs implications sur les capacités computationnelles de la mRF ils montrent l'apport potentiel des algorithmes évolutionnistes aux neurosciences computationnelles.

Bibliographie

- [Albert and Barabási, 2002] Albert, R. and Barabási, A. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1) :47–97. [cited at p. 8]
- [Ashby, 1952] Ashby, W. (1952). Design for a brain. [cited at p. 31]
- [Berntson and Micco, 1976] Berntson, G. and Micco, D. (1976). Organization of brainstem behavioral systems. *Brain Research Bulletin*, 1(5) :471–483. [cited at p. 6]
- [Berridge, 1989] Berridge, K. (1989). Progressive degradation of serial grooming chains by descending decerebration. *Behavioural brain research*, 33(3) :241–253. [cited at p. 6]
- [Birkmayer and Pilleri, 1966] Birkmayer, W. and Pilleri, G. (1966). The brainstem reticular formation and its significance for autonomic and affective behavior. [cited at p. 6]
- [Bowsher, 1970] Bowsher, D. (1970). Place and modality analysis in caudal reticular formation. *The Journal of Physiology*, 209(2) :473–486. [cited at p. 5]
- [Deb, 2001] Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. Wiley. [cited at p. 20]
- [Deb et al., 2002] Deb, K., Pratap, A., Agarwal, S., and Meyerivan, T. (2002). A fast and elitist multiobjective genetic algorithm : Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2) :182–197. [cited at p. 20]
- [Doncieux and Mouret, 2009] Doncieux, S. and Mouret, J. (2009). Single step evolution of robot controllers for sequential tasks. In *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, pages 1771–1772. ACM. [cited at p. 42]
- [Doncieux et al., 2004] Doncieux, S., Mouret, J., Muratet, L., and Meyer, J. (2004). The robur project : towards an autonomous flapping-wing animat. *Proceedings of the Journées MicroDrones*. [cited at p. 19]
- [Eccles et al., 1976] Eccles, J., Nicoll, R., Rantucci, T., Taborikova, H., and Willey, T. (1976). Topographic studies on medial reticular nucleus. *Journal of Neurophysiology*, 39(1) :109–118. [cited at p. 5]
- [Girard et al., 2003] Girard, B., Cuzin, V., Guillot, A., Gurney, K., and Prescott, T. (2003). A basal ganglia inspired model of action selection evaluated in a robotic survival task. *Journal of integrative neuroscience*, 2 :179–200. [cited at p. 1, 12, 14, 31, 33]

- [Girard et al., 2008] Girard, B., Tabareau, N., Pham, Q., Berthoz, A., and Slotine, J. (2008). Where neuroscience and dynamic system theory meet autonomous robotics : a contracting basal ganglia model for action selection. *Neural Networks*, 21(4) :628–641. [cited at p. 15]
- [Glickman and Schiff, 1967] Glickman, S. and Schiff, B. (1967). A biological theory of reinforcement. *Psychological Review*, 74(2) :81–109. [cited at p. 6]
- [Grantyn et al., 1987] Grantyn, A., Ong-Meang Jacques, V., and Berthoz, A. (1987). Reticulo-spinal neurons participating in the control of synergic eye and head movements during orienting in the cat. *Experimental Brain Research*, 66(2) :355–377. [cited at p. 8]
- [Guillot, 1988] Guillot, A. (1988). Contribution à l'étude des séquences comportementales de la souris : approches causale, descriptive et fonctionnelle. *Diplôme de Doctorat de l'Université Paris 7, Spécialité Biomathématiques*. [cited at p. 41]
- [Hammer Jr et al., 1981] Hammer Jr, R., Lindsay, R., and Scheibel, A. (1981). Development of the brain stem reticular core : an assessment of dendritic state and configuration in the perinatal rat. *Developmental Brain Research*, 1(2) :179–190. [cited at p. 7]
- [Humphries et al., 2005] Humphries, M., Gurney, K., and Prescott, T. (2005). Is there an integrative center in the vertebrate brain-stem? a robotic evaluation of a model of the reticular formation viewed as an action selection device. *Adaptive Behavior*, 13(2) :97–113. [cited at p. 1, 4, 12, 33, 34, 37]
- [Humphries et al., 2006] Humphries, M., Gurney, K., and Prescott, T. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proceedings of the Royal Society B : Biological Sciences*, 273(1585) :503–511. [cited at p. 1, 4, 7, 8, 10, 11]
- [Humphries et al., 2007] Humphries, M., Gurney, K., and Prescott, T. (2007). Is there a brainstem substrate for action selection? *Philosophical Transactions of the Royal Society B : Biological Sciences*, 362(1485) :1627–1639. [cited at p. 5, 6, 7, 8, 23, 29]
- [Humphries and Prescott, 2006] Humphries, M. and Prescott, T. (2006). Distributed action selection by a brainstem neural substrate : An embodied evaluation. *From Animals to Animals 9*, pages 199–210. [cited at p. 7, 13, 36]
- [Jones, 1995] Jones, B. (1995). Reticular formation : cytoarchitecture, transmitters, and projections. *The rat nervous system*, pages 155–171. [cited at p. 5]
- [Jouvet, 1967] Jouvet, M. (1967). Neurophysiology of the states of sleep. *Physiological Reviews*, 47(2) :117–177. [cited at p. 6, 10]
- [Kilmer et al., 1969] Kilmer, W., McCulloch, W., and Blum, J. (1969). A model of the vertebrate central command system. *International Journal of Man-Machine Studies*, 1(3) :279–309. [cited at p. 1, 7, 8, 10, 43]
- [Langhorst et al., 1983] Langhorst, P., Schulz, B., Schulz, G., Lambertz, M., and Krienke, B. (1983). Reticular formation of the lower brainstem. a common system for cardiorespiratory and somatomotor functions : discharge patterns of neighboring neurons influenced by cardiovascular and respiratory afferents. *Journal of the autonomic nervous system*, 9(2-3) :411–432. [cited at p. 5]
- [Lovick, 1972] Lovick, T. (1972). The behavioural repertoire of precollicular decerebrate rats. *The Journal of physiology*, 226(2) :4P–6P. [cited at p. 6]

- [Magoun and Rhines, 1946] Magoun, H. and Rhines, R. (1946). An inhibitory mechanism in the bulbar reticular formation. *Journal of neurophysiology*, 9(3) :165–171. [cited at p. 6]
- [Mouret and Doncieux, 2009a] Mouret, J. and Doncieux, S. (2009a). Overcoming the bootstrap problem in evolutionary robotics using behavioral diversity. In *Evolutionary Computation, 2009. CEC'09. IEEE Congress on*, pages 1161–1168. IEEE. [cited at p. 42]
- [Mouret and Doncieux, 2009b] Mouret, J. and Doncieux, S. (2009b). Using behavioral exploration objectives to solve deceptive problems in neuro-evolution. In *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, pages 627–634. ACM. [cited at p. 42]
- [Mouret and Doncieux, 2010] Mouret, J. and Doncieux, S. (2010). Sferesv2 : Evolving in the multi-core world. In *Evolutionary Computation (CEC), 2010 IEEE Congress on*, pages 1–8. IEEE. [cited at p. 52]
- [Parvizi and Damasio, 2003] Parvizi, J. and Damasio, A. (2003). Neuroanatomical correlates of brainstem coma. *Brain*, 126(7) :1524–1536. [cited at p. 6]
- [Pinville et al., 2011] Pinville, T., Koos, S., Mouret, J., and Doncieux, S. (2011). How to promote generalisation in evolutionary robotics : the progab approach. [cited at p. 41]
- [Prescott et al., 1999] Prescott, T., Redgrave, P., and Gurney, K. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behavior*, 7(1) :99–127. [cited at p. 7]
- [Ramón-Moliner and Nauta, 1966] Ramón-Moliner, E. and Nauta, W. (1966). The isodendritic core of the brain stem. *The Journal of Comparative Neurology*, 126(3) :311–335. [cited at p. 5]
- [Redgrave et al., 1999] Redgrave, P., Prescott, T., and Gurney, K. (1999). The basal ganglia : a vertebrate solution to the selection problem ? *Neuroscience*, 89 :1009–1024. [cited at p. 7]
- [Scheibel and Scheibel, 1967] Scheibel, M. and Scheibel, A. (1967). Anatomical basis of attention mechanisms in vertebrate brains. *The neurosciences : A study program*, pages 577–602. [cited at p. 7, 8]
- [Schuz, 1998] Schuz, A. (1998). Neuroanatomy in a computational perspective. In *The handbook of brain theory and neural networks*, pages 622–626. MIT Press. [cited at p. 8]
- [Segundo et al., 1967] Segundo, J., Takenaka, T., and Encabo, H. (1967). Somatic sensory properties of bulbar reticular neurons. *Journal of neurophysiology*, 30(5) :1221–1238. [cited at p. 5]
- [Siegel, 1979] Siegel, J. (1979). Behavioral relations of medullary reticular formation cells. *Experimental Neurology*, 65(3) :691–698. [cited at p. 43]
- [Siegel and McGinty, 1977] Siegel, J. and McGinty, D. (1977). Pontine reticular formation neurons : relationship of discharge to motor activity. *Science*, 196(4290) :678. [cited at p. 42]
- [Siegel et al., 1979] Siegel, J., Wheeler, R., and McGinty, D. (1979). Activity of medullary reticular formation neurons in the unrestrained cat during waking and sleep. *Brain Research*, 179(1) :49–60. [cited at p. 43]
- [Torvik and Brodal, 1957] Torvik, A. and Brodal, A. (1957). The origin of reticulospinal fibers in the cat. an experimental study. *The Anatomical Record*, 128(1) :113–137. [cited at p. 5]

[Woods, 1964] Woods, J. (1964). Behavior of chronic decerebrate rats. *Journal of Neurophysiology*. [cited at p. 6]

Appendices

Annexe A

Notes techniques sur l'implémentation

Nous tenons à évoquer les outils informatiques que nous avons utilisés pour mener à bien ce projet, car ils représentèrent d'une part un travail très important au cours du stage, et d'autre part il est toujours intéressant d'avoir une vision concrète des moyens techniques utilisés pour trouver les résultats du rapport, que ce soit par curiosité, en vue de reproduire les expériences ou bien encore pour trouver des pistes de solutions pour implémenter ses propres modèles.

Nous voulons également mettre l'accent sur le fait que les travaux réalisés au cours du stage n'ont nullement pour but unique la production de ce rapport mais aussi de donner la possibilité à de potentiels futurs successeurs d'être rapidement opérationnels sans qu'ils aient besoin de construire leurs propres outils informatiques à partir de rien. Nous rejoignons en cela l'initiative Plume du CNRS (<http://www.projet-plume.org>), dont le but est de Promouvoir les Logiciels Utiles, Maîtrisés et Économiques (d'où l'acronyme PLUME) à destination de la communauté de l'Enseignement Supérieur et de la Recherche.

À cet effet, une attention particulière a été portée sur la qualité du développement et tout le code est disponible sous la licence libre CeCILL (<http://www.cecill.info/>) à l'adresse TODO. Le tableau A.1 montre quelques statistiques sur le code source ainsi que sur les scripts écrits afin d'automatiser certains processus et analyser les résultats (contenus dans des fichiers de logs du programme principal).

Les réseaux de neurones ainsi que les algorithmes évolutionnistes ont été développés en C++ pour des raisons de rapidité, les algorithmes évolutionnistes deman-

Language	files	blank	comment	code
C/C++ Header	14	371	915	2571
C++	4	337	630	1940
Python	5	128	139	537
MATLAB	5	51	193	163
Bourne Shell	1	9	13	39
SUM:	29	896	1890	5250

FIGURE A.1: Statistiques sur le code source du programme et des scripts d'analyse écrits en plus du framework Sferes2 et des bibliothèques existantes. La taille élevée des C/C++ headers s'explique par l'utilisation intensive des templates.

dant des ressources computationnelles très importantes, en s'appuyant sur le framework Sferes2 (<http://pages.isir.upmc.fr/mouret/sferes2/>) développé à l'ISIR [Mouret and Doncieux, 2010]. Pour les besoins du développement, nous avons porté Sferes2 sous Windows. Le programme est multi-threadé et repose sur les bibliothèques Boost, notamment la Boost Graph Library qui permet de gérer aisément les graphes, TBB (Intel Threading Building Blocks) pour le multi-threading, Eigen2 pour les calculs matriciels et SDL pour le rendu graphique de la tâche de survie minimale que nous verrons plus tard.

Graphviz a été utilisé pour générer graphiquement les réseaux de neurones représentant les mRF, et le programme ffmpeg a été utilisé pour générer des vidéos à partir de ces graphes. MATLAB a été retenu pour analyser les résultats.

Enfin, nous avons fait un usage intensif de la grappe de serveurs du laboratoire ISIR pour exécuter notre programme, ce dernier nécessitant une forte puissance de calcul. L'ensemble des programmes utilisés dans ce projet tournent aussi bien sous Windows que sous Linux.

L'ensemble du code source ainsi que des scripts d'analyse est disponible à l'adresse http://pages.isir.upmc.fr/evorob_db/moin.wsgi/mRF2011.

Miroir : <http://bit.ly/mRF-xp>

Annexe B

Exemple d'une mRF à 4 clusters

Cf. page suivante.

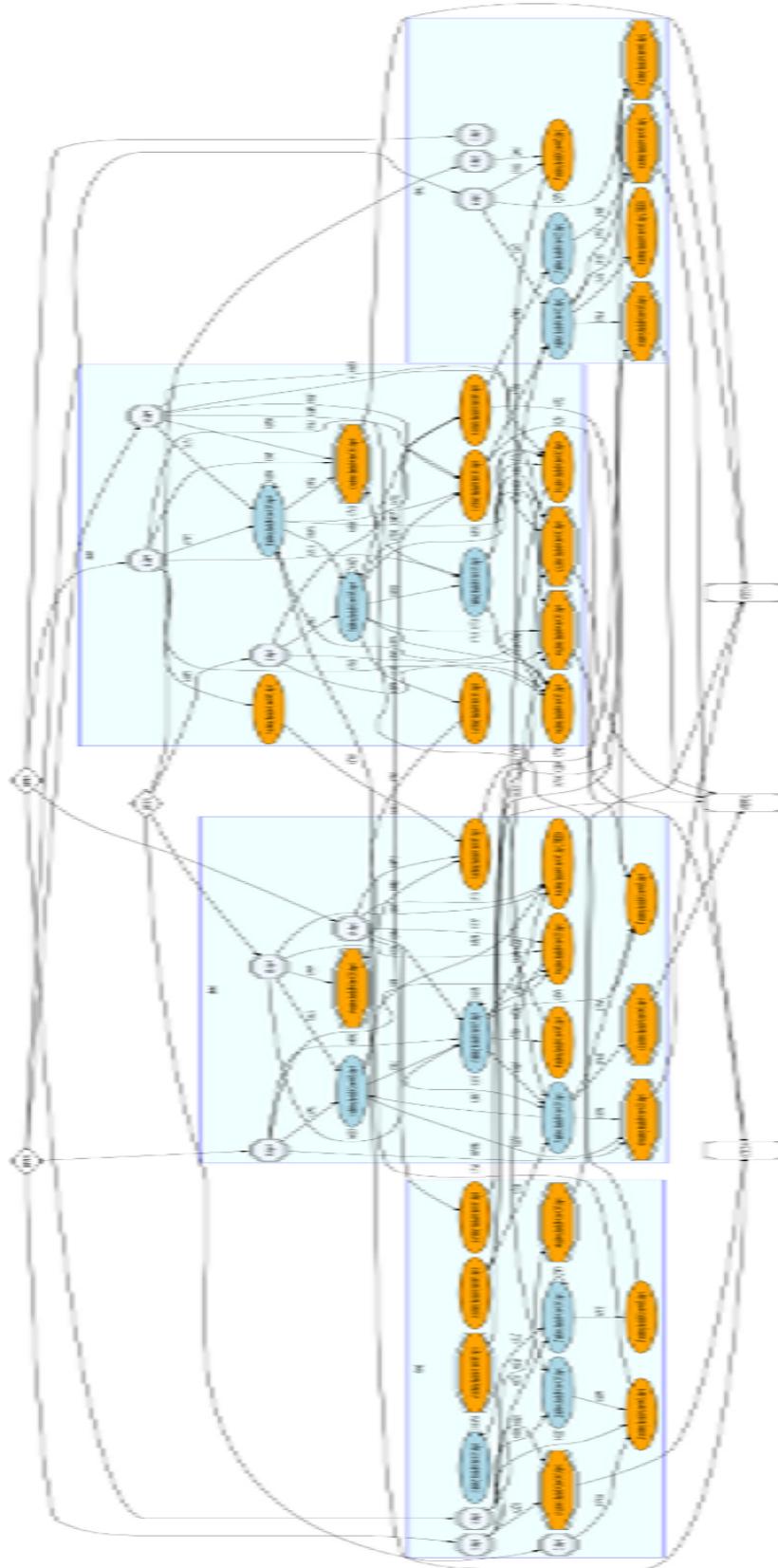


FIGURE B.1: Exemple d'une mRF à 4 clusters. Les neurones oranges sont excitateurs, les neurones bleus foncé sont inhibiteurs. Un cluster correspond à un rectangle bleu. Les neurones situés en dehors des rectangles bleus représentent les entrées que reçoit la mRF ainsi que les neurones vers lesquels elle projette.

Annexe C

Exemple d'un cluster de la mRF

Cf. page suivante.

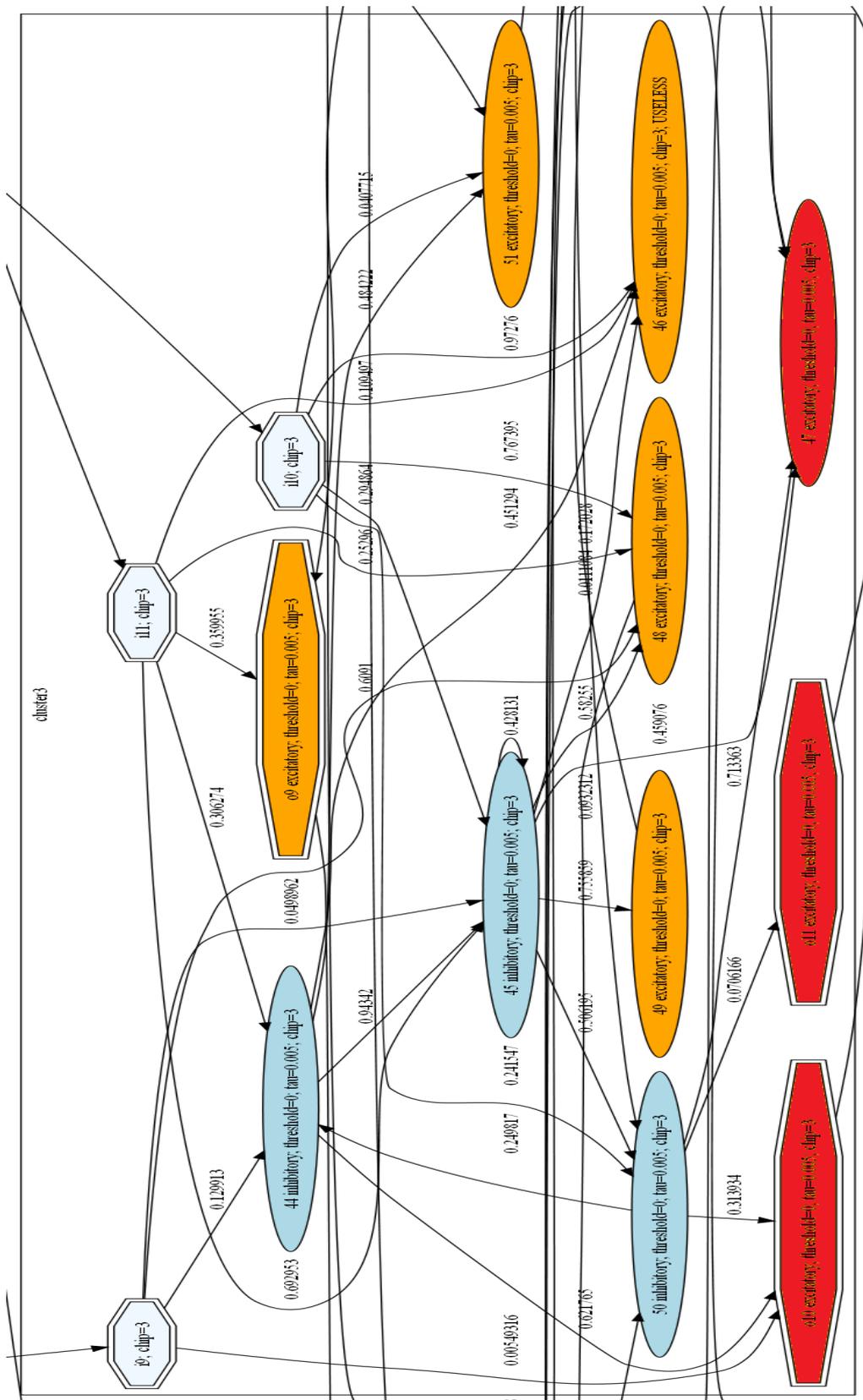


FIGURE C.1: Exemple d'un cluster de la mRF. Les neurones oranges sont excitateurs, les neurones bleus foncé sont inhibiteurs. Chaque connexion synaptique à un poids entre 0 et 1. Les 3 neurones en bleu clair sont les entrées (neurones d'entrée), les 3 neurones en rouge sont les sorties de la mRF (neurones de sortie).