



HAL
open science

Selection of action with a cortically-inspired model

Hervé Frezza-Buet, Frédéric Alexandre

► **To cite this version:**

Hervé Frezza-Buet, Frédéric Alexandre. Selection of action with a cortically-inspired model. Seventh European Workshop on Learning Robots, 1998, Edimbourg, pp.13-21. inria-00098709

HAL Id: inria-00098709

<https://inria.hal.science/inria-00098709v1>

Submitted on 26 Sep 2006

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

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

Selection of action with a cortically-inspired model

Hervé Frezza-Buet and Frédéric Alexandre*

**LORIA, 615 rue du Jardin Botanique, BP 101, F-54600 Villers-lès-Nancy, France.
email: Herve.Frezza@loria.fr, Frederic.Alexandre@loria.fr*

Abstract. We aim to show in this paper that a cortically-inspired control model applied to an autonomous robot can provide strong capabilities. First, such an approach enables a robot to extract spatio-temporal regularities during its “life” in an unknown environment, these regularities being stored in a more expressive way than the Q-values. Second, a compromise between perception and drives emerges from the distributed mechanism of spreading activities through the cortical net. Endly, the multi-modal management of information flows endows the robot with capabilities to generalise sequences of perceptions and use this knowledge to transfer learning from those learnt sequences to further behaviour.

Although our experiments with a simulated robot clearly exhibit these properties, they reveal a lack of planning capabilities that encourage us to extend our work toward a model of prefrontal cortex, grounded on the current results.

1 Introduction

The observation of animal or human behaviour is a classical way of inspiration to build models that allow a robot to behave autonomously in an unknown environment. This observation can be done at different levels of abstraction, from the psychological to the neurobiological level. Although we are also interested in the psychological domain, our work is mainly inspired with neurobiological data. Indeed, neurosciences are today able to give us many hints about the architecture, functioning and learning of the neural substratum that allows an animal to adapt to an environment. Accordingly, our approach is clearly based on a functional view of the cerebral system and more precisely of the cortex, main basis for an intelligent behaviour.

Despite ongoing works on robotic experiments, the work that we report here is the result of a computer simulation. That is the reason why, in this paper, we will speak about an animat and not (at the moment) about a robot.

We present below the main characteristics of our model, then we present its application to an autonomous behaviour task and conclude on a discus-

sion about the main features of our approach. We aim to show that a system that is strictly based on local computation, inspired from biological description of the cortex, is able to manage multimodal perceptions, temporal representations and the constraints of the environment. This paper then rather tends to exhibit the features of the model than to present a sharp control architecture.

2 The cortical model

The cortical model that we now present was first proposed from a biological point of view (Burnod, 1989). It was then adapted as a computational model (Alexandre et al., 1991) for perceptive and associative tasks. We are now studying it for the design of sensorimotor loops, including drive signals. Though its origins are clearly neurobiological, the computational model presented here must be understood as a basic functional view of the cortex and thus a very rough description of it.

2.1 The map level

The cortex has been described for a long time as a set of more or less overlapping functional maps (Brodmann, 1909). Whatever the exact definition of these areas, it is clear that the cortex can be partly described as a set of sensorimotor poles and a set of associative information flows, linking these poles and defining multimodal regions.

Implications in our model are twofold:

1. Different kinds of maps are designed, including sensory, motor and associative maps. Sensory maps receive information from the external world (e.g. visual information) or from an internal state of the animat (e.g. drives). They are mono-modal maps. On the contrary, associative maps are multi-modal (bi-modal in the work presented here). They combine two sensorimotor flows or two “lower” associative maps (Ballard, 1986).
2. Three different information flows can be defined in such a scheme. The feed-forward flow corresponds to the input signal, from sensorimotor poles toward associative regions. Conversely, the feed-back flow corresponds to the reverse direction and can be generally explained as a control flow. We will thirdly make an extensive use of the lateral flow, inside maps, that will allow inhibition and sequencing mechanisms, as described below. These flows are sketched in figure 1 below.

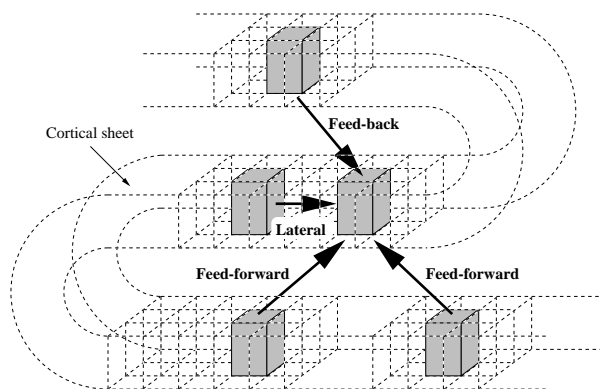


Figure 1: Three kinds of information flow

2.2 The maxicolumn level

Several models of cerebral functioning can be found in literature. Some of them are focused on

the smallest elements of the cortex, the neurons, and are applied to relatively elementary tasks. On the opposite side, models from symbolic artificial intelligence can be viewed as functional models of high cognitive capabilities. As we tend on the one hand to apply our model to a complete behavioral problem, and as we pretend on the other hand that the connectionist organisation of the cortex is a main cue to understand cognitive functions, we decided to ground the present model on the cortical maxicolumn. A maxicolumn can indeed have complex functioning rules, due to the high number of neurons it has, and is also a replicated elementary structure of the cortex.

The human cortex has been extensively described through its modular organisation (Mountcastle, 1978). As deeply described in (Alexandre et al., 1991), there are many different types of neurons with different connection pattern, inhibitory or excitatory mode, neurotransmitters, etc., but a regular structure, the cortical column, a functional set of more than one hundred neurons, can be observed throughout the cortex. This structure will thus be chosen as the basic unit of our cortical model. A cortical column is said to be a functional unit in that sense that it can specialise, with learning, on a specific tuning on its input information flow. For example, Hubel and Wiesel (Hubel and Wiesel, 1977) have described columns in the visual cortex, responding to a specific orientation of a visual stimulus.

Maxicolumns are sets of columns, having the same input, but performing different operations on it. Concerning the visual case evoked above, a maxicolumn is the set of columns responding on the same input flow (same location on the retina), but for different orientation selectivity. The set of maxicolumns whose information flow has the same modal origin (e.g. retina, cochlea, muscles) determines the cortical map.

The three information flows listed above are in fact the three kinds of links a column (and thus a maxicolumn) has. A column is tuned to answer to a specific distribution of an input information flow (feed-forward flow). This flow can be mono-modal. In this case, the column belongs to a sensory map and performs a filtering process on the input (Alexandre, 1996). The flow can be bi-modal. The column belongs to an associative maps. Its activation will correspond to a logical AND between its two inputs (Alexandre and Guyot, 1995). In both cases, the activation of a column will correspond to the detection (and thus the representation inside the cortex) of an event in the external world.

Similarly, activity of a column in a motor map will indicate the triggering of a motor event. In all these cases, detection of an event will lead to a specific activity of the column, the excitation activity.

As we will more deeply explain in the next section about learning, a column can have another activity, the call activity. We have shown above that a column can be seen as the representation of a specific event and will be excited if this event is occurring. It is thus possible to state that a column can be viewed as one of the possible goal the system can reach or ask for, which is the occurrence of the corresponding event (Burnod, 1989). For reasons rising from other parts of the cortex and detailed below, a goal can be desired, i.e. the excitation of the corresponding column can be asked for. In this case, the column will have a specific activity, the call activity, that will spread through feed-back and feed-forward links to other columns. This activity aims to trigger actions that might lead to the realisation of the desired goal.

Inhibitory mechanisms, carried out by lateral links will select, among a wide set of called columns, relevant calls. A column called with relevance and excited will be said to be satisfied: the event the cortex is asking for occurs.

Lateral links inside a map thus support inhibitory mechanisms. It is also the basis for one of the learning mechanisms that we now describe.

2.3 Learning

Three different kinds of learning can occur inside a column. First, a column has to learn its mask, i.e. the external event the column will represent. Classical non supervised learning, like Self Organisation Maps (Kohonen, 1989) or its derivatives (Martinetz and Schulten, 1991) can be successfully used for this kind of tasks (Alexandre, 1996). Such algorithms extract prototypes best representing the input information. Columns in sensory maps will be tuned to answer to these prototypes.

Concerning associative maps, we explained above that they basically perform a logical AND on their inputs. Another learning mechanisms can take place in these maps. Let Aa be an associative maxicolumn representing the co-activation of events A and a in two sensorimotor maps. Of course, other columns, representing coactivation of event A with other events exist in the associative map, and in some other associative maps. Now, suppose that event a is more often satisfied at the same time as associative maxicolumn Aa than any other maxicolumn associating a . In this case, A will be considered

to be a cause of the excitation of a and this association will be reinforced on feed-forward links of the associative column.

The third learning mechanism is a sequence learning mechanism, for sequences occurring inside a map. For example, suppose an animat is turning left; the object it is facing will move on the right on its retina and so will do its image on the visual map. It is thus interesting to learn this regularity of the external world.

The learning mechanism, using lateral links, is the following: If a maxicolumn a is “often” satisfied “after” the excitation of a maxicolumn b in the same map, b is assumed to be one of the possible events required for the occurrence of a . Then, calling b is useful when a is called. This utility of b will be realised by reinforcing the lateral link from a to b so that the call of a can generate a call of b . b is said to be a sub-goal of the goal a , noted $b \leftarrow a$. This simple mechanism can lead to the merging of two sequences, going through the same column. To prevent this undesirable phenomenon, we implemented a specialisation mechanism (also reported from a biological point of view (Burnod, 1989)), specialising a column in a maxicolumn, for each sequence, as represented in figure2.

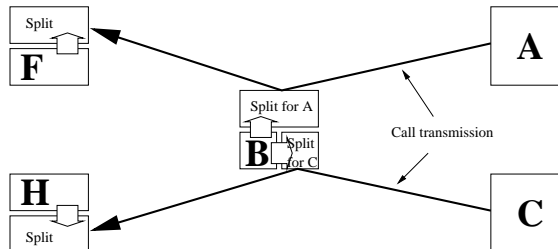


Figure 2: Because of the splitting mechanism, sequences $F \leftarrow B \leftarrow A$ and $H \leftarrow B \leftarrow C$ do not merge in B

From the latter example, the call propagation mechanism can be more easily illustrated: suppose that your animat has learned the regularity *if I turn left, the object I am facing is moving on the right* (in fact, it won't learn this rule, but columns detecting the presence of an object and columns representing movements will learn their causal relationships, in a distributed way), then, the animat can use this knowledge to reach a goal. Suppose that it wants (for any reason) an object to be on a specific location with regard to it. The column representing the presence of an object on this place will thus have a call activity. This state will propagate on the visual map until it reaches the actual location of the

object. Relevant call on motor maxicolumns will trigger a movement, yielding a new location, triggering a new movement, etc. until the object is at the correct location. Relevant calls are the result of a compromise between drives that initiates calls on corresponding sensory maps and current perceptions. The five rules discussed later in section 2.4 concerning the computation of relevant calls are designed in order to provide that choice.

Such phenomenon (and other more complex) will be reported in section 3 below. Before that, we have to be more precise on the computation of the learning parameters. Deeper studies about cortex-like learning rules can be found in (Sutton and Barto, 1981; Guigon, 1993).

All our learning rules are founded on correlation rules. Among these rules, hebbian rules are certainly the most famous, but they also have weaknesses (Sutton and Barto, 1981). To remedy some of these weaknesses, we used rules derived from Rescorla-Wagner theory for event contingency evaluation. A full discussion about that can be found in (Allan, 1993). Basically, strengths of associations V_i , for events e_i occurring at time t , with regard to event G , are computed as follows :

$$dV_i = \alpha_i(t)(\lambda_G - \sum_{e_j} V_j) \quad \text{if } e_i \text{ occurs at } t$$

$$dV_i = 0 \quad \text{otherwise}$$

where λ_G is a positive constant if event G occurs at time t and is null otherwise. This constant represents the total amount of association that can be distributed among events e_i and thus leads to a competitive view of learning.

To apply this rule to *temporal* sequence learning, three states **active**, **end** and **inactive** are defined for each column. A column is in state **active** if it had an excitation activity less than τ periods before (τ is a time constant); it is in state **end** if it was excited exactly τ periods before and in state **inactive** otherwise. During the active state of M_i , the efficiency α_i exponentially decreases, starting from a constant α . Our temporal adaptation of the Rescorla-Wagner learning rule, for a maxicolumn M_i is summarised in table 1 below.

This competitive temporal learning mechanism will lead to the reinforcement of the most frequent and closest in time predictor, that will be considered as a subgoal of G .

2.4 Functioning rules

Let us sum up the model we have presented so far. Its architecture can be described in terms of senso-

	States of M_i		
	active	end	inactive
Satisfac- tion of G oc- curs.	$dV_i = \alpha_i(t)(\lambda_G - \Sigma_G(t))$ Now G considers that M_i is inactive, until M_i be- comes excited again.		\emptyset
Rele- vant call in G .	\emptyset	$dV_i = -\beta\Sigma_G(t)$ The fact that M_i is not considered as inactive implies that M_i wasn't active the last time G was satisfied.	\emptyset
other- wise	\emptyset		

Table 1: Adapted Rescorla-Wagner learning rules

rimotor and associative maps. For one given application, it is built with regard to the perception and action capabilities of the animat. Three information flows link the maxicolumns belonging to these maps. Learning allow them to evaluate their spatial and temporal contingencies. External phenomena excite columns and, for a given task, excitation of a column can be desired. This will imply the spreading of a wide search (call) activity in maps. From this wide pre-activation, relevant calls must be selected.

Relevant calls are obtained by a competition between all called maxicolumns inside a map. We insist on the fact that this competition emerges from the local functioning of each maxicolumn and does not require any supervision. The competition is provided by the following mechanisms. Each column computes a competition value, and the maxicolumns whose competition value is lower than the maximum competition value of the maxicolumns in the map are inhibited, through lateral links. An inhibited maxicolumn only sends its calls through lateral links corresponding to learnt sequences. A cortical map in that sense can be viewed as a “vertical” filter of calls that allows only relevant calls to spread “vertically” toward other maps. We give hereafter the different criteria used to compute the competition value of a maxicolumn A , ordered from

the most relevant to the less significant.

1. The strength of call in A : this criterion enables the cortex to solve the most critical problems first.
2. The excitation of B , if the sequence $\dots \leftarrow B \leftarrow A \leftarrow \dots$ has been learnt. Thus, the maxicolumns that are called from learnt sequential paths and whose sub-goal is obtained are favoured. They are supposed to be the “next step” to the satisfaction of the drive.
3. The call that A receives through feed-back paths : this call is the result of the competition of an “upper” map.
4. Maximal excited pre-conditions of A : this value is computed if A is a successor of a goal G which have relevant call, (i.e. $\dots \leftarrow A \leftarrow \dots \leftarrow G \leftarrow \dots$). The value computed is the max of the excitation of the maxicolumns A receives feed-forward from. If A associates lower events α and β , we consider the excitation of one of them as a precondition to excitation of the other. Perceptive maxicolumns are considered as conditions producing a call on actions. This is thus not possible with the excitation of motor maxicolumns.
5. The relevant call that A receives, coming from feed-forward paths corresponding to the second learning mechanism described above : a maxicolumn from a “lower” map needs the occurrence of the other associated event to fire.

Endly, propagation of the relevant calls terminates in motor areas where the corresponding motor columns (and thus actions) are triggered. If learning was efficient, this action will allow to satisfy a sub-goal in the causal chain and, recursively, the original goal. This process is illustrated in the experiment described below.

3 Experiments

The features of the task we decided to assess our model on, have been defined within the framework of a project, partly funded by the *G.I.S Sciences de la cognition*, gathering French laboratories.

3.1 The framework

Our problem consists in enabling an animat to survive in a 2D environment, where it can find food and water : food area are red and water area are blue.

The animat perceives its external environment from a colour video camera, that sends the image to a retina. As our paper only deals with selection of actions, perceptive modules will not be described here. In our simulation, they are very simple representations of the environment. The actions the animat can perform are turn-left, turn right, go ahead, eat, drink. The animat survives as long as its internal levels of food and water have correct values.

We also wish to underline that no a priori knowledge is given to the animat. It will have to learn both the regularities of the environment (e.g. the colour of food) and the strategies to apply to satisfy a need (reaching a food area in order to eat).

3.2 Architecture of the animat

The animat has six perceptive kinds of inputs, and the corresponding six primary cortical maps. In a map receiving inputs of a particular modality, there is one maxicolumn for each discrete possible state of the modality. The different maps are :

- The *drives* map : this map includes two maxicolumns, each associated with one type of drive : need of food and need of water. The lower the internal food level is, the stronger is the call on the corresponding maxicolumn. The food maxicolumn is excited when the level of food increase. The water maxicolumn works in the same way.
- The *colours* map : this map includes several maxicolumns, one per colour that is taken account of. A relevant call on a maxicolumn causes an attention point to move on the retina to the closest patch of the corresponding colour. Then, the excitation of the maxicolumn is proportional to the intensity of the patch.
- The *focus-X* map : this map includes a row of maxicolumns, each corresponding to an abscissa of the attention point. A call on those columns has no direct effect on the environment, and excitation occurs on the appropriate abscissa-maxicolumn when the attention point succeeds in finding the patch asked by the *colours* map.
- The *focus-Y* map : this map is the same as the *focus-X* map, except that it manages the ordinate of the focalisation point.
- The *motion* map : this motor map contains three maxicolumns, each corresponding to a motion : turn left, turn right, go ahead. A

relevant call causes the motion to be executed. When the action is done, the corresponding maxicolumn is excited. The excitation of motor maxicolumns is not used to compute pre-condition competition criterion number 4 (see 2.4).

- The *ingestion* map : this motor map contains two maxicolumns : drinking and eating. A relevant call causes the corresponding action to be performed. The action succeeds (i.e. the maxicolumn is excited) whenever the animat is on suitable area or not, but the corresponding increase of internal level will happen only if the animat ingests on the correct area. As this map is a motor map, it has the same particularities as the *motion* map.

The aim of the animat’s cortex is to detect that all combinations of perceptions are not randomly distributed, that some of them are necessary to get others. Even if this basic architecture does not correspond to real biological systems, the coordination task the cortex will have to perform and the capacity required to perform the task have similar features with biological action decision problem.

The six primary cortical maps described above are connected through associative maps, and the global associative cortex is represented on figure 3, plain arrows representing feed-forward paths.

3.3 Experimental results

A set of experimental situations was proposed to the animat, for it to learn regularities of such environments and the consequences of its actions. The learnt sequences are shown on figure 3.

On the one hand, the second learning mechanism enables a call from “lower” maps to generate a call on the “upper” map containing the associative maxicolumn the lower maxicolumn needs. Results of this mechanism are shown by the black colour of the concerned associative maxicolumn. We can see that the animat has learnt association between each drive and each colour of the corresponding food and water. The colours are also associated with ingestions, and ingestions are consistently associated with drives. Both drives have been satisfied when the focalisation point was on a particular *XY* position. It is due to the fact that eating an drinking both satisfy the drives when the patch of the appropriate colour is centred on the bottom of the retina, i.e. the resource is near animat’s feet.

On the other hand, longer sub-goals sequences, resulting from the third learning mechanism, can

be found in visual and motion maps. They are represented with empty arrows : $a \leftarrow b$ meaning that a receives a call from b , because the occurrence of a is a sub-goal of the occurrence of b .

Now let us see what happens if the animat is hungry (see section 2.4). Initially, hunger causes a relevant call on the appropriate maxicolumn in the drives map. The effect of criterion number 5 makes the call on black maxicolumns in figure 3 be relevant, which immediately implies, due to criterion number 3, relevant calls on primary maps : a call for searching red patch on colour map and a call on eating action.

The focus position “middle-X, bottom-Y” is relevant-called in the same way. Figure 4 illustrates the following step by step description of the selection of the appropriate action. Note that the learnt sequences within Focus-X/Focus-Y map are not as numerous as they are in figure 3. Figure 4 is an intermediate stage of learning, useful for further descriptions. Let us suppose that the animat position is so that the red patch appears on the upper left corner of the retina (see 4-a). The upper-centre maxicolumn becomes now relevant, due to the criterion number 2 (see 4-b). As the upper-Y position is already obtained, call only goes down feed-back path to the middle-X maxicolumn of the focus-X map, which becomes relevant due to criterion number 3. This call propagates towards the three maxicolumns that associate “centre-X” with each motion. From the middle one, it goes through lateral sequential paths to the two connected maxicolumns (see 4-c). One of them, representing the association of “left-X” and “turn left”, will be relevant, due to criterion number 4, because left-X is excited (see 4-d). The effect of this relevance is a call on “turn left”, because of criterion number 3, and produces rotation of the animat to the left. The perceptive result is that the patch is now centred at the top. Criterion number 2 makes multicolumn “middle-X, middle-Y” be relevant (see 4-e). Perception on X is satisfied and call only propagates through Focus-Y map toward the map that associates motion and focus-Y, the mechanism being the same as the one we have just described for centering the patch (see 4-f). The effect will be the animat going ahead (see 4-g).

This mechanism shows how learnt sequences exploiting current perceptions can constraint the animat to first turn in order to centre the patch, and then go ahead until the patch arrives on the bottom of the retina. The most important thing is that sequences like those in *XY* map can be learnt only when looking for food (if the animat is never

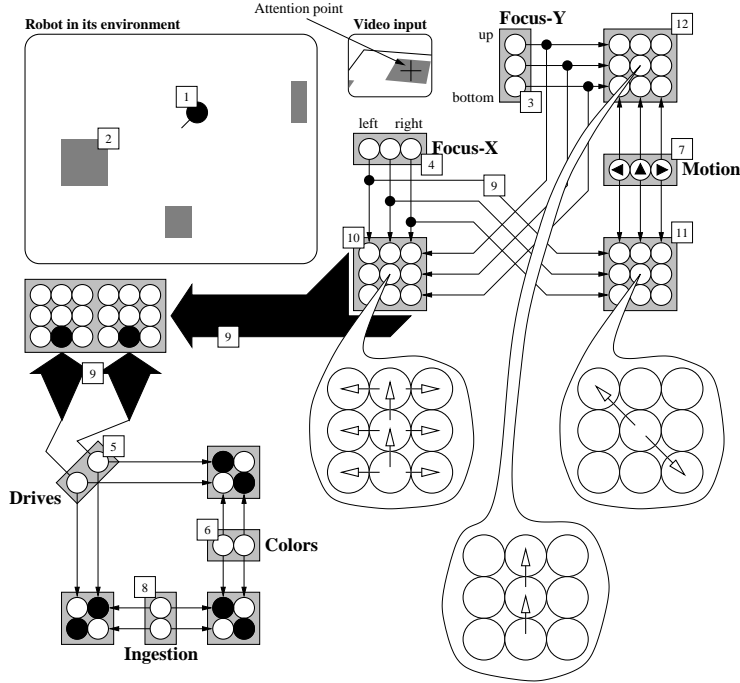


Figure 3: Associative cortex of the animat: (1) *the animat*, (2) *coloured areas*, (3-6) *sensory maps*, (7-8) *motor maps*, (9) *feed-forward streams*, (10) *associative visual map* that learns that the patch must be first centred and then translated toward the bottom, (11) *associative X-motion map* that learns to move to centre the patch, (12) *associative Y-motion map* that learns to move to make the patch translate toward the bottom.

thirsty), and then applied to learn to search a blue patch when needed. The sequence learnt within a map is a modal generalisation of all sequences of perception involving that sequence in that map. Due to the inhibitory mechanism of relevant calls, this generalisation ensures a capacity to transfer learning, i.e. to use modal generalities, learnt during past multimodal experiences, to constraint further learnings.

These properties can be illustrated from the intermediate learning stage shown in figure 4. The sequences learnt within the Focus-X/Focus-Y map only concern the centering and reaching of patches far from the animat. Let us suppose that the patch is now center-right on the retina. In this intermediate stage, the animat doesn't know that it has first to center it, because the appropriate sequence has not been learnt yet. The Focus-X/Focus-Y map can not elicit a relevant subgoal, and call spreads toward Focus-X and Focus-Y maps simultaneously (see 4-1). Nevertheless the animat, during centering and reaching far objects tasks did learn two things : when using the Focus-X/Motion map, it can center a patch regardless to its height on the retina, and when using the Focus-Y/Motion map, it can make a

patch be on the bottom of the retina, regardless to its horizontal position. These two capabilities have been implicitly generalized from previous reaching tasks. In the current case, as we see on figure 4-1, the spreading of calls will be constrained by these two learnt sequences, so that one of the two relevant called motions will be executed randomly. One of that motions, going ahead, will prevent the animat from seeing the patch going out of perception, and the sequence "left-X, middle-Y" \leftarrow "middle-X, middle-Y" will be learnt. This example shows how low-modal generalized sequences, even if they are not very reliable, can be used to constrain action and exhibit a correct behavior, until this behavior is learnt in "upper" appropriate map.

We tried to put random noise when the animat moves, adding a random value to the module and to the angle of the motion vector at each movement. The regularities learned in the maps allow the animat to keep direction (i.e. to keep the target centred), correcting continuously the effect of the noise.

The learnt behaviour appears to be reactive, directly exploiting current perception and the current dominant drive. We have added a reflex mechanism

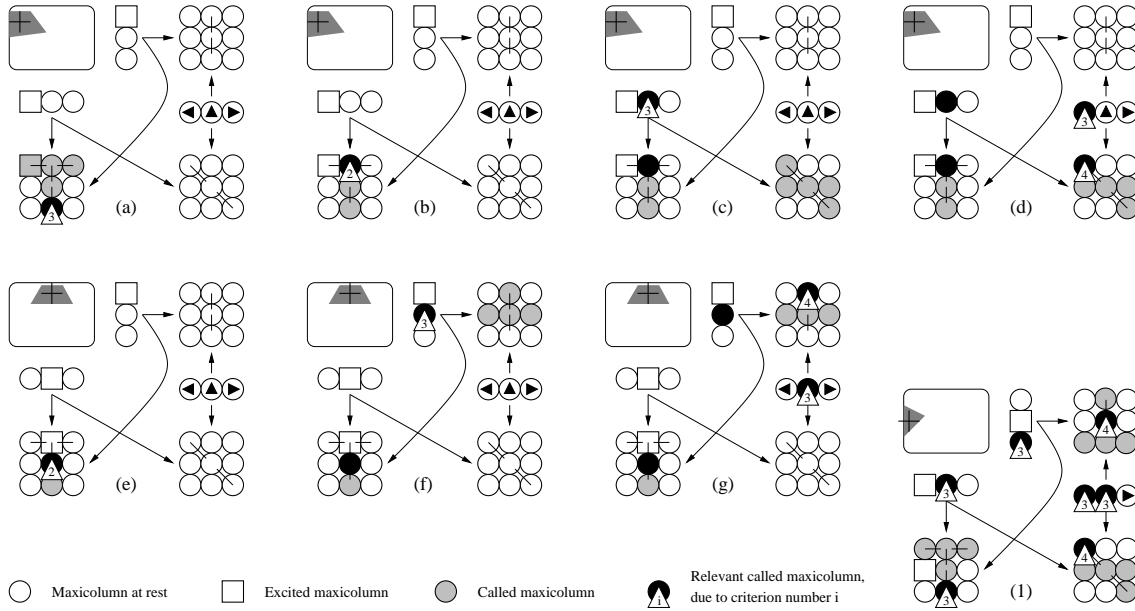


Figure 4: Selection of action according to current perceptions and needs.

consisting in turning left if no patch of the called colour is found in the video picture. This reflex acts as a picture exploring mechanism. The experiments shows that the animat can extract and exploit regularities of such environments, in order to reach an area corresponding to its current drive. Although the architecture ensures the learning and exploitation of contingencies of perceptions, the experiments clearly illustrate a lack of planning capabilities. For example, when thirsty, the animat can move near a food area without eating there, because eating isn't the strongest drive at that time. Our experiments leads our work toward the design of prefrontal cortex functionalities. A model of prefrontal cortex, grounded on the associative cortex presented here, will “think” before acting, i.e. will use associative regularities to virtually try different plans, in order to maximise a global satisfaction of drives and not necessarily the most urgent one.

4 Discussion

At the moment, experiments reported in section 3 have only been simulated in an artificial environment. Ongoing works correspond to the implementation of this application on a real robot Nomad200. Other works in our team are interested in adding to this model episodic memory and planification capabilities. This is respectively done with the integration of hippocampus and prefrontal cortex models.

As our approach is from many aspects close to other neural and modular models applied on robotics (Gaussier and Zrehen, 1995; Edelman, 1992), we wish here to insist on the specificities of our model. More precisely, the basic unit of this model is not the formal neuron, as usual, but a higher level unit, the cortical maxicolumn. At this level of granularity, such units can be regarded as functional units, representing the detection of an event, providing inside the network a sub-symbolic representation of information, easy to manipulate and to interpret. At a even higher level of abstraction, different kinds of spatial and temporal learning allow to learn, in a distributed way, regularities including causal relationships of the external world. This specific learning mechanism, grounded on the spread of activities and vertical filtering gives the model its power.

If we compare this approach with another reinforcement learning rule, as the famous Q-learning algorithm introduced by Watkins and discussed in (Jaakkola et al., 1994), the richness of the internal representations for our approach is obvious. The only internal representation of the external world in the Q-learning paradigm is the set of Q-values, coupling each action in each configuration of perceptions with the expected cumulative reward, from that configuration until the goal. Our approach aims to show that a cortically-inspired model can enable an animat to build internal representations of the regularities of its environment in order to

perform an efficient selection of action. As we described in the paper, these representations, more expressive than the Q-values, are based on competition and multimodality integration. They allow the animat to learn sequences of actions toward its satisfaction.

At last, another interest of this modular localist expressive approach is that it easily allows extensions, not only concerning local functioning rules, but also concerning other cortical maps, or even extra-cortical structures. Our ongoing works on the hippocampus and the prefrontal cortex exploits this convenience.

References

- Alexandre, F. (1996). Connectionist Cognitive Processing for Invariant Pattern Recognition. In *Proceedings International Conference on Pattern Recognition*, Vienna (Austria).
- Alexandre, F. and Guyot, F. (1995). Neurobiological Inspiration for the Architecture and Functioning of Cooperating Neural Networks. In *Proceedings International Workshop on Artificial Neural Networks*, Malaga (Spain).
- Alexandre, F., Guyot, F., Haton, J.-P., and Burnod, Y. (1991). The Cortical Column: A New Processing Unit for Multilayered Networks. *Neural Networks*, 4:15–25.
- Allan, L. G. (1993). Humain contingency judgments : Rule based or associative. *Psychological Bulletin*, 114(3):435–448.
- Ballard, D. H. (1986). Cortical connections and parallel processing : Structure and function. *The Behavioral and Brain Sciences*, (9):67–129.
- Brodmann, K. (1909). *Vergleichende Lokalisationstheorie der grobhirnrinde*. J.A. Barth, Leipzig.
- Burnod, Y. (1989). *An adaptative neural network : the cerebral cortex*. Masson.
- Edelman, G. (1992). Bright air, brilliant fire. on the matter of the mind. Basic books.
- Gaussier, P. and Zrehen, S. (1995). Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems journal*, 16(2-4):291–320.
- Guigon, E. (1993). *Modelisation des proprietes du cortex cerebral : Comparaison entre aires visuelles, motrices et préfrontales*. PhD thesis, École centrale de Paris. in english.
- Hubel, D. H. and Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Ferrier Lecture Proc. Roy. Soc. London*, pages 1–59.
- Jaakkola, T., Jordan, M. I., and Singh, S. P. (1994). On the convergence of stochastic iterative dynamic programming algorithms. *Neural Network*, 6:1185–1201.
- Kohonen, T. (1989). *Self-Organization and Associative Memory*, volume 8 of *Springer Series in Information Sciences*. Springer-Verlag.
- Martinetz, T. M. and Schulten, K. J. (1991). A "neural gas" network learns topologies. In Kohonen, T., Makisara, B., Simula, O., and Kangas, J., editors, *Artificial neural networks*. Elsevier science publishers.
- Mountcastle, V. B. (1978). An organizing principle for cerebral function. The unit module and the distributed system. In *The mindful brain*, Cambridge. MIT Press.
- Sutton, R. S. and Barto, A. G. (1981). Toward a modern theory of adaptative network : Expectation and prediction. *Psychological Review*, 88(1):137–170.