



The medial Reticular Formation (mRF): a neural substrate for action selection? An evaluation via evolutionary computation.



Franck Deroncourt
franck.deroncourt@gmail.com

*Advisors: Stéphane Doncieux and Benoît Girard
Project funded by the ANR (ANR-09-EMER-005)*



Paris, June 2011

Contents

Contents	i
List of Figures	iii
1 Introduction	3
1.1 Preliminary remarks	3
1.2 Action selection	5
1.3 Anatomical data	7
1.4 Proof of $P(l) > 45 \times P(p)$	9
1.5 Previous models	10
1.5.1 The Kilmer-McCulloch model - 1969	10
1.5.2 Humphries model - 2006	12
1.6 Approach and objectives of this report	14
2 Theoretical material	15
2.1 Neural networks	15
2.2 Evolutionary algorithms	17
2.2.1 Definitions	17
2.2.2 Application	19
3 Disembodied task: the abstract vector task	22
3.1 Experiment	22
3.2 Results	25
4 Embodied task: the survival task	31
4.1 Experiment	31
4.2 Results	34
5 Discussions and perspectives	40
6 Conclusion	44
Bibliography	46

A	Technical notes on the implementation	51
B	Example of an mRF with 4 clusters	53
C	Example of a mRF cluster	55

List of Figures

1.1	Brainstem in the human brain (part in red)	4
1.2	Location of the mRF	4
1.3	Projections from and to the mRF	4
1.4	Decerebration lines rostral to the brainstem	6
1.5	Sagittal section of the mRF	7
1.6	mRF and small-world structures	11
1.7	Humphries model	13
2.1	Example of a neuron with 2 inputs	16
2.2	Example of an mRF with 4 clusters	17
2.3	Example of a cluster of the mRF	17
2.4	Functioning of an evolutionary algorithm	19
2.5	Example of a 2-dimensional Pareto front	21
2.6	Single vs multi-objective	21
3.1	Disembodied task	23
3.2	2D Pareto front	26
3.3	3D Pareto front	27
3.4	2D Pareto front without any constraints in the mutations	29
3.5	2D Pareto front of the evolution without any constraint	30
4.1	Setting of the survival task	32
4.2	Comparison of random controllers and WTA in the survival task	35
4.3	2D Pareto front for the evolution of the mRF controller	36
4.4	Comparison of the controllers in the survival task	37
4.5	Comparison of controllers in the survival task	38
4.6	2D Pareto front of the evolution of a mRF controller with modulation	38
4.7	2D Pareto front of the evolution of a mRF controller without salience	39
A.1	Statistics on the source code	52
B.1	Example of an mRF with 4 clusters	54
C.1	Example of a mRF cluster	56

Abstract

The medial Reticular Formation (mRF) is located in the brainstem: it receives many sensory inputs and it can control motor actions through its projections on the spinal cord and cranial nerves. The mRF is phylogenetically one of the oldest neural structures of the brainstem, the latter being regarded as one of the oldest centers of the central nervous system. Subsequently it seems to be a low-level system for action selection.

The first model of the mRF was proposed by [Kilmer et al., 1969], who already proposed that the mRF could be a "mode selector". [Humphries et al., 2005] tested the efficiency of this model in the minimal survival task defined in [Girard et al., 2003]. It performed poorly, but another version of it that included artificially evolved weights performed quite honorably. As a result, [Humphries et al., 2006] proposed a second model of the mRF, based on neural network formalism and taking into account new anatomical data. Nevertheless, it showed poor performances in the minimal survival task and turns out not to be anatomically very plausible.

In this Master's Thesis, we propose a new model of the mRF:

- constrained by anatomical information about its structure,
- constructed based on neural networks generated by artificial evolution,
- assessed on tasks of action selection.

The model we obtained successfully manages the selection tasks, indicating that the mRF can be used as an action selection system. We will also demonstrate an anatomical property of the mRF, which coupled with the results of the paper [Humphries et al., 2006] shows that it is very likely that the mRF network has a small-world structure.

This project was funded by the ANR (ANR-09-EMER-005-01) in the project EvoNeuro (<http://pages.isir.upmc.fr/EvoNeuro>).

Acknowledgements

*First and foremost, I would first like to express my deepest gratitude to my two supervisors, **Stéphane Doncieux** and **Benoît Girard**: in addition to their invaluable advice needed to carry out this project, they meticulously followed the progress of my work throughout the internship and their high standards pushed me to give the best of myself. Their complementary skills, Stéphane being a specialist in evolutionary computation and Benoît being an expert in neuro-computational models of action selection, turned out to be crucial in achieving the work that we present in this report and helped me drastically improve my skills on these two issues within a few months.*

*I also warmly thank **Jean Liénard**, **Tony Pinville** and **Paul Tonelli**, who are PhD students at ISIR lab: their generous help and spontaneous opinions guided my thinking toward more fruitful paths, in particular by sharing their PhD experiences.*

*Finally, my thoughts of thanks go to all the other CogMaster and ISIR people with whom I could discuss my internship project and helped me to mature this project and move forward in my research, especially my CogMaster/HEC advisor, **Philippe Mongin**, whose kindness, insight and openness to name a few of his qualities are endless. Many things would have not been possible without him.*

Chapter 1

Introduction

This chapter presents an overview of the median reticular formation: after a few preliminary remarks, we will summarize studies that suggest that the mRF is a proto-system of action selection. In a second step, we will present a synopsis of all the known anatomical data, which can be used to model the mRF. Lastly, we will explore the only two existing models of the mRF to see their shortcomings, while leveraging their strengths to build our own.

1.1 Preliminary remarks

The reticular formation, which includes the medial reticular formation, is a component of central nervous system which takes its name (from the Latin textsl reticulum meaning net) from its dense, intricate, anatomical presentation. Located on the floor of the brainstem between the medulla oblongata and the midbrain, it projects and receives nerve fibers throughout the spinal cord as well as many other components of the central nervous system such as the cerebral cortex, hypothalamus or cerebellum (see figures 1.1, 1.2 et 1.3).

We will focus in this report to a part of the reticular formation called median reticular formation, which we will shorten by mRF.

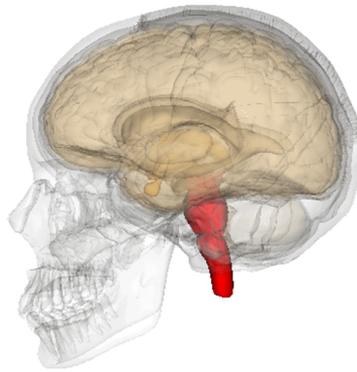


Figure 1.1: Brainstem in the human brain (part in red)

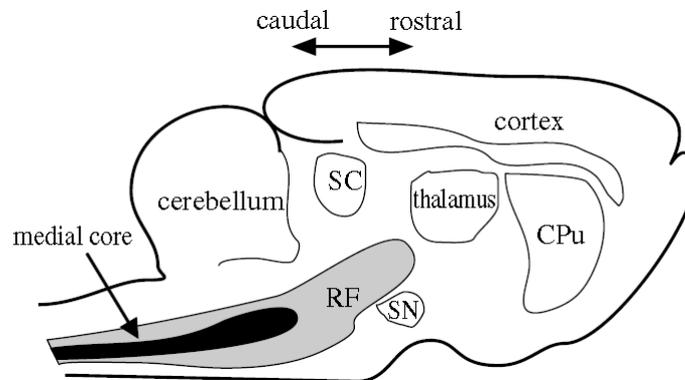


Figure 1.2: Location of the MRF (black area) in the brain of one cat. RF : reticular formation. CPu : caudate-putamen. SC : superior colliculus. SN : substantia nigra. Source : [Humphries et al., 2006]

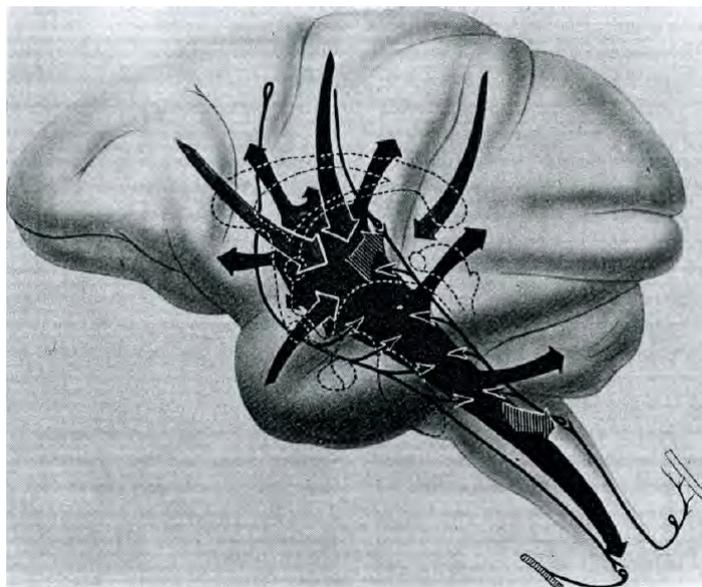


Figure 1.3: Projections from and to the mRF. Source : [Humphries et al., 2005]

The phylogeny of the nervous system shows that the mRF is one of the earliest neural structures of the brainstem, which in turn is what can be considered as one of the oldest centers of the central nervous system. Thus, among the different animal species, the mRF is very similar, as shown for example by [Ramón-Moliner and Nauta, 1966] between sharks and humans, allowing us to aggregate data directly from studies on different animals.

1.2 Action selection

Action selection can be defined by the crucial issue facing any autonomous agent, be it animal or robotic in nature, which is to continuously select and coordinate their behavior with a view to carry out its long-term goals, such that the survival, reproduction, or any other task defined by the designer in the case of a robot. The animals must necessarily implement effective solutions for selecting the action, hence our research within the nervous system of such a mechanism.

Several data suggest that the mRF is involved in the selection of the action, we will briefly summarize in this section.

First, the mRF seems to have all the information which monitoring systems have access to, and both external and internal senses of an animal: it thus receives a considerable amount of sensory input, such as synthesized by [Humphries et al., 2007], in particular from of the sensory, respiratory, visceral, vestibular, proprioceptive, nociceptive or cardiovascular systems. These data are corroborated by the various recordings made on the mRF showing that it reacts to very different stimuli [Segundo et al., 1967, Bowsher, 1970, Langhorst et al., 1983].

In a complementary manner, the mRF projects in mass on all levels of the spinal cord and cranial nerves [Torvik and Brodal, 1957, Eccles et al., 1976, Jones, 1995], which gives it the ability to control both the axial musculature that face. Therefore, the mRF has the inputs and outputs necessary for any candidate action selection system.

It has been experimentally shown that rats who had undergone a complete cut in the posterior brainstem, specifically posterior to thalamus and the hypothalamus, by removing the entire brain rostral to this cross-section (see the 3 decerebration lines in Figure ??), had a surprisingly coherent behavior [Woods, 1964], except for errors caused by the loss of sight, smell and problems of regulation hormone. The rats of the experiment were still able to make selections of low-level action, such as eat, move, drink or sleep according to the stimuli. The very important results of this experiment were later confirmed by [Lovick, 1972, Berntson and Micco, 1976, Berridge, 1989],

which emphasized the fact that rats can perform combinations of coordinated actions, such as holding, biting and chewing food, combinations more complex than simple reflexes which may emanate from the spinal cord.

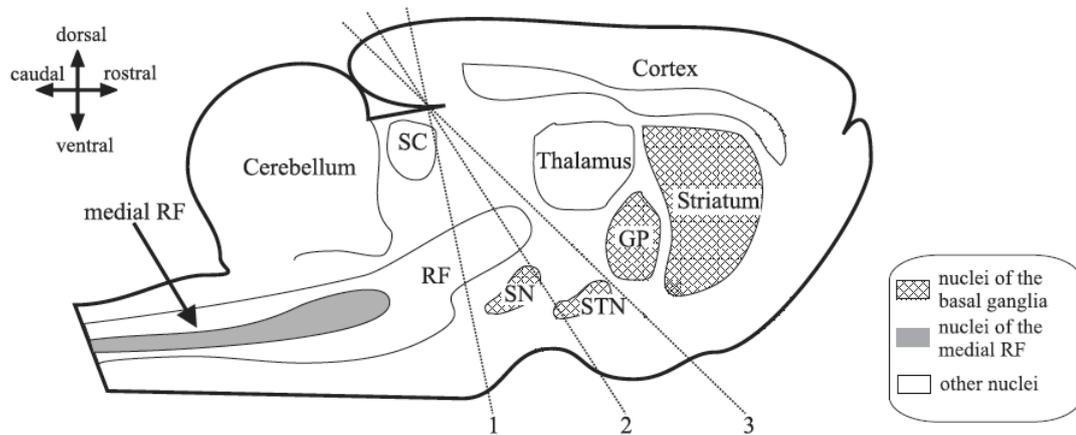


Figure 1.4: Decerebration lines rostral to the brainstem. The dashed lines show the location of the three most-common decerebration lines all the brain rostral to the line is removed, leaving hindbrain and spinal cord intact. GP : globus pallidus. RF : reticular formation. SN : substantia nigra. STN : subthalamic nucleus. SC : superior colliculus. Source : [Humphries et al., 2007]

To verify the role of the mRF, local electrical and chemical stimulation were made on normal animals. These stimuli made them change their behavior, such as eating, sleeping, drinking, escape, or seek to wash [Magoun and Rhines, 1946, Glickman and Schiff, 1967]. These results show the magnitude of the different behaviors that control at least part of the MRF.

Conversely, other studies have focused on the impact of injuries to the mRF. They demonstrate severe behavioral disorders, including sleep disorders, the study subjects showed a phase shift of sleep between the brain and the rest of the body [Birkmayer and Pilleri, 1966] as well as frequent alternations between deep sleep and frequent extreme rage [Jouvet, 1967]. [Parvizi and Damasio, 2003] have even shown that lesions in certain parts of the mRF can cause coma and even death in humans.

Lastly, unlike most neural structures, mRF cells exist at birth [Hammer Jr et al., 1981], which may hint of the importance of their presence for the survival of the individual.

In light of these studies, it seems reasonable to conjecture that the mRF is a proto-

system of selection of the action.

1.3 Anatomical data

In this section, we will gather all the anatomical data that will be useful to construct a model of the mRF.

First general remark, the neuroscience literature shows that the action selection mechanisms are implemented in two different architectures in the mammalian brain:

- a centralized architecture style, where each neural module projects onto a central decision-making. The basal ganglia is supposed to have this type of architecture [Redgrave et al., 1999, Prescott et al., 1999].
- a modular architecture, where each neural module is competing, inhibiting or exciting the other modules. The mRF would be supposed to have this type of architecture [Humphries et al., 2007].

One feature of the mRF is its cluster organization (also known as chips or stacks) [Scheibel and Scheibel, 1967], as shown by the sagittal section of the mRF shown in Figure 1.5.

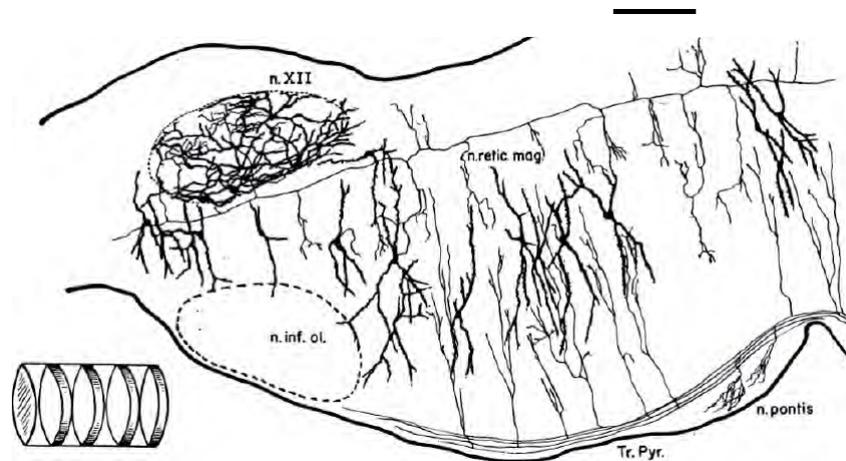


Figure 1.5: Sagittal section of the mRF of a rat, showing its cluster organization. Source : [Scheibel and Scheibel, 1967]

The mRF of a rat measures approximately 7 mm on the rostral-caudal axis and contains between 35 and 75 clusters [Humphries et al., 2006]. The mRF of a frog contains circa 0.75 million neurons, whereas the mRF of a human being contains 2

million [Kilmer et al., 1969]. Clusters have the same size in terms of size and number of neurons [Humphries et al., 2007]. They all receive the same inputs and project to the same areas [Humphries and Prescott, 2006].

There are two main types of neurons in the mRF:

- The **interneurons**: of small or medium size, they are **inhibitory** (some exceptions exist though) and they project almost exclusively within their own cluster, both to projection neurons and to interneurons. They represent about 20% of neurons in the mRF [Humphries et al., 2006].
- The **projection neurons**: of medium to high size, they are **excitatory** (some exceptions exist though) and they project almost exclusively outside their own cluster to other clusters or outside the mRF [Humphries et al., 2006]. Approximately 45% of the input synapses of projection neurons are GABAergic, that is to say inhibitory, i.e. interneurons [Humphries et al., 2007]. Projection neurons represent about 80% of neurons in the mRF.

The probability $P(c)$ that a projection neuron projects on a given cluster is the subject of two different models in the literature. According to [Grantyn et al., 1987], $P(c) = 0.25$ regardless of the source cluster (i.e. where the projection neuron's nucleus is located) and the destination cluster (to which the projection neuron projects). According to [Kilmer et al., 1969], the probability $P(c)$ depends on the distance between the cluster source and the destination cluster: $P(c) = d^{-a}$, where d is the distance and a a parameter greater than or equal to 1.

The probability $P(p)$ that a projection neuron projects on a given neuron, given that projects to the cluster of the latter, and the probability $P(l)$ that an interneuron forms a connection with a given neuron's cluster are both low, probably less than 0.1 according to [Schuz, 1998, Albert and Barabási, 2002].

To sum up, given the anatomical constraints, six parameters completely describe the structure of the network:

- c : the number of clusters (between 35 and 75) ;
- n : the number of neurons in a cluster (about $1500000/55 \approx 30000$) ;
- p : the percentage of projection neurons (about 80%). The percentage of interneurons is therefore $1 - p$;
- $P(c)$: the probability that a projection neuron projects on a given cluster ($P(c) = 0.25$ or $P(c) = d^{-a}$ with d being the distance and a a parameter) ;

- $P(p)$: the probability that a projection neuron projects on given neuron, given that it projects into the cluster of the latter ($P(p) < 0.1$) ;
- $P(l)$: the probability that an interneuron forms a connection with a given neuron's cluster ($P(l) < 0.1$).

After presenting the known anatomical data on the MRF, we will now present a proof that we have made while reviewing the anatomical data that $P(l) > 45 \times P(p)$.

1.4 Proof of $P(l) > 45 \times P(p)$

Let's define a few notations, in addition to those seen in the previous section:

- $P(I \rightarrow I)$ the probability that an interneuron forms a connection with a given interneuron in the same cluster (the sign *rightarrow* symbolizes a connection in the graph) ;
- $P(I \rightarrow P)$ the probability that interneuron forms a connection with a given projection neuron in the same cluster ;
- nb_I the number of interneurons in a cluster of the mRF ;
- nb_P the number of projection neurons in a cluster of the MRF.

Let's assume that :

- an interneuron projects only in its own cluster (the number of projections from interneurons outside their cluster is negligible) ;
- $P(c) = 0.25$. This is one of the two known anatomical models for $P(c)$. The second model, by conditioning $P(c)$ on the distance between the source cluster and the destination cluster, promotes a structure of type small-world compared to 0.25 as shown in Figure 1.6. Therefore, this assumption will generalize the results of the demonstration in the second model, since a small-world structure, intuitively and as also shown in Figure 1.6, requires a high number of connections in a cluster (defined by $P(l)$) compared to the number of inter-cluster connections (defined by $P(p)$), which favors the second model in comparison to the first one ($P(c) = 0.25$).

By construction, we have:

- $P(l) = P(I \rightarrow I) + P(I \rightarrow P)$ (because we do not consider that interneurons project in their own cluster) ;

- $nb_P/(nb_P + nb_I) = 80\%$ (the proportion of projection neurons, known in the literature), therefore $nb_P = 4 \times nb_I$.

As we have seen previously, 45% of the synapses of the projection neurons are GABAergic. In other words, 45% of the incoming connections on a projection neuron come from an interneuron, because interneurons are inhibitory and GABAergic (there may be exceptions, but negligible).

Formally, this means that $\frac{nb_I \times P(I \rightarrow P)}{nb_I \times P(I \rightarrow P) + nb_P \times P(p) \times P(c) \times c} = 45\%$
 which is equivalent to $nb_I \times P(I \rightarrow P) = (45\%/55\%) \times (nb_P \times P(p) \times P(c) \times c)$
 which can be also written $P(I \rightarrow P) = 4 \times (45\%/55\%) \times (P(p) \times P(c) \times c)$ (car $nb_P = 4 \times nb_I$).

Replace it with the values $P(c) = 0.25$ et $c = 55$ (average between 35 and 75) :

This therefore gives $P(I \rightarrow P) = 45 \times P(p)$

Yet $P(l) = P(I \rightarrow I) + P(I \rightarrow P)$

Hence $P(l) > 45 \times P(p)$

This result strongly underpins the theory that the mRF has a small-world structure [Humphries et al., 2006], as shown in Figure 1.6.

After presenting the structure of the mRF, we will now analyze the only two published models of the mRF.

1.5 Previous models

1.5.1 The Kilmer-McCulloch model - 1969

During the 1960s, W.L. Kilmer, W.S. McCulloch, and J. Blum published several articles offering the first model of the mRF based on anatomical studies of Scheibel & Scheibel. In 1969 they synthesized their research in a single article, which became a reference article on the subject [Kilmer et al., 1969].

The common thread in their model is the concept of mode of behavior: an animal at a given time follow one and only one mode of behavior, such as eating or sleeping. The mRF serves to switch from one mode to another. To support this hypothesis, when the RF is damaged we notice pathological mode switching [Jouvet, 1967]. It is hypothesized that a cluster is associated with exactly one mode of operation.

Three model variants are available:

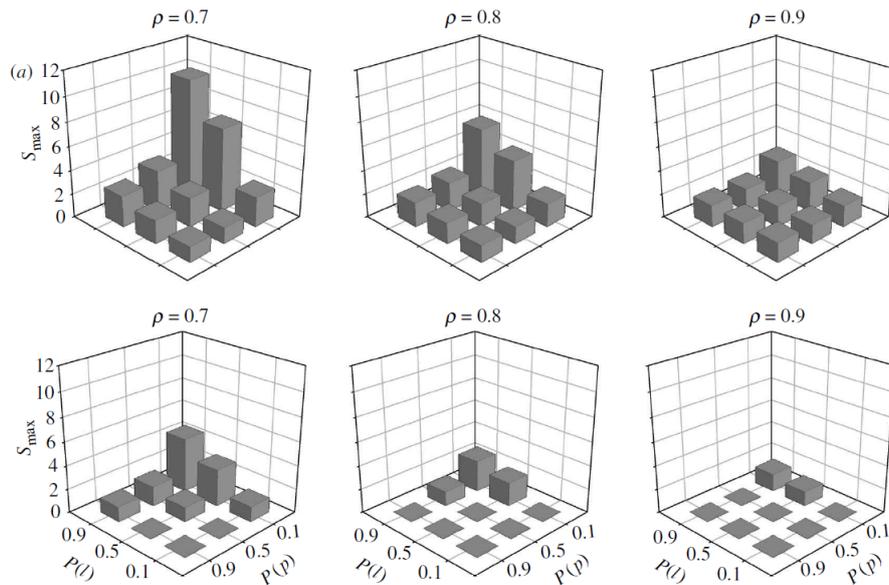


Figure 1.6: This figure shows the degree to which the mRF can be considered as a small-world network according to the values of $P(l)$, $(P(p)$ and p . On the Y axis, the value S_{max} shows the degree of small-worldness: if it is greater than 1, then the network is considered to be small-world. We clearly see that the proof of $P(l) > 45 \text{ times } P(p)$ greatly increases the chances that the mRF has a small-world structure. Source : [Humphries et al., 2006]

1. S-RETIC (S stands for Simple, Retic for Reticular): As the name suggests, this first model is simple and directly derived from anatomical data. This model consists of a dozen modules that receive stimuli and determine what mode to choose, each module corresponding to a cluster of mRF. These modules are provided with information to allow them to be both general, in order to remedy any default modules, and specialized, so as to be able to make a decision. The modules are more or less linked together according to the distance between them. Each module makes a decision and gives it a probability, and the final decision takes into account all the modules. If the consensus for a mode is large enough, then there is convergence to this mode. This model can easily be extended to more modes and modules. This model gives correct results, however it suffers from several shortcomings, the second variant will try to circumvent them.
2. STC-RETIC (STC stands for Spatio-Temporal Conditioning): This model enriches the S-RETIC model by introducing the concepts of development, generalization, discrimination, habituation and conditioning. To introduce these properties, additional information must be provided in order to indicate whether a stimulus or a mode of behavior is good, neutral or bad, like in reinforcement

learning. This learning occurs individually in each module and cooperatively between the various modules. The results for this model are not detailed. The article notes that STC-RETIC have several unattractive features: its connections do not strictly abide by the anatomical data of the RF and it can only change the mode when a new stimulus arrives.

3. H-RETIC (H stands for Hardware): It is a hardware version of the STC-RETIC model, but being designed at the time of writing the article and then never finished, as STC-RETIC software implementation had become too complicated to maintain and develop effectively with the tools of the time.

Nevertheless, the results obtained by the authors with this model are never really exposed, and inaccuracies about its description make it difficult to implement the model for concrete analysis. In addition, many considerations of the article are interesting but not explored: thus exposed, the model would not satisfy the requirements of current scientific literature. As a result, this first model gives a great synopsis of the mRF data while providing some interesting ideas, although not precisely described and actually implementable.

For almost 40 years, no new model of the mRF was proposed. In 2005, Mark Humphries [Humphries et al., 2005] tried to reproduce the Kilmer-McCulloch model and to evaluate its performance: he implemented the model in a simulated robot and a real robot. The latter was placed in the survival task inspired by [Girard et al., 2003] that we will detail later in the chapter 4.1: in this experiment, the robot has access to four variables (mRF inputs), from which he must choose an action (output of the MRF) amongst 5 actions it has at its disposal. The goal for the robot is to survive as long as possible and that the survival time reflects the quality of action selection.

The results were somewhat disappointing: while the model of the mRF often gets better results than a purely random model (that is to say a model in which decisions are made randomly without considering the input variables), the mRF model is well below a simple model of type Winner-Takes-All (WTA), whose decisions simply correspond to the highest of input variables.

1.5.2 Humphries model - 2006

In 2006, Mark Humphries and his colleagues presented their own model of the mRF in [Humphries and Prescott, 2006], the second model in the literature of the mRF. They adopted the classical formalism of neural networks and chose to use a population model, where each neuron model was a set of real neurons of mRF, as shown in the diagram 1.7. They kept the Kilmer-McCulloch hypothesis which assumes that a cluster is associated with exactly one action.

Each cluster consists of two neurons, each modeling the average activity of each two types of neurons:

- 1 excitatory neuron, projecting on all other neurons except those of its cluster;
- 1 inhibitory neuron, projecting only on himself and on the excitatory neuron of its cluster.

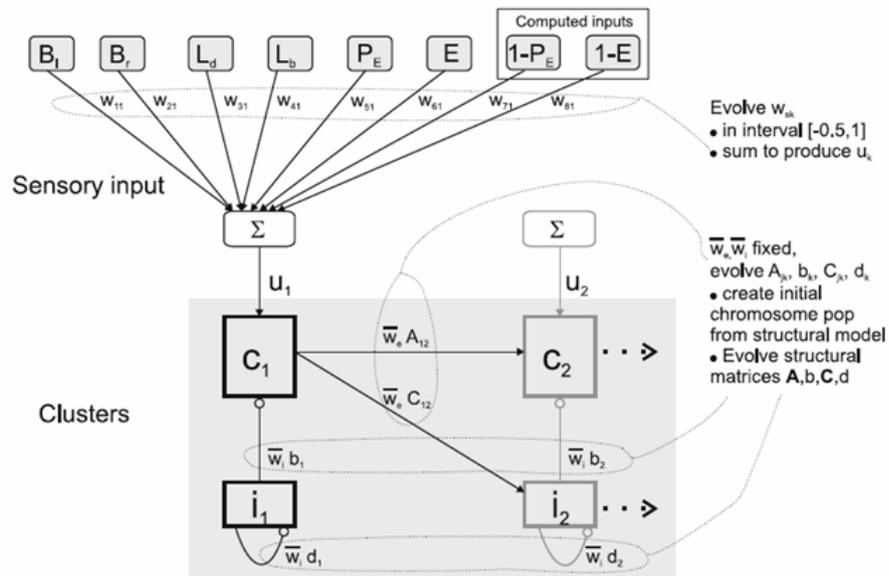


Figure 1.7: Humphries model. This diagram shows two clusters, the inputs the right cluster receive and the neural connections with the left cluster. Each cluster contains one inhibitory neuron i and one excitatory neuron c . Source : [Humphries and Prescott, 2006]

The weights of the connections on the one hand reflect the probabilities of connections from anatomical data and on the other hand the weight of actual connections in the mRF, unknown in the literature.

This model was also evaluated with a survival task inspired by [Girard et al., 2003] and some parameters have been optimized with evolutionary algorithms, which will be further discussed in chapter 2.2. The results also proved somewhat disappointing in terms of survival time, most simulations of this model were only slightly better compared to the purely random model.

1.6 Approach and objectives of this report

Although their results did not have very good performance in the survival task compared with a random controller, both models of the mRF had the merit to offer interesting food for thought as well as in the case of the second one, methods of evaluation of the model.

Our approach is to build a new model of the mRF, observing strictly the formalism of neural networks and in accordance with the anatomical data, while reducing the level of abstraction by increasing the number of neurons per cluster. We decided to remove the assumption made by the designers of the previous two models that a cluster corresponds to one and only one action, because no data to support this hypothesis and given the diversity of actions, it seems reasonable not to keep such a constraint.

In order to avoid defining all the network settings by hand and find the best network - best in terms of tasks for action selection and respect of anatomical constraints - we use evolutionary algorithms like Humphries. Unlike the latter who only defined a single goal in its evolution, we use a multi-objective evolutionary algorithm, which will allow us to optimize better our networks and analyze them more accurately.

We will evaluate our model with two action selection tasks: a *disembodied* task, showing a static capacity of a network to pick stocks, and an *embodied* task that will put the mRF model in real life conditions, inspired by the survival task which Humphries also uses for his model evaluation.

The next chapter will detail our approach and the tools we used to carry out this work.

Chapter 2

Theoretical material

To begin with, we will explain how the mRF will be formalized in our model with a neural network. Then, as we will evolve it by evolutionary algorithms, we present how they work and how we can make use of them. The technical aspects of the implementation of these theoretical tools will be discussed in appendix A.

2.1 Neural networks

A neural network consists of a set of neurons and a set of oriented connections linking neurons between them. Formally, we can consider it as a weighted directed graph where each node corresponds to a neuron.

There are different types of neurons: in our model, we use a variant of neurons with discharge rate of type leaky integrators, called IPDS (locally Projected Dynamical Systems) as they allow to model a population of neurons. We chose the IPDS because of their interesting stability properties, having been shown, for example, that the stability (in the meaning of contraction) of a nonlinear system consisting of IPDS directly results from the stability of the same IPDS-free system, which has not been shown with standard leaky integrators [Girard et al., 2008]. As we seek here to build a system for action selection, stability is preferable to instability.

An IPDS a neuron is characterized by two parameters:

- τ , corresponding to the time constant,
- *threshold*, corresponding to the activation threshold.

For the sake of simplicity, we fix τ to 5ms and set the *threshold* to 0 in order not

to increase the number of free parameters. We also fix the iteration step dt , which by construction must be always less than τ , to 1ms.

The first operation performed by the neuron is a sum of input values, weighted by the synaptic coefficients, i.e. the sum $w_1x_1 + \dots + w_mx_m = \sum_{j=1}^m w_jx_j$, where x_i are the inputs and w_i the synaptic coefficients.

We need to add the threshold $threshold$ to this formula: $\left(threshold + \sum_{j=1}^m w_jx_j \right)$

We will use the integration of IPDS by the Euler method. The activation function will therefore be the following, where a_n is the current internal value of the neuron, a_{n+1} the next internal value, equal to the output value:

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

As x is the weighted sum of the quantities received entries, it eventually gives us:

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

The \pm presents in the formula reflects the fact that a neuron can be either excitatory IPDS or inhibitory.

Figure 2.1 shows an example of a neuron, and the graph B.1 illustrates a neural network derived from our model corresponding to an mRF with four clusters.

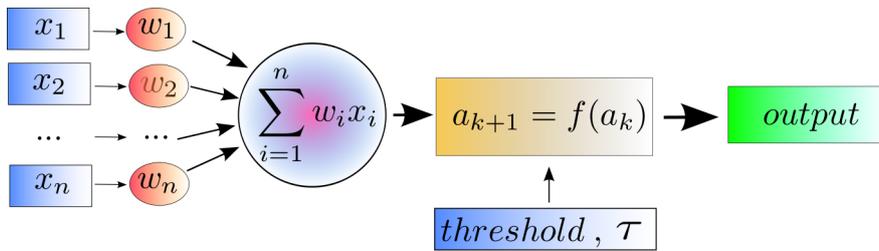


Figure 2.1: Example of a neuron with 2 inputs and an activation function with threshold.

Each cluster in the mRF has the same number of inputs and receives the same values. Similarly, each cluster of the mRF has the same number of outputs. At the global level of the mRF, the output values correspond to the average output values of each cluster. Figure C.1 shows a cluster. A cluster has various numbers of neurons and connections.

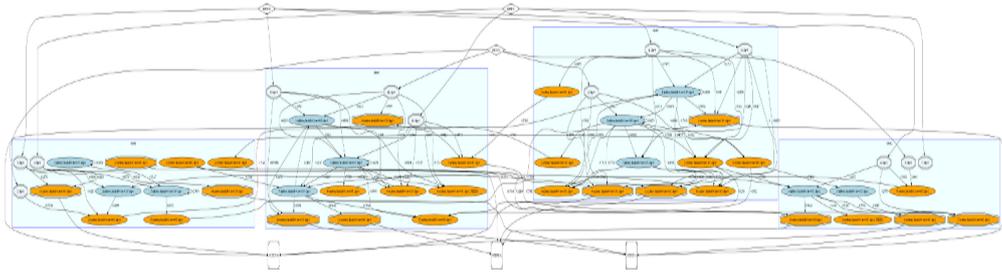


Figure 2.2: Example of an mRF with 4 clusters. Excitatory neurons are orange and dark blue neurons are inhibitory. A cluster represented by a blue rectangle. Neurons located outside the blue rectangles represent inputs received by the mRF and the neurons to which it projects. This figure can also be found enlarged in Appendix B.

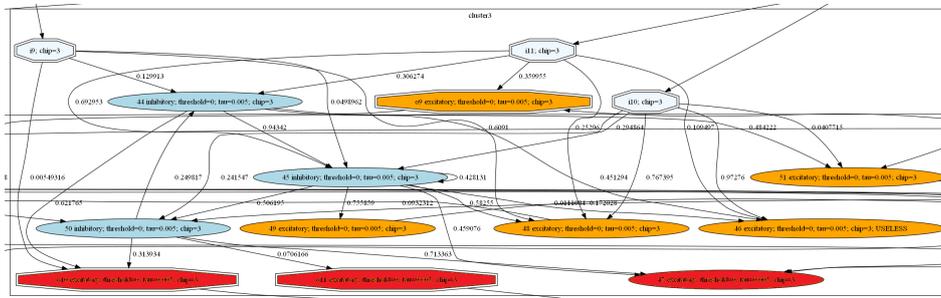


Figure 2.3: Example of a cluster of the mRF. Excitatory neurons are orange and dark blue neurons are inhibitory. Each synaptic connection has a weight between 0 and 1. The three neurons in light blue are the inputs (input neurons), the three neurons in red are the outputs of the mRF (output neurons). This figure can also be found enlarged in Appendix C.

As these networks contain a large amount of neural connections and parameters, it would be tedious to optimize them by hand to study how their structure enables action selection. Therefore, we chose to use evolutionary algorithms to find solutions since this optimization method has interesting properties for our problem as we shall see in the next section.

2.2 Evolutionary algorithms

2.2.1 Definitions

Evolutionary algorithms are a family of optimization algorithms based on the principle of **Darwinian natural selection**. As part of natural selection, a given environment

has a population of individuals that compete for survival and reproduction. The ability of each individual to achieve these goals determines their chance to have children, in other words to pass on their genes to the next generation of individuals, who for genetic reasons will have an increased chance of doing well, even better, in realizing these two objectives.

This principle of continuous improvement over the generations is taken by evolutionary algorithms to optimize solutions to a problem. In the **initial generation**, a **population** composed of different **individuals** is generated randomly or by other methods. An individual is a solution to the problem, more or less good: the quality of the individual in regards to the problem is called **fitness**, which reflects the adequacy of the solution to the problem to be solved. The higher the fitness of an individual, the higher it is likely to pass some or all of its genotype to the individuals of the next generation.

An individual is coded as a **genotype**, which can have any shape, such as a string (genetic algorithms) or a vector of real (evolution strategies). Each genotype is transformed into a **phenotype** when assessing the individual, i.e. when its fitness is calculated. In some cases, the phenotype is identical to the genotype: it is called **direct coding**. Otherwise, the coding is called indirect. For example, suppose you want to optimize the size of a rectangular parallelepiped defined by its length, height and width. To simplify the example, assume that these three quantities are integers between 0 and 15. We can then describe each of them using a 4-bit binary number. An example of a potential solution may be to genotype 0001 0111 01010. The corresponding phenotype is a parallelepiped of length 1, height 7 and width 10.

Last definition before applying these theories to our model of the mRF, during the transition from the old to the new generation are called **variation operators**, whose purpose is to manipulate individuals. There are two distinct types of variation operators:

- the **mutation operators**, which are used to introduce variations within the same individual, as genetic mutations;
- the **crossover operators**, which are used to cross at least two different genotypes, as genetic crosses from breeding.

We chose the evolutionary algorithms because they have proven themselves in various fields such as operations research, robotics, biology, finance, or cryptography. In addition, they can optimize multiple objectives simultaneously and can be used as black boxes because they do not assume any properties in the mathematical model to optimize, which allows us in our case to optimize a dynamic and nonlinear as a

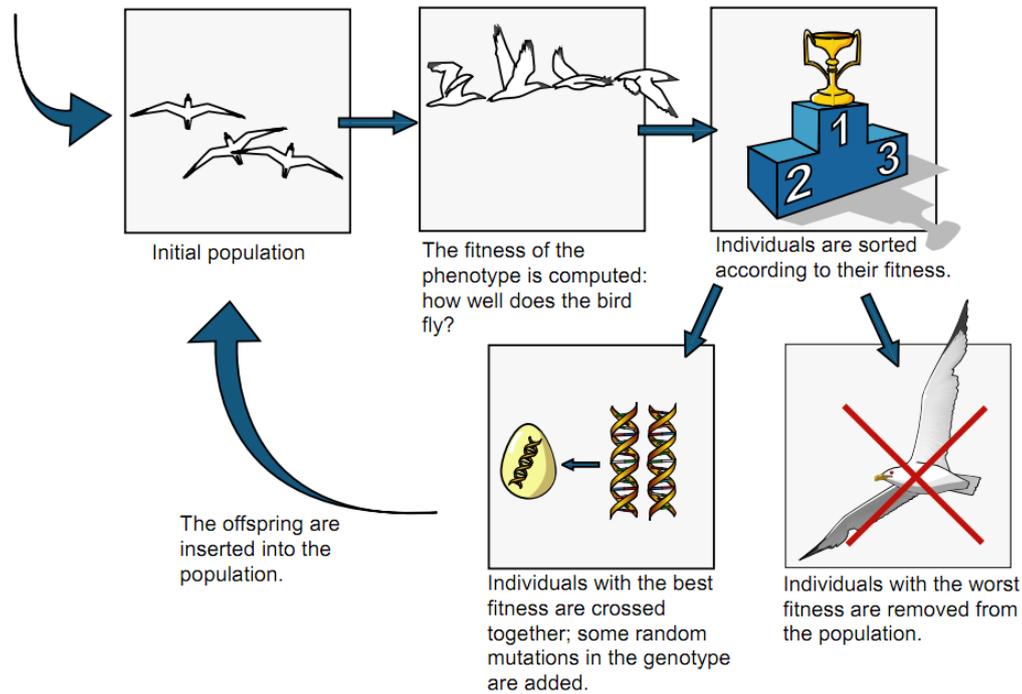


Figure 2.4: Functioning of an evolutionary algorithm: from an initial population of solutions, they are ranked according to their fitness, the worst ones are eliminated and the best ones are used to produce new solutions. Source: [Doncieux et al., 2004]

neural model. Their only real limitation is the computational complexity, hence the decision to code our program in a fast (C++), multi-threaded language, which we run on a computer cluster. Appendix A gives details on the technical aspects of the implementation.

2.2.2 Application

In our model, the mRF is modeled as a neural network. The chosen genotype for the implementation is a set of neural networks, each corresponding to a cluster of mRF, and a vector containing all the connections between clusters, which we call **interconnections**. The phenotype is obtained from the genotype by copying each of these networks within a large network, the mRF, to which we add the interconnections.

Our mutation operators are:

- Add/delete a neuron ;

- Add/delete/modify a connection (intra-network) or an interconnection (inter-network).

We could have over the mutations change other settings, such as certain properties of neurons (e.g. inhibitory / excitatory), however we preferred to limit the degree of freedom of the evolution. Also, we haven't used crossover operators: although intuitively we might think it would be interesting to cross mRFs so as to allow them to mix their clusters, such an operation is a very difficult to implement because on one hand the interconnections are unique to each cluster and to each mRF and on the other hand such crossings are not really interpretable in terms of evolution as the role of each cluster is not defined a priori.

One of the hardest part was the implementation of the anatomical constraints of the mRF, so that evolution produces neural networks consistent with the anatomical data. We have implemented them in two complementary levels:

- upstream, at the level of mutation operators: at each mutation, we strive to remain close to anatomical data;
- downstream, during the calculation of fitness: we used a multi-objective evolutionary algorithm, which allows us to define an objective of anatomical plausibility, pushing the networks to meet the anatomic constraints.

The way objectives are defined significantly impacts the results. We have established an objective for anatomical plausibility, as well as specific objectives for action selection tasks that we will detail in the next section.

Lastly, we chose to use the algorithm **NSGA-II** [Deb, 2001, Deb et al., 2002], which is to date one of the most efficient multi-objective evolutionary algorithms and by far the most used. Unlike a single-objective algorithm where there is only one best individual (possibly with equally placed individuals), the best individuals from a multi-objective evolution will form a front called the Pareto front, of a size equal to the number of objectives. Figure 2.5 shows a 2-dimensional Pareto front and Figure 2.6 compares the overall performance of a single-objective algorithm over the entire results of a multi-objective algorithm.

Now that we have presented one the one hand the mRF and on the other hand the theoretical tools that we used for the project while explaining their role in modeling the mRF, we will detail the action selection experiences we carried out and analyze the results in the next chapter.

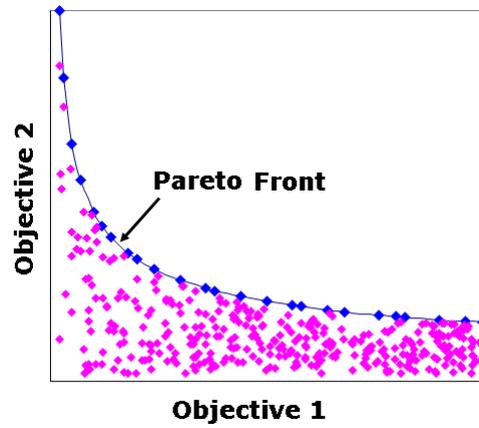


Figure 2.5: Example of a 2-dimensional Pareto front: unless otherwise stated, evolutionary algorithms maximize objectives unlike most optimization algorithms whose goal is to minimize them.

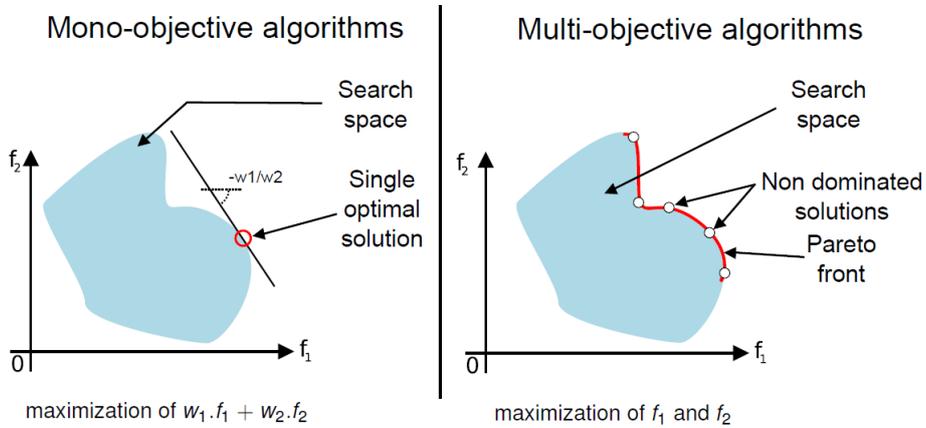


Figure 2.6: Single vs multi-objective. A single-objective algorithm will give a result, while a multi-objective algorithm gives a set of results. On the left figure, f_1 and f_2 correspond to scores for objectives 1 and 2. w_1 and w_2 are weights assigned to the two scores, the linear combination corresponding to a single objective in order to use a single-objective algorithm. Source : *Stéphane Doncieux et Jean-Baptiste Mouret*.

Chapter 3

Disembodied task: the abstract vector task

In this chapter we will present the first experiment to evaluate our model of the mRF. First, we will describe the experiment, then we will analyze the results.

3.1 Experiment

The disembodied task is an abstract selection task. The principle is simple: we provide a vector of dimension 3 as the input to the mRF, with values between 0 and 1. This input vector represents a set of values from various data sources.

The purpose of the mRF is to select one action. We got our inspiration from [Humphries et al., 2007]: we decided that the mRF will be considered having chosen the right action if and only if its output vector, also of dimension 3, have its maximum value in the same dimension as the one containing the maximum value in the input vector. For example, if the input vector is (0.3, 0.5, 0.1), the MRF will select the right action if and only if its output vector is of the form (a, b, c), where $b > a$ et $b > c$, as illustrated by figure 3.1.

In order that the evolution does not over-learn from a small set of vectors, which would prevent a good generalization, the mRF will be assessed on its ability to select an action on the set of 3D vectors where we vary each element from 0 to 1 with a 0.1 step, and removing vectors having a maximum number of components so as to eliminate ambiguous cases. Here is formally written the set of vectors, which contains a total of 1155 elements:

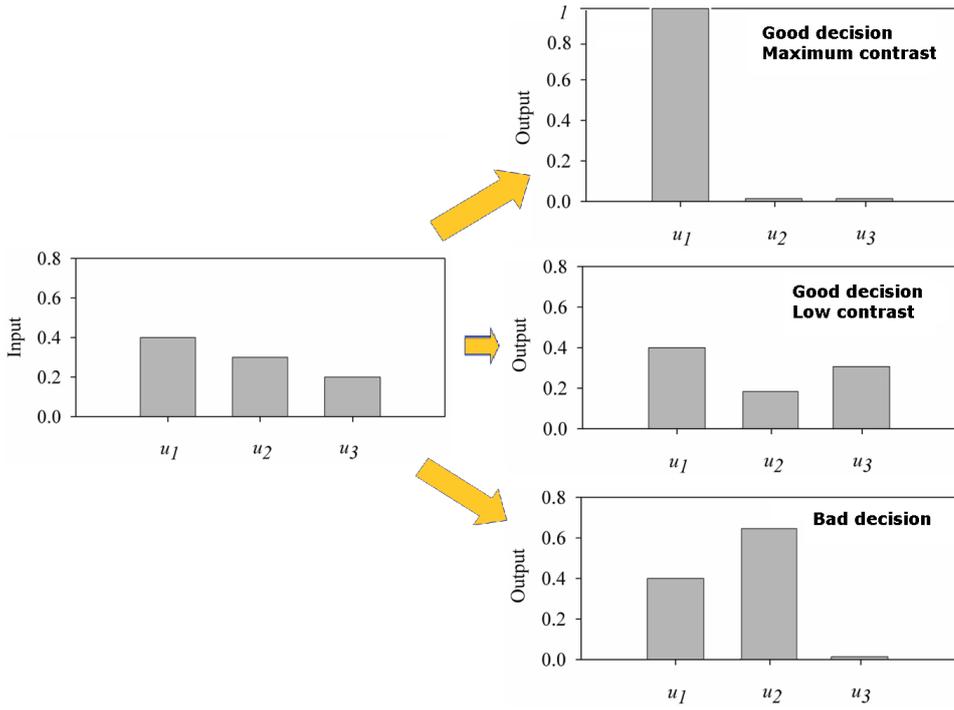


Figure 3.1: This diagram shows the first two goals of the task disembodied: the network receives as input a 3D vector whose values are between 0 and 1, and it must make the right decision (in other words select the right dimension) while maximizing the contrast.

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

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

Nevertheless, this first objective allows uninteresting solutions, such as a neural network that simply copies the input in the output. To make a real action selection, we defined a second goal forcing the mRF to maximize the contrast between the maximum value of the output vector and other values. The general formula for calculating the contrast is as follows, where x_i is the i^{th} element of vector X and n its dimension:

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

Take for example the vector $(0.6, 0.3, 0.5)$, the contrast is calculated as follows:

$$\text{contrast} = \sqrt{\frac{((0.6-0.5)^2 + (0.6-0.3)^2)}{2}} \approx 0.22. \quad \text{The higher the contrast, the higher the action is clearly selected.}$$

The third and final objective will be to observe the anatomical constraints for the evolution of networks get as close as possible to a mRF-like network structure. The score for this objective of anatomical plausibility is a linear combination of scores of sub-goals, each of them representing one of the following anatomical constraints (we use the notations of chapter 1.3):

- p must be equal to 80%, the formula for the score is $score_p = (-1) \times (0.8 - p)^2$,
- $P(c)$ must equal 25%, the formula for the score is $score_{pc} = (-1) \times (0.25 - pc)^2$,
- the synapses of projection neurons are 45% GABAergic, the formula for the score is $score_{gabaergic} = (-1) \times (0.45 - percent_{gabaergic})^2$,
- the interneurons can not project outside their cluster, let $nb_interneuron_not_within_chip$ be the number of interneurons which do not respect this constraint,
- the projection neurons may not project into their cluster, let $nb_projection_within_chip$ be the number of projection which do not respect this constraint.

We did not put constraints on $P(p)$ and $P(l)$ as known anatomical data about them are quite vague. As we remarked in Chapter 2.2, evolutionary algorithms maximize the objectives, hence the negative scores as approaching the constraints is equivalent to a score approaching 0.

The formula for calculating the overall score of the third goal is:

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

The weights 8, 1 and 0.5 were arbitrarily chosen, the only criterion being that each has a value weighted scores very roughly the same order of magnitude, to maximize the chances that evolution optimizes each of them, not just some potentially at the expense of others.

In summary, here are the three goals we have identified:

- Objective 1: number of right decisions. Minimum: 0, maximum: number of vectors tested,
- Objective 2: value of the contrast. Minimum: 0, Maximum: 1 (because all vector values are between 0 and 1),

- Objective 3: anatomical plausibility. minimum: $-\infty$, maximum: 0.

In addition to these three goals downstream, we also applied anatomical constraints upstream, that is to say, within the definition of mutation operators to push the evolution to comply as much as possible with the data anatomy. We already mentioned in chapter 2.2.2 this double application of constraints. Here are the constraints we have included within the mutation operators:

- impossibility for a projection neuron to project in their cluster,
- impossibility for an interneuron to project outside their cluster,
- $P(c)$ must be close to 25% ;
- p must be close to 80%.

As we see, these constraints only contain a portion of the known anatomical data: the interest to add them to the mutation operators is that evolution produces fewer individuals clearly implausible from an anatomical point of view. However, putting too much stress in mutation operators would be risky, because besides its computational cost, restricting the search space excessively can make it more difficult to obtain a good solution.

Notwithstanding its apparent simplicity, this disembodied task allows us to validate or invalidate the possibility of an anatomical structure similar to the mRF to make selections. We will analyze the results in the next section.

3.2 Results

First, here are the parameters that we use throughout the experiments, unless otherwise indicated:

- population size: 500 individuals,
- number of generations: 500,
- number of clusters: 4,
- initial number of neurons per cluster (in addition to the input and output neurons): uniform random between 3 and 10, the clusters may have a different number of neurons in the same mRF,
- probability of adding/deleting a neuron/connection (inter- or intra-cluster): 0.05

- probability of modifying the weight of a connection (inter- or intra-cluster): 0.1,
- number of iterations to propagate an input in the network: 100.

After 300 generations, the computed evolutions generate individuals that make more than 95 % of correct decisions, in some cases 100 %, almost perfectly respecting anatomical constraints and whose output vectors have a contrast greater than 0.6. Chart 3.2 shows a 2D Pareto front typically obtained when the scores represents the objectives of good decisions and contrast obtained by each individual. Chart 3.3 shows a 3D Pareto with the scores of the 3 objectives.

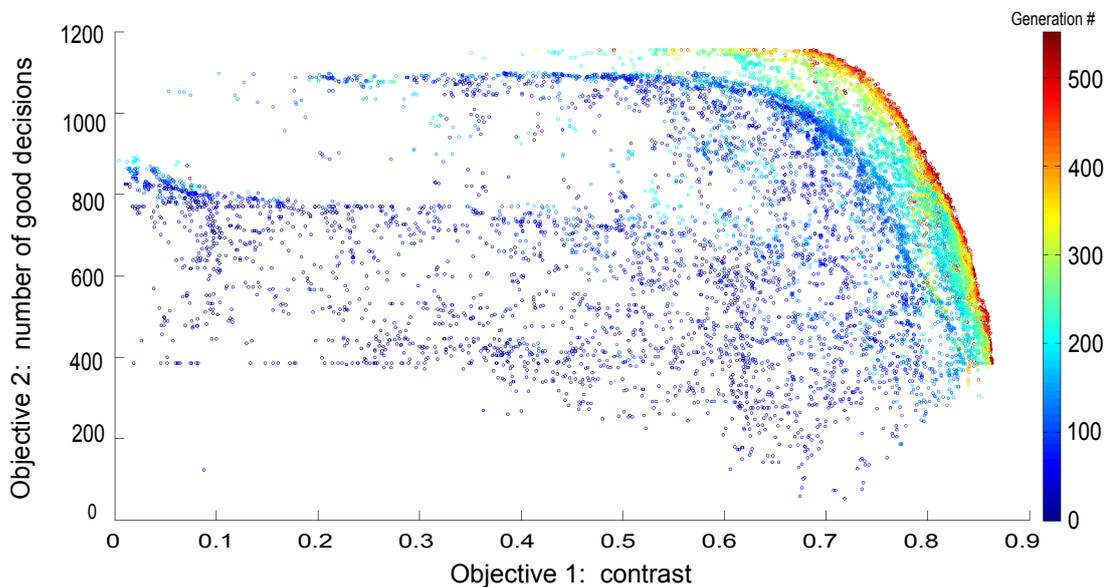


Figure 3.2: 2D Pareto front showing the scores for the contrast objective on the x-axis and of the good decision objective on the y-axis obtained by each individual of each generation during the evolution.

Let's take the 5 best individuals having the maximum number of good decisions (1155) and look at the average results:

- number of good decisions score: 1155 (which corresponds to the theoretical maximum);
- contrast score: 0.68761 (the theoretical maximum is 1) ;
- anatomical plausibility score: $-1.20792e-15$ (the theoretical maximum is 0).

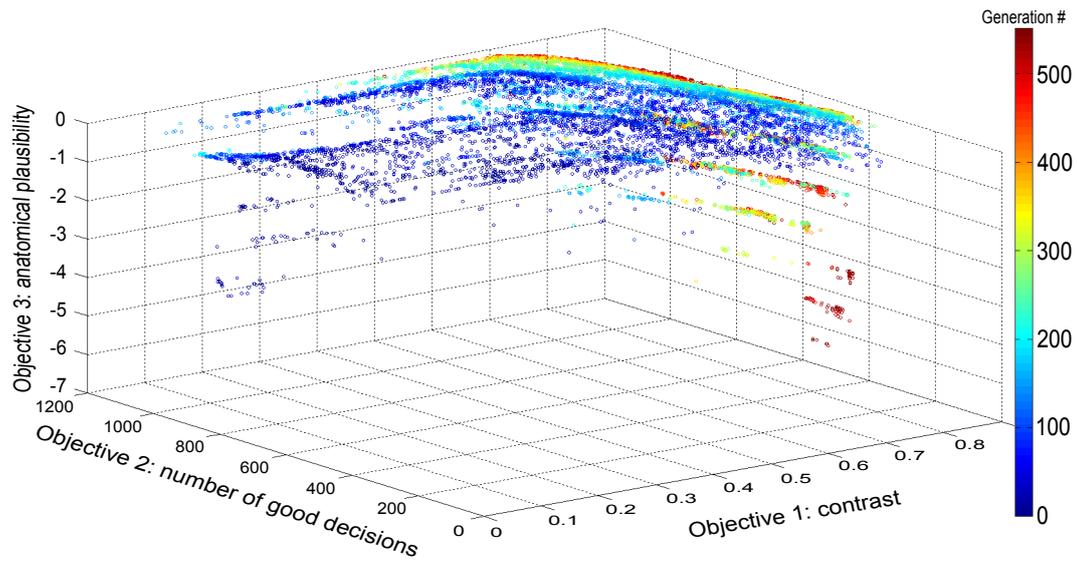


Figure 3.3: 3D Pareto front showing the scores for the contrast, good decision and anatomical plausibility objectives obtained by each individual of each generation during the evolution.

The following table summarizes the average anatomical statistics of these five individuals.

	Statistics of the top 5 individuals	Constraint on mutation operators	Constraint in objectives	Known anatomical data
total number of neurons	57	None	None	NA
total number of connections	117	None	None	NA
p	80%	80%	80%	80%
$percent_{gabaergic}$	45%	None	45%	45%
$P(c)$	25%	25%	25%	25%
$P(p)$	8.73626%	None	None	< 10%
$P(l)$	8.61552%	None	None	< 10%
Number of interneurons projecting outside their cluster	0	0	0	0
Number of projection neurons projecting in their cluster	0	0	0	0

This result shows that the mRF can perform a task of action selection while respecting the known anatomical data. For comparison, [Humphries et al., 2007] from which this task was inspired gets about 75% of good decisions, without taking into account the contrast, and considers that this is sufficient to show the possibility of an action selection.

We have also run the program by removing the constraints located in the mutation operators. As shown in figure 3.4, this has the effect of increasing the number of individuals with low anatomical plausibility, since we see that the 2D front representing the scores of the objectives of good decisions and contrast obtained by each individual is much less clear than the one obtained in the initial experience. This observation corresponds to the intuition that we had. Still, the evolution always generates individuals - certainly rarer than in the initial experiment - almost perfectly anatomically plausible, with a contrast greater than 0.5 and whose rate of good decisions is greater than 95%.

Another variant we tried was to remove the constraints of the mutation operators

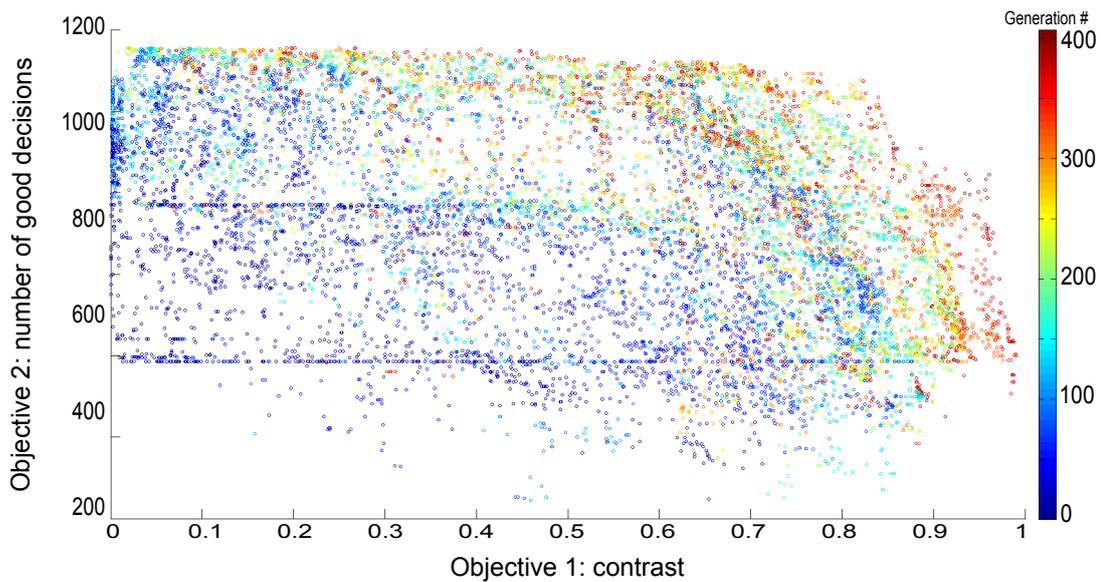


Figure 3.4: 2D Pareto front without any constraints in the mutations. The front shows the scores of the objectives of good decisions and contrast obtained by each individual.

and eliminate the objective of anatomical plausibility. In other words, we wanted to see how a system would evolve without anatomical constraint on the selection task.

The results show that on the one hand the networks achieved performance similar to those obtained with the constrained networks or even slightly better (see chart 3.5) if we consider the contrast, and on the other hand through statistical analysis on the structure of these networks we see they do not tend to mRF-like structures. This means that the known anatomical data on the mRF represent neither an advantage (because there are other network structures just as effective) nor a disadvantage for selection.

This first series of experiments based on a disembodied task showed the computational capacity of the mRF to perform a selection task. Our model is more efficient than that Humphries' in this task and we have added an additional constraint, namely the contrast. We will now make a second series of experiments to analyze the performance of the mRF in an embodied task of robotic simulation.

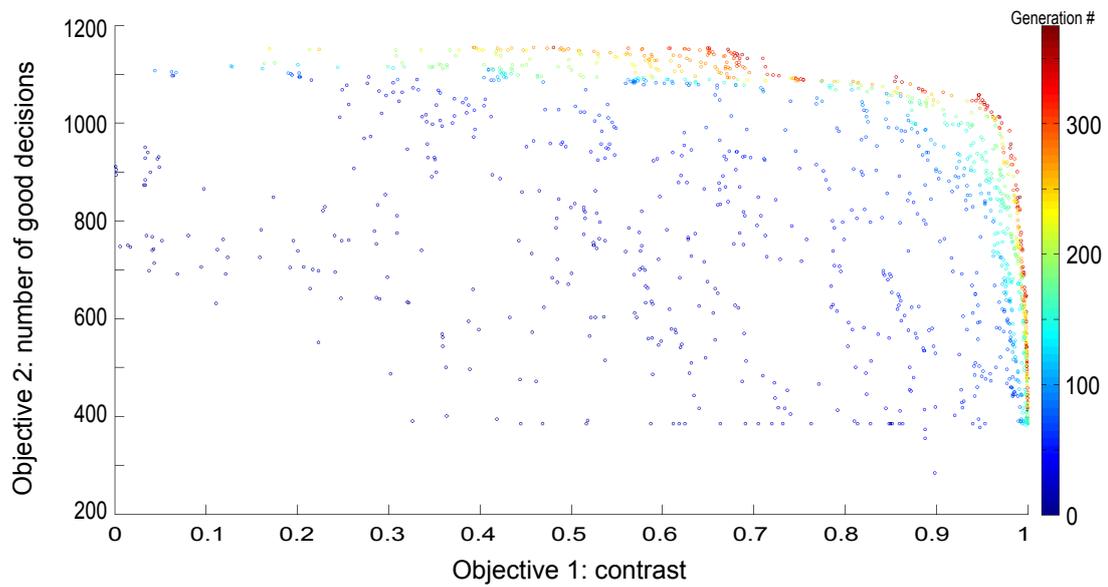


Figure 3.5: 2D Pareto of the evolution without any anatomical constraint (neither in mutation operators, nor in objectives). The front shows the scores objectives contrast on the abscissa and on the ordinate good decisions made by each individual in each generation during evolution.

Chapter 4

Embodied task: the survival task

Like the previous chapter, we will in this chapter first present the experience at first, and secondly we will analyze the results.

4.1 Experiment

The survival task that we use here is strongly inspired by [Girard et al., 2003]. In this task, a robot has to select effective behaviors in order to ensure its survival, maintaining its internal state variables in tolerable ranges, the area of viability [Ashby, 1952]. This survival depends directly on the robot's ability to refuel with two different types of resources in a limited time by its level of recharge. The use of two different resources forces the robot to move in the environment to access them and then puts it in conflict to determine which resource is a priority at any given time, which is likely to generate behavioral oscillations. This task will be simulated on a computer.

The robot will be placed in an environment where it can find two types of resources: *ingestion* areas where it can stock up and *digestion* areas where it can *assimilate* these stocks and turn them into usable energy. Since all the behaviors of the robot consumes energy, it will therefore have to alternate *ingestion* phases with *digestion* phases in order to survive.

The experimental environment is a 2D surface of 400x400 units surrounded by walls. It is covered with 25 tiles of 80x80 units, with three different tile types: 21 gray tiles (neutral zones), two black tiles (ingestion areas), whose resources are inexhaustible, and 2 white tiles (digestion areas). Figure 4.1 shows the setting.

The robot is a disc of radius 20 units and has two internal variables:

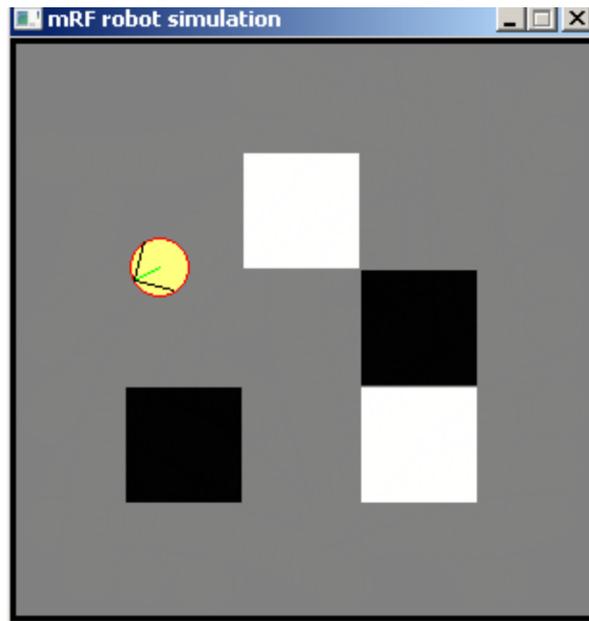


Figure 4.1: Setting of the survival task. We see 21 gray tiles, two black tiles and 2 white tiles. All tiles have the same size. The yellow circle represents the robot.

- the **Potential Energy** (PE) corresponds to the reserves drawn from the ingestion areas of intake, between 0 and 1,
- the **Energy** (E) is the energy actually used to survive in the environment; it is obtained by digestion of the PE on light areas. To survive, the robot must maintain E above 0 and the maximum of E is 1.

The robot has access to 4 external sensors:

- a **left bumper** (B_L), binary value, 1 if contact with the wall, 0 otherwise,
- a **right bumper** (B_R), binary value, 1 if contact with the wall, 0 otherwise,
- a **floor sensor for black tiles** (L_D pour light darkness), binary value, 1 if above a black square, 0 otherwise;
- a **floor sensor for white tiles** (L_B pour light brightness), binary value, 1 if above a white square, 0 otherwise;

The robot has five actions available to it:

- **Wander**: the robot moves randomly (random rotation between 0 and 9 followed by a forward move of 5 units). Note that in the absence of navigation

capabilities and memory on the environment, this behavior is the only way for the robot to find areas of refills (black or white tiles). This action takes 2 time units.

- **Avoid an obstacle:** the robot performs a backward move of 60 units followed by a rotation of 180 °. This action takes 2 time units.
- **Reload on black tile:** the robot stops and reloads its PE: $\delta PE = 0.027 \times L_D$. The robot can only recharge its PE if it is on a black tile. This action takes one time unit.
- **Reload on white tile:** the robot stops and reloads its E: $\delta E = 0.027 \times L_B$ and $\delta PE = -0.027 \times L_B$. The robot can only recharge its E if it is on a white area. This action takes one time unit.
- **Rest:** the robot does nothing. This action takes one time unit.

At each unit of time, the robot consumes 0,002 of its energy, even when he chooses to rest. If the energy becomes negative or zero, then the robot dies.

To choose an action, the mRF receives as input the **saliency** of each action calculated from the internal and external variables. The saliency is the degree of urgency or motivation to perform an action. The formulas used here to calculate the saliency are the same as in the evaluation of the mRF model with [Humphries et al., 2005] and in the evaluation of a model of the basal ganglia by [Girard et al., 2003].

- $S_{wander} = -B_L - B_R + 0.8(1 - P_E) + 0.9(1 - E)$
- $S_{avoid} = 3B_L + 3B_R$
- $S_{reload.on.dark} = -2L_B - B_L - B_R + 3L_D(1 - P_E)$
- $S_{reload.on.light} = -2L_D - B_L - B_R + 3L_B(1 - E)\sqrt{1 - (1 - P_E)^2}$

If the mRF fails to converge with the saliency input data, then the action off is selected. In our experiment, a new vector of saliencies is propagated during 100 iterations (one iteration lasts 1 ms) in the mRF and we consider that there is convergence when over the last 50 iterations the variation of each value of the output vector of the mRF is less than 0.001. When the mRF converges, then we consider the action selected is the highest output of the mRF. We will also test a variant in which the selected action is modulated by the contrast of the output vector of the mRF.

4.2 Results

To evaluate each robot controller, we simulated five survival tasks in the fitness computation, each time by placing the robot at a random location on the map and initialized with 0.5 of energy and 1 of potential energy. As each unit of time the robot uses 0,002 of its energy, its minimum life time is subsequently 500 units of time. So that the assessment takes a reasonable time of computation, we limited the simulations to 3000 time units each.

First, to ensure a sufficient complexity of the task, we assessed a random controller, which randomly chose one of the five possible actions. The results of this controller shows a median survival between 500 and 600 units of time, confirming that the task can not be resolved by a random controller and provides a basis for comparison.

A second prior verification we conducted was to test a Winner-Takes-All (WTA) type controller, the latter choosing the action based on the highest salience. Our first results have shown that these controllers had a life often approaching 3000 units of time, which meant that the task was too simple to evaluate our model of the mRF correctly. We subsequently tried to find a factor complicating the task and we found that the speed of the robot, not specified in [Humphries et al., 2005], greatly influences the results. Initially, in our experiment, when the robot controller chooses the action to explore randomly, it moves forward 10 units after performing a random rotation. When we reduce the displacement to 5 good speed units, the task becomes more difficult and a controller WTA has only average life of about 1250 units of time, which is near the maximum life span (3000 time units). Therefore, the task seems nontrivial to achieve. Figure 4.2 compares the survival time for random controller and WTA in 1000 survival tasks .

The results of the mRF controller show that within a few generations the robot is able to live more than 2500 units of time or even the maximum 3000. However, it takes several hundreds of generations for the contrast to become important. Figure 4.3 shows the evolution of the 2D Pareto front of a mRF controller after 500 generations, the survival time is near or equal to the maximum, the contrast is about half the theoretical maximum for the best individuals. Scores for the anatomical plausibility objective show that the mRF meets almost exactly the known data on the mRF. Regarding the time spent on average on each activity, the figure 4.5 shows that there is no significant difference between the different controllers, with the exception of the trivial random controller.

We tried to force the mRF to have a higher contrast by modulating the actions based on the contrast value. To this end, we have redefined every five actions to include the value of the contrast, f is the modulation function of the contrast:

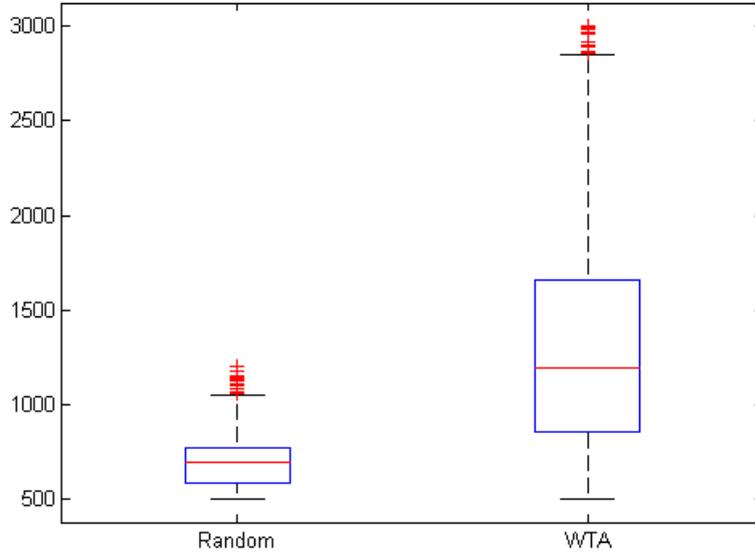


Figure 4.2: Comparison of survival time for random controllers and WTA over 1000 survival tasks.

- **Wander:** the robot moves randomly (random rotation between 0 and $f(\text{contrast}) \times 9^\circ$ followed by a forward move of $f(\text{contrast}) \times 5$ units).
- **Avoid an obstacle:** the robot performs a backward move of $f(\text{contrast}) \times 60$ units followed by a rotation of $f(\text{contrast}) \times 180^\circ$.
- **Reload on black tile:** $\delta PE = f(\text{contrast}) \times 0.027 \times L_D$.
- **Reload on white tile:** $\delta E = f(\text{contrast}) \times 0.027 \times L_B$ et $\delta PE = -f(\text{contrast}) \times 0.027 \times L_B$.
- **Rest:** the robot does nothing.

Taking as a modulation function $f(x) = \sqrt{x}$ and evaluating each network over 5 survival tasks, mRF networks we obtained have survival times which are similar or slightly lower than those WTA controllers. However, the contrast is a little better than when we introduce a modulation function, as shown in figure 4.6. Therefore, the modulation function introduced a selection pressure favoring the contrast at the expense of the survival time. The anatomical plausibility objective always has a score between -1 and 0, which means that networks have an mRF-like structure.

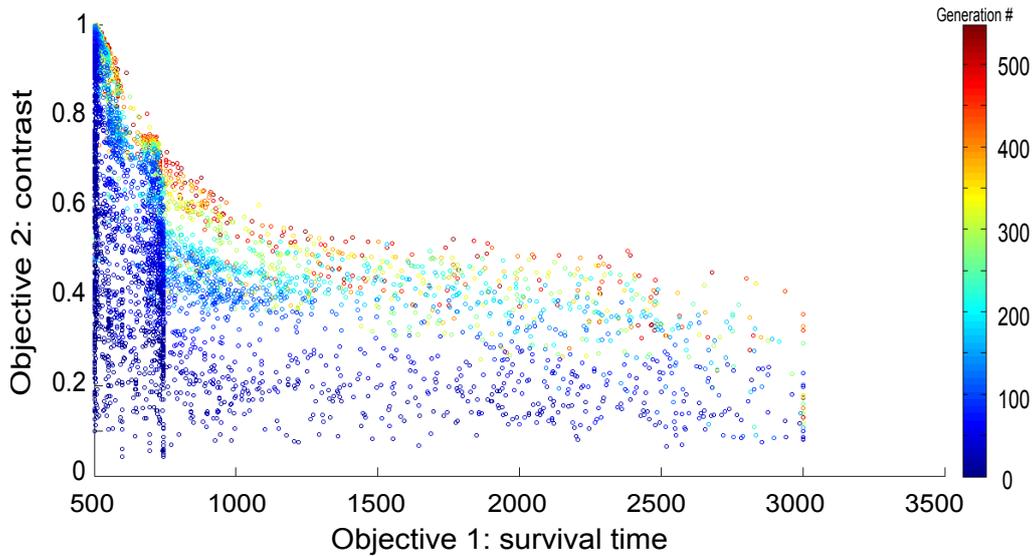


Figure 4.3: 2D Pareto front for the evolution of the mRF controller representing the scores of the objectives survival time (abscissa) and contrast (ordinate) obtained by each individual in each generation during the evolution. The mRF controller was evaluated on five survival tasks.

As another variant of the initial experiment, we tried to make the task more realistic by removing the computation of saliences and giving directly as the network's inputs the 4 external variables (B_L , B_R , L_D et L_B) as well as the 2 internal variables (E et EP), which has the effect of complicating the task. We have also given as inputs $1 - E$ and $1 - EP$ so as to lower the task's complexity, following the experiment's settings of [Humphries and Prescott, 2006]. The networks have therefore in this variant 8 inputs and 4 outputs.

The chart 4.7 shows the Pareto fronts obtained after 1000 generations. We see that some networks have a lifetime exceeding 2000 units of time, however the contrast is very low (below 0.1). Compared with the results obtained with the model of Humphries [Humphries and Prescott, 2006], the latter indicates that most of the networks obtained only manage to do slightly better than a random controller, but the actions are modulated according to contrast like what we did in the previous experiment. Here, our networks clearly succeed better than a random controller, however the actions are not modulated.

These results show that the mRF is generally more effective than a WTA network and a controller even more random, as summarized in Figure 4.4, unlike the experiments

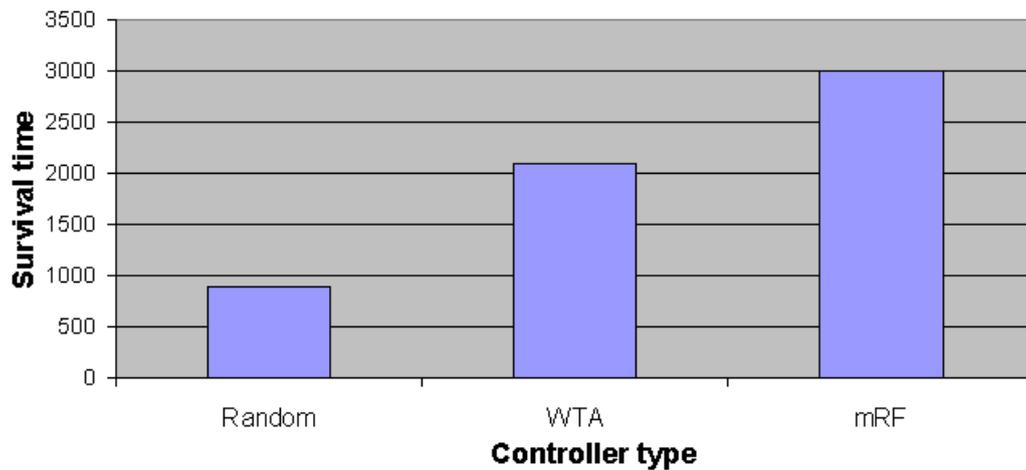


Figure 4.4: Comparison of the best average survival time on 5 survival tasks of the random, WTA and mRF controllers. For each type of controller, we performed 20,000 evaluations, each consisting in 5 survival tasks and average duration of survival. We have kept here that the highest found average for each type of controller.

of [Humphries et al., 2005] who failed to evolve the mRF optimally enough to exceed the WTA. This means that the mRF is not only able to perform action selection, but that it can also cope with complex situations where a WTA network is not efficient enough. However, in the variants in which we tried to make the task more realistic survival by modulating the actions based on the contrast of the mRF output vectors or giving it directly the internal and external variables without prior calculation of salience, the results we have obtained are less convincing: it would be interesting to further study these variants. It would also be interesting to quantify the propensity of our mRF networks to generalize their performance by evaluating them over a bigger number of survival tasks.

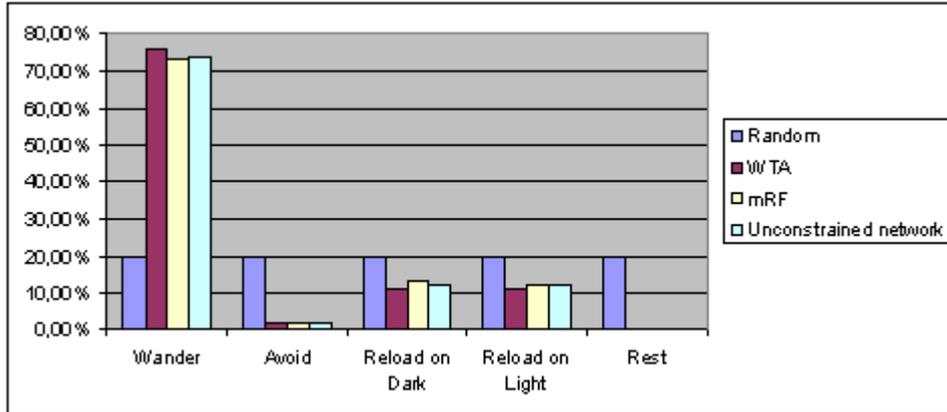


Figure 4.5: Comparison of time spent on each activity by the random, WTA and mRF controllers. These statistics aggregate the data of 500 survival tasks for each of the controllers with the highest survival times among the 50,000 simulated survival tasks.

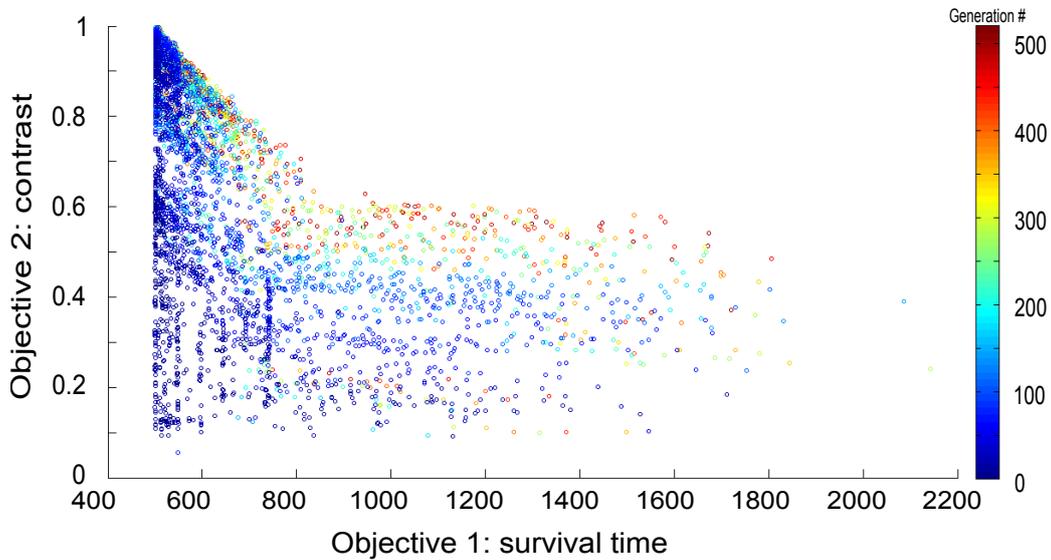


Figure 4.6: 2D Pareto front of the evolution of a mRF controller representing the scores of the objectives of survival time (abscissa) and contrast (ordinate) obtained by each individual in each generation during evolution, with modulation of action based on the contrast value. The modulation function is $f(x) = \sqrt{x}$ and the mRF controller was evaluated on five survival tasks.

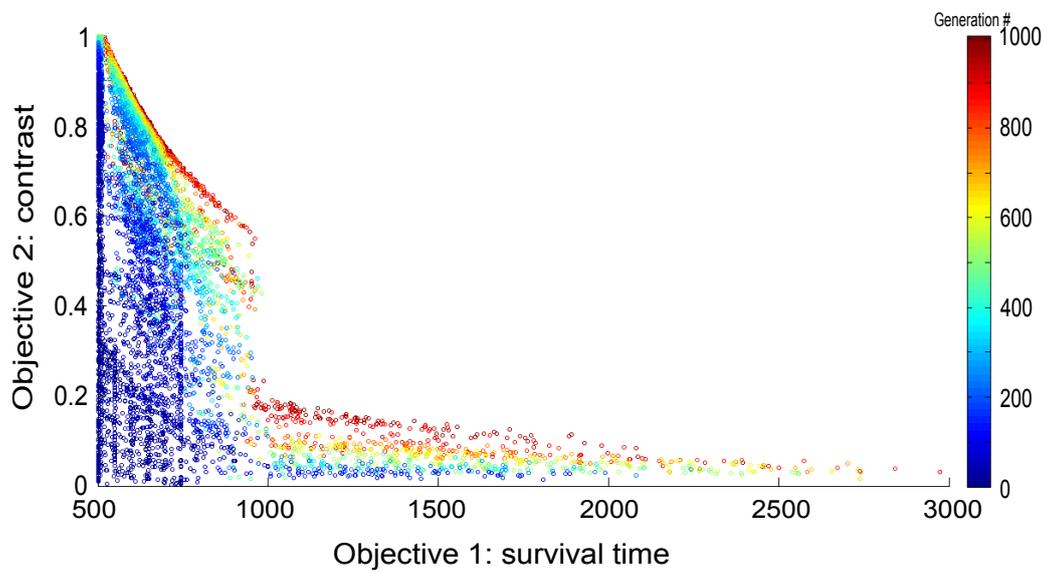


Figure 4.7: 2D Pareto front of the evolution of a mRF controller without salience representing the scores of the objectives of survival time (abscissa) and contrast (ordinate) obtained by each individual in each generation during the evolution. The network receives direct input from all external and internal variables, and the mRF controller was evaluated over five survival tasks.

Chapter 5

Discussions and perspectives

The nature of this modeling work of the mRF was primarily exploratory. Only two models existed and the number of articles on this area of the brain is quite low. As a result, we had to make a number of choices and focus our research on some points that seemed the most important to initiate such a modeling work. However, many areas of study remain to be explored, we will briefly mention them in this chapter.

To begin with, we have not detailed the parameters $P(p)$ and $P(l)$ of the network: we have seen in the first part of this report that literature suggests that these values are below 0.1. In our experiments, we put neither $P(p)$ nor $P(l)$ the anatomical constraints, however the structures of the best individuals we have obtained show that the average values of $P(p)$ and $P(l)$ both rotate around 0.09. This is consistent with the known anatomical data of the mRF, yet we have shown from the known anatomical data that $P(l) > 45 \times P(p)$, which is a relationship that we do not find in our results. Therefore, it would be interesting to add this constraint to see the performance of the evolved mRF.

We also left out the analysis and the impact of the neural population model chosen, i.e. the IPDS, which we described in chapter 2.1: first, we could let free the parameters we have fixed (τ to 5ms and *threshold* to 0). Second, there are other models of neural population that we could compare the efficiency with our model of the mRF with respect to the IPDS. Lastly, the networks that we developed are of sufficient size for modeling, especially as we have them evolve through evolutionary algorithms, with which it is best to handle networks of small size for reasons of computational speed and ease of interpretation: therefore, it would be interesting to quantify the impact of the maximum number of neurons that may contain a cluster.

About the embodied task, we compared our results indirectly with those of Humphries via comparison with the performance of random and WTA controllers. It would be interesting to directly compare our results with those of Humphries by implementing its model in our program.

In addition, the embodied task has a certain level of abstraction. A further assessment of the mRF would be to incorporate it as a controller of a known metabolic activity of an animal, such as eating behaviors and dipsiques mouse [Guillot, 1988], and compare the selections made by the mRF with the actual behavior of the animal.

It would also be interesting to analyze the neural networks obtained from the changes on one hand to better understand how they manage to solve a selection task, and on the other hand to try to extract the similarities between obtained networks. By the same token, we could study the properties of these networks, such as the presence of structures like small-world, scale-free or others.

On the neuro-evolution side, we made an extensive use of evolutionary algorithms without assessing their usefulness and their impact on the networks obtained after evolution. Among the various aspects that would be useful to further study, introducing new targets could yield interesting results, including:

1. a **generalization objective**, to ensure that decisions made by the mRF in the original task can be generalized to other similar tasks. Specifically, in the disembodied task it would mean ensuring that if we input a vector of the mRF not found in all input vector set that we used, for example (0.3, 0.5, 0.68) the mRF also gives the right output (0, 0, 1). In the disembodied task, the generalization may lead to assess the mRF in different maps than the one we used [Pinville et al., 2011].
2. a **diversity objective**, with which the evolution would check whether within each population individuals (i.e. mRFs) are not all alike. This objective would address the *genetic drift* (convergence to a local optimum), which is often found, like natural selection has formed sub-groups due to geographical constraints. We can find as many local optima as sub-groups, thereby achieving better results by getting closer to the global optimum [Mouret and Doncieux, 2009a, Mouret and Doncieux, 2009b, Doncieux and Mouret, 2009].

Another technique that we could use to try to improve the performance of evolutionary algorithms would be to perform evolutions in stages. Indeed, when we try to have the networks perform a difficult task, the search space is very important and evolution may experience difficulties to cross certain thresholds. For example, we have seen that the variant of the embodied task in which we gave to the mRF directly to internal

and external variables without calculating salience generated networks whose output vector contrast was very low, even after more than 1,000 generations. One possibility to help the evolution to further optimize the contrast could have been a first step to have the network learn the calculation of salience, and then to solve the survival task. As we showed in the disembodied task, mRF-type networks can act as a WTA, and a WTA controller can perform the survival task from the calculation of salience in a moderately effective (about 1,500 units of survival time, up to 3,000) but with a maximum contrast: it is possible that such an introduction of a temporary objective (learning calculation of salience) in the development allows to cross the thresholds more easily in the optimization goals life and contrast.

On the neurophysiological side, the literature gives a number of ideas to develop models for the mRF, in particular to better take into account the diversity of neuronal nuclei and divide the mRF on its three regions (midbrain, pontine, bulbar) and consider further the functions of the mRF. Various studies of the neurophysiological activity of the reticular formation (RF) represent a useful source of inspiration, especially the series of physiological studies on the RF produced by JM. Siegel in the 1970 and 1980. Thus,

- [Siegel and McGinty, 1977] found that neurons in the pontine reticular formation (PRF) had a high activity rate when high ocular activity was observed by electrooculography, which is consistent with previous studies. The authors also show the existence of a link between the discharges of neurons in the PRF and motor activity. The data found suggest a major role for PRF neurons in the regulation of motor activities. The study also gives an interesting indication on projections from the 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] shows the existence in cats of three classes of cells in the bulbar RF (*medullary reticular formation cells*), classification based on the correlation between their discharge and type of movements: a class discharges when the movement is laterally asymmetric (54% of the cells), the other class discharges when the movement is laterally symmetrical (38% of cells). The 8% of remaining cells correspond to the third class and are not related to motor behavior. There is no correlation between RF and bulbar eye movement, unlike the pontine RF.
- [Siegel et al., 1979] indicates the presence in cats of a correlation between the discharges of neurons in the medulla oblongata of the mRF (aka. *medial medullary reticular formation*) and certain motor activities during the phases of activity and during REM sleep. The study notes that the three cell types

mentioned above can also be found in the pons. It also gives an estimate of their proportions and their location.

In the long term, one of our goals is to connect our mRF model with existing models for the basal ganglia , which in turn will raise a number of important issues: how the networks interact, what is the precise role of each, and so on. As a historical sidenote, the original article introducing the first model of the mRF ended with a final chapter of cybernetic considerations which specifically referred to this kind of connections between brain areas and imagining their equivalents in a robotic controller [Kilmer et al., 1969].

Chapter 6

Conclusion

The objective of this study was to propose a new model of the mRF closer to the known anatomical data than the two previous existing models in the literature, and evaluate its ability to make the action selection. Our approach based on evolutionary algorithms has allowed us to show that a mRF-like network has the capacity to make decisions and strongly select them:

- the first series of experiments based on a disembodied task showed the computational capacity of the mRF to perform a selection task (chapitre 3);
- the second series of experiments based on an embodied task shows that the mRF is able to perform a task of action selection in simulated condition (chapitre 4).

The results we obtained are better than those achieved by Humphries' model in the two tasks we have discussed: refining our model by adding more neurons and removing the hypothesis of the Kilmer-McCulloch model taken up by Humphries that each cluster represents one action enabled us to improve the performance in terms of selection while better respecting known anatomical on the mRF.

However, the mRF-like structure does not appear to represent a particular advantage over a neural network without constraint. Therefore, to answer the original question, the mRF can be a substrate for action selection, but it does not appear that its structure represents an asset in particular.

In addition, the proof of $P(l) > 45 \times P(p)$ in chapter 1.4 virtually proves that mRF has a network structure of small-world type. This result can turn out to be very useful in future analysis.

Given the exploratory nature of this work to model the mRF, many areas of research that we have mentioned in the discussion remain to be further explored so as to refine the model and deepen the results. However, the results of this study are encouraging and besides their implications on the computational capabilities of the MRF, they show the potential contribution of evolutionary algorithms in computational neuroscience.

Bibliography

- [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. 5]
- [Berridge, 1989] Berridge, K. (1989). Progressive degradation of serial grooming chains by descending decerebration. *Behavioural brain research*, 33(3):241–253. [cited at p. 5]
- [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 Meyarivan, 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. 41]
- [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, 13, 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. 6]
- [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, 22, 28]
- [Humphries and Prescott, 2006] Humphries, M. and Prescott, T. (2006). Distributed action selection by a brainstem neural substrate: An embodied evaluation. *From Animals to Animats 9*, pages 199–210. [cited at p. 8, 12, 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, 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. 5]
- [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. 41]
- [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. 41]
- [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]
- [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. 42]
- [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. 42]

- [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. 5]

Appendices

Appendix A

Technical notes on the implementation

We would like to discuss the technical tools we used to carry out this project, because they represented on the one hand a very important work during the internship, and on the other hand it is always useful to have a concrete vision of technical tools used to find the results of the report, whether out of curiosity, to reproduce the experiments or even to find possible solutions to implement one's own models.

We also want to emphasize that the work carried out during the internship by no means has for sole purpose to produce this report but also to provide an opportunity for potential future successors to be quickly operational without needing to build their own tools from scratch. We thus join the initiative Plume from CNRS (<http://www.projet-plume.org>), whose purpose is to promote the open-source software (in French, the acronym PLUME stands for Promouvoir les Logiciels Utiles, Maîtrisés et Économiques) destined for the community of higher education and research.

To this end, particular attention was paid to the quality of development and all the code is available under the CeCILL license (<http://www.cecill.info/>) at the address <http://franck-dernoncourt.com/publications.php>. The table A.1 shows some statistics on the source code as well as scripts written to automate certain processes and analyze the results (contained in log files of the main program).

Neural networks and evolutionary algorithms have been developed in C++ for computational reasons, the evolutionary algorithms demanding very important computational resources, based on the framework

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: Statistics on the source code of the program and of the analysis scripts written in addition to the Sferes2 framework and existing libraries. The large size of C/C++ headers can be explained by the intensive use of templates.

Sferes2 (<http://pages.isir.upmc.fr/~mouret/sferes2/>) developed by ISIR [Mouret and Doncieux, 2010]. For the purposes of development, we ported Sferes2 to Windows. The program is multithreaded and is based on Boost libraries, including the Boost Graph Library for managing the graphs easily, TBB (Threading Building Blocks Intel) for multi-threading, Eigen2 for matrix calculations and SDL for graphics rendering the survival task.

Graphviz was used to generate graphical neural networks representing the mRF and the program fpmeg was used to generate videos from these graphs. MATLAB was chosen to analyze the results.

Lastly, we made an intensive use of the computer cluster of ISIR to run our program, the latter requiring high computing power. All programs used in this project run as well on Windows as on Linux.

The entire source code and scripts of analysis is available at <http://franck-dernoncourt.com/publications.php>

Mirror 1: http://pages.isir.upmc.fr/evorob_db/moin.wsgi/mRF2011.

Mirror 2: <http://bit.ly/mRF-xp>

Appendix B

Example of an mRF with 4 clusters

PTO.

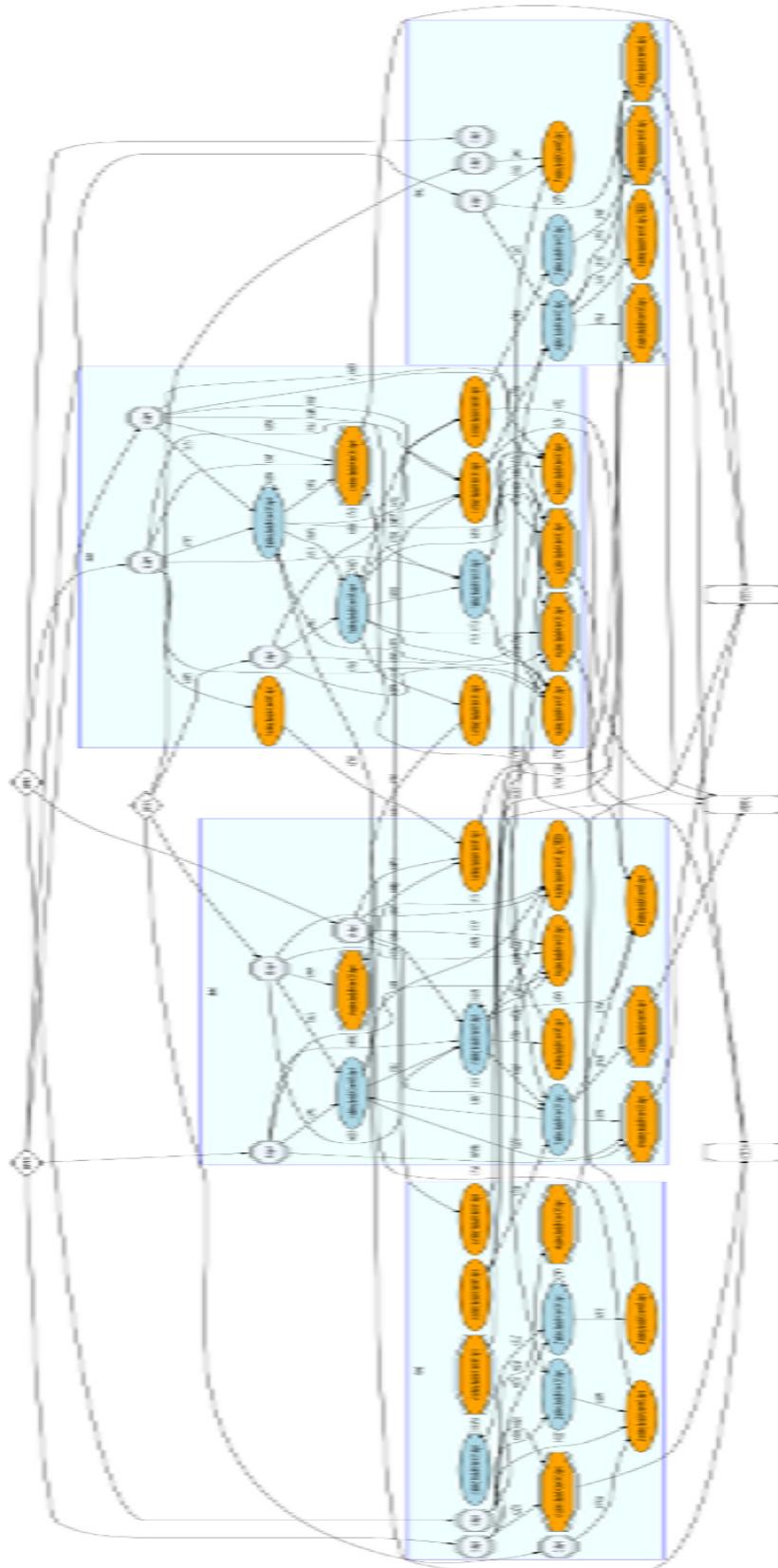


Figure B.1: Example of an mRF with 4 clusters. Excitatory neurons are orange, dark blue neurons are inhibitory. A cluster is a blue rectangle. Neurons located outside the blue rectangles represent inputs received by the mRF and the neurons to which it projects.

Appendix C

Example of a mRF cluster

PTO.

