AN INTEGRATION OF TWO CONTROL ARCHITECTURES OF ACTION SELECTION AND NAVIGATION INSPIRED BY NEURAL CIRCUITS IN THE VERTEBRATES: THE BASAL GANGLIA

B. GIRARD*, D. FILLIAT**, J.-A. MEYER*, A. BERTHOZ*** AND A. GUILLOT*

*AnimatLab/LIP6, CNRS-University of Paris 6,
8 rue du Capitaine Scott,
75015 Paris, France
E-mail: {benoit.girard;jean-arcady.meyer;agnes.guillot}@lip6.fr

**DGA/Centre Technique d'Arcueil,
16 bis av. Prieur de Côte d'Or
94110 Arcueil Cedex, France
E-mail: david.filliat@etca.fr

***Laboratoire de Physiologie de la Perception et de l'Action,
CNRS-Collège de France,
11, place Marcelin Berthelot
75231 Paris Cedex 05, France
E-mail: alain.berthoz@college-de-france.fr

A control architecture of action selection, inspired by the neural loops of the dorsal part of the basal ganglia –subcortical nuclei of the vertebrate's brain– was proved to be able to solve a minimum survival task. The present paper concerns the connection of this architecture with a navigation system. This connection is inspired by recent hypotheses concerning the role of a ventral nucleus of the basal ganglia in integrating spatial, motivational and sensorimotor information. The ventral loop selects locomotion actions generated by various navigation strategies and modulated by motivations. The dorsal loop is in charge of non-spatial task selection and of coordination with the ventral loop. Implemented in a simulated robot performing the same survival task as in the previous experiment, the whole architecture improves the robot's survival thanks to map building and path planning abilities. Furthermore, the robot is also able to occasionally overlook the information recorded in its cognitive map in order to behave opportunistically, i.e. to reach an unexpected but visible resource, instead of a memorized but remote one. These results are discussed in terms of biological and robotic contributions.

2

1. Introduction

Contemporary robots are predominantly single-task systems, operating in specially-designed environments, in which they perform pre-programmed sequences of actions. While such systems have proved appropriate and useful on the factory floor, many future applications of robotics will require control systems with much greater flexibility. An effective strategy for designing flexible control systems is to reverse-engineer the biological control systems evolved in animals.

Among the essential functions for an autonomous system is the ability to decide which action should be done next to achieve its goals as well as the ability to orient itself in an environment in order to reach the various resources required by these goals [6]. Various computational models have proposed efficient mechanisms for these functions, respectively called action selection and navigation (see [16,3,14] for reviews). However, most action selection models use rudimentary navigation strategies while navigation models neglect the motivational choices which should guide the selection of their goals. To our knowledge, the integration of spatial information in a control architecture of action selection was not addressed as a specific issue.

Our objective is to design a bio-inspired architecture interfacing action selection and two navigation strategies of different complexity levels, taking inspiration from recent hypotheses about the role of the basal ganglia (BG) –a group of subcortical nuclei in the vertebrate's brain. These neural structures are assumed to be involved in action selection.

The BG are organised into segregated channels which receive inputs from the whole cortex and project to both brainstem nuclei and frontal parts of the cortex via thalamus nuclei, forming cortex-BG-thalamus-cortex loops. The selection occurs between the channels, a limited number of which being activated at any time. Each channel is supposed to be associated with one behaviour –a simple motor act like 'turn left' or an action pattern like 'feeding'. The BG select these channels by disinhibition [1]: the output nuclei of the basal ganglia are inhibitory and tonically active, therefore, in the absence of cortical input, they keep their targets –thalamus and brainstem nuclei– under constant inhibition. Specific patterns of activation in the cortex remove this inhibition in some channels and the related targets are consequently able to contribute to the behaviour.

There are three main cortex-BG-thalamus-cortex interconnected loops which are functionally distinct: motor (dorsal), associative (dorsal) and limbic (ventral). Low-level action selection using sensory-motor data is

supposed to be located in the motor loop and more cognitive processes are assumed to be managed by the associative loop while the integration of motivation and topologic navigation involves the limbic loop [18]. In this specific loop, the BG input nucleus –the nucleus accumbens (NAcc)– receives spatial information provided by the hippocampus (localisation [15]) and the prefrontal cortex (path planning [17]). However, the coding of this information in the NAcc core is still not clearly deciphered [11].

In a previous work, we have tested a computational model of action selection inspired by the dorsal loop, designed by Gurney *et al.* [7,8], in a survival task. It was implemented on a Lego Mindstorms robot which had to select efficient sequences of actions in order to find vital resources and it was able to solve this action selection problem [5]. However this robot followed the most rudimentary navigation strategy to find these resources: a random walk. To improve the survival of an artificial system in a complex environment, our objective is to add to this architecture a second circuit –simulating the ventral loop– which would select locomotor actions according to more complex navigation strategies: *object approach* and *topological navigation*.

After having described the survival task and the characteristics of the simulated robot, we will introduce the details of our computational model. The results of two experiments focusing on the contribution of navigation to action selection issues will be discussed in terms of biological and robotic advances.

2. Survival task

The model is evaluated in the same survival task as in [5]. The simulated robot has to maintain above 0 a constantly decreasing internal variable called *Energy*. To do so, it has to find "ingesting" places where it can reload its *Potential Energy*, then find "digesting" places to turn its *Potential Energy* into *Energy*. *Potential Energy* and *Energy* fluctuate between 0 and 1, and the consumption of a full charge of *Energy* without reloading allows the robot to survive only 16 min.

3. Environment and robot

The experiments are performed in 2D environments, where the resources are represented by $50cm \times 50cm$ squares of specific colours. The robot is circular ($30cm$ diameter) and equipped with an omnidirectional linear camera providing every 10° the colour of the nearest object. A resource is

4

considered available when it occupies more than 60° in the visual field (its centre is within a range of less than $70cm$ from the robot's centre).

The robot has three internal variables: *Energy*, *Potential Energy* and *Disorientation* (linked with navigation). For each type of resource –*Energy* and *Potential Energy*– three external variables (*Prox*, *maxProx* and *Avail*) are computed using the camera's image. *Prox* is a vector measuring the proximity of the resource in every direction (used for the *object approach*), *maxProx* is the maximum of this vector and *Avail* is a boolean which is true if the resource is available (*distance* $< 70cm$). The possible actions are: *Reloading Energy*, *Reloading Potential Energy* and *Moving in a given direction*.

4. Model

Both ventral and dorsal loops are simulated by the Gurney *et al.* model –named GPR according to the initials of the authors [7,8]. The GPR was also chosen to model the ventral loop because it is –anatomically and physiologically– very similar to the dorsal loop [12].
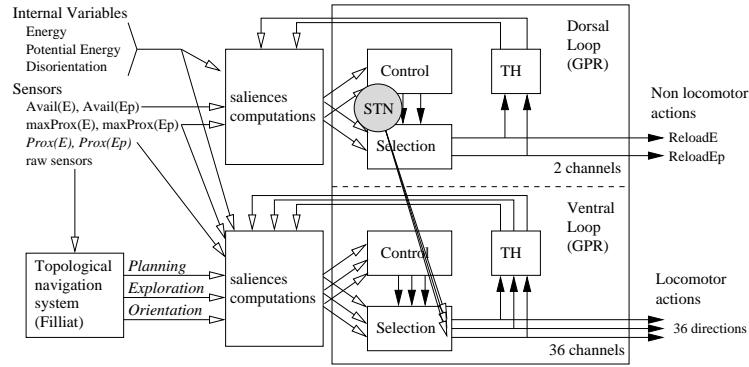

Figure 1.   Integration model of navigation and action selection, calling upon two basal ganglia loops. The STN of the dorsal loop, which provides the interconnection with the ventral loop, is highlighted. GPR: Gurney *et al.* model; black arrows: inhibitions; white arrows: excitations; italics: 36 value vectors.

Only the main characteristics of the GPR will be summarized here (for details, see [7,8]). The GPR identifies two sub-circuits in the basal ganglia (Fig. 1, GPR boxes): a *selection circuit*, which processes the cortical input and generates the inhibitory output of the basal ganglia and a *control circuit*, which regulates the activity of the selection circuit and insures that

inhibitions are always contrasted (disinhibition of the selected channel and strong inhibition of the others). The thalamo-cortical loop (TH) provides positive feedback to cortical inputs.

In the GPR, each channel is represented by one leaky-integrator neuron in every nucleus of the BG. The interconnections between nuclei are based on their neuro-anatomy, while the connection weights are fixed and derived from the connection densities. The *selection circuit* operates through two mechanisms: local recurrent inhibitory connections in the D1 striatum act as a "winner-takes-all" which disinhibits the winning channel, while the global "off-centre on-surround" structure of the network reinforces the inhibition of the losing channels and thus the discrepancies between the winning channel and the others. The *control circuit* is similar to the *selection circuit*, except that its output is directed to the *selection circuit*. This modulation of the *selection circuit* enhances its selectivity and allows for effective selection irrespective of the number of channels.

For every channel, the processed input is called *salience*, which represents the commitment towards displaying the action associated to that channel. It is computed by a sigma-pi neuron whose inputs are the external and internal variables as well as the feedback value provided by the thalamo-cortical loop, processed by transfer functions (linear, square and square root combinations). Because the system currently lacks any learning capacity, all the parameters of the transfer functions and the weights of the neurons are hand-tuned.

4.1. *Ventral loop*

Our main hypothesis is that the ventral loop selects directions of movement, given by simple sensor processing (for the *object approach* strategy) and by a navigation model (for the *topological navigation* strategy). The *object approach* generates movements induced by the proximity and direction of interesting in-sight resources. The *topological navigation* system (a modified version of the model of Filliat [2]) builds a topological map of the environment, using allothetic (vision and sonars) and idiothetic (self-movement) data. This map is made up of nodes representing locations in the environment. Links between nodes represent movements the robot has previously done between two locations. Within such a map, the robot's position is represented and calculated probabilistically according to algorithms inspired by hidden Markov models.

There are 36 channels, each of them coding for a 10° direction range.

6

Their saliences are influenced by vectors of 36 intensities computed by each strategy. The *object approach* is implemented by the two *Prox* vectors, in which directions leading to close resources are favoured. The *topological navigation* system produces three direction vectors: a *planning* vector, indicating the directions of the known resources, an *exploration* vector, favouring the directions leading to unmapped areas and an *orientation* vector, leading to well-known areas.

This *topological navigation* system puts some constraints on action selection because the robot is committed to regularly return to previoulsy mapped areas in order to ensure the building of a precise map. This need is expressed by a *Disorientation* variable produced by the system, which increases when the robot enters unexplored areas and decreases when it goes back in known areas. *Disorientation* is used as an additional internal variable in the computation of saliences.

At the output of the model, the selected direction is the result of a "soft-switching" process in which every channel disinhibited below a given threshold influences the final decision according to its level of disinhibition.

4.2. *Dorsal loop and loops interconnection*

The dorsal loop is kept to handle the non-locomotor behaviours. In our implementation, it corresponds to the reloading actions (reloading *Energy* and *Potential Energy*). As these two actions are not compatible, they are subject to "hard-switching": only the most disinhibited one is activated.

These reloading actions imply that the robot stops moving. Consequently, when the dorsal loop selects one of these actions, it has to be able to simultaneously influence the ventral loop in order to prevent locomotion. According to the so-called "trans-subthalamic pathway" hypothesis [10], this is modelled by excitatory projections from the part of the subthalamic nucleus (STN in Fig. 1) dedicated to one BG loop, to the output of the other loop. These additional excitations, generated by the selection of an action in the dorsal loop, and fed into the output nucleus of the ventral loop, increase its output inhibitions thus preventing the selection of any direction.

5. Experiments

Two experiments were carried out to respectively test the efficiency of the topological navigation integration and the capacity of the model to use its two navigation strategies in a situation requiring opportunistic decisions.

5.1. *Topological navigation efficiency*

In this experiment, the robot runs in the environment ($7m \times 9m$) depicted in Fig. 2 (left): one resource of each type is available, but it is impossible to see one resource from the vicinity of the other one. Consequently, a "reactive" robot following an *object approach* strategy only has to rely on random exploration to find hidden resources. On the contrary, after a first phase of random exploration and map building, a robot using a topological map should be able to reach desired resources.
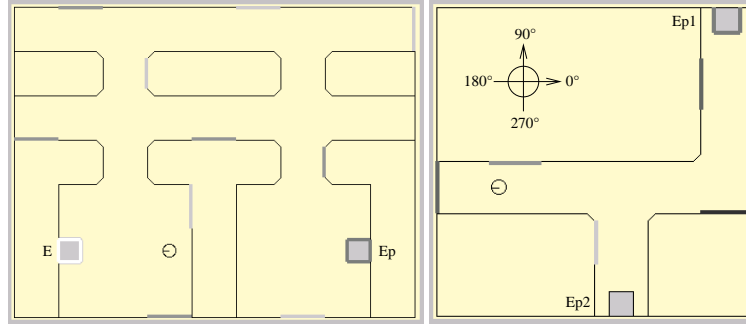
Figure 2.   Test environments used in experiment 1 (left) and experiment 2 (right). E: "digesting" resource; Ep: "ingesting" resource.

Ten tests, with a four-hour duration limit, were run for two simulated robots, one with topological navigation capacities (robot A) the other without (robot B). *Energy* and *Potential Energy* are initially set to 1. Table 1 (left) shows that robot A is able to survive significantly longer ($p < 0.01$, Mann-Whitney test) than robot B.

Table 1.   Experiment 1 (left): comparison (U Mann-Whitney test) of experiments duration medians of robots A and B. Experiment 2 (right): resources choice using three different configurations of weights for planning (Plan) and object approach (OA) vectors.

Durations (s)	Median	Range
Robot A	14431.5	2531:17274
Robot B	4908.0	2518:8831
U test	U = 15	$p < 0.01$

Weights		Selection	
Plan	**OA**	**Ep1**	**Ep2**
0.65	0.55	13	2
0.55	0.55	7	8
0.45	0.55	2	13

In [5], action selection was only constrained by the artificial metabolism. Here, the addition of the *topological navigation* system generates a new constraint: if the robot is too much disoriented, it has to come back to known

8

areas in order to build a precise and useful map. Despite this additional constraint, the integration of the *topological navigation* system into the action selection system proves to be efficient in enhancing the life span of robots.

However, during three robot A tests, the robot "died" because of a bad tuning of the salience computations: in situations characterized by a huge need of *Potential Energy*, when the robot approaches the *Potential Energy* resource, the dorsal loop increases, via STN, the output of all the channels of the ventral loop and stops the robot too far away from the desired resource. This shows the limitation of hand-tuning saliences and demonstrates the need for an adaptive learning mechanism capable of modifying the salience computations online.

5.2. *Coordination of the navigation strategies*

This experiment takes place in the environment ($6m \times 6m$) depicted in Fig. 2 (right), in which the robot may exhibit an opportunistic behaviour: when reaching the edge, it has to orientate towards an unexpected visible resource instead of following a planned path leading to a farther resource.

In a preliminary learning phase, the robot is allowed to build a topological map of this environment in which only the *Potential Energy* resource Ep1 is present, without metabolic constraints. In the test phase, a resource of the same nature, Ep2, is added to the environment and the initial *Potential Energy* level is set to 0.5. Following the planned path leading to Ep1, the robot begins to go to direction 0° but when it reaches the edge, it detects the new resource Ep2 and has to choose between two conflicting orientations: the planning vector leading to Ep1 (0°) or the object approach vector leading to Ep2 (270°).

Fifteen tests were run with various weights attributed to both planning and object approach vectors for salience computation. The results of Table 1 (right) show that, depending on which navigation strategy is favored, the robot is able to behave opportunistically thus choosing the closest resource (Ep2) or to ignore this unexpected resource and follow its initial plan (Ep1).

The indecision of the architecture, when both strategies have the same weight, is caused by the choice we made to merge orientations given by all the navigation strategies in the NAcc core and to select directions instead of selecting the strategies themselves. This alternative, suggested in a BG model [13], should be tested in a future implementation.

6. Discussion and Perspectives

We proposed an architecture of action selection integrating object approach and topological navigation strategies inspired by the structure and role of the basal ganglia in vertebrates. The advantage derived from the integration of the topological navigation strategy was demonstrated by an improved survival capability. The interest of using two different navigation strategies has been proven in a dedicated experiment.

From a neurobiological point of view, the model reaches the limits of current knowledge. Firstly, the coding of spatial information in the NAcc core being still undeciphered, a simplistic hypothesis of direction selection was made. The second experiment has suggested to try an alternative hypothesis of explicit navigation strategy selection. The link between these hypotheses and electrophysiological data remains to be considered. Secondly, the loop interconnection was achieved through the modelling of the trans-subthalamic pathway. However, the connections from one BG loop to another concern only a limited part of them, while, in our model, they concern the whole ventral loop output. Neurobiological data concerning the BG loops interconnections and their precise role is however currently too sparse to allow a more accurate model.

From the autonomous robotics point of view, this model proposes an integration of both action selection and navigation, taking into account the constraints of both issues. A preliminary experiment in a larger and more complex environment, including dangerous areas and in which various adaptive capacities were committed, showed that the robot can survive a very long time (up to 20 hours) [4]. The need to add learning abilities, as evoked in the first experiment, is taken into account in work currently in progress [9], which considers the role of dopamine as a reinforcement signal.

This model constitutes the first base of the control architecture of an "artificial rat", the *Psikharpax* robot. This integrative project aims at testing the coherency and the completeness of the current knowledge of the rat's nervous system and at building a robot which would be autonomous in the choice of its goals and actions.

References

[1] G. Chevalier and M. Deniau, Disinhibition as a basic process of striatal functions, *Trends Neurosci.*, **13**, 277 (1990).
[2] D. Filliat, Cartographie et estimation globale de la position pour un robot mobile autonome, *PhD. Thesis, Univ. Paris 6, France*, (2001).

10

[3] D. Filliat and J.-A. Meyer, Map-based navigation in mobile robots - I. A review of localization strategies, *J. Cog. Sys. Res.*, **4(4)**, 243 (2003).

[4] B. Girard, Intégration de la navigation et de la sélection de l'action dans une architecture de contrôle inspirée des ganglions de la base, *PhD. Thesis, Univ. Paris 6, France*, (2003).

[5] B. Girard, V. Cuzin, A. Guillot, K. N. Gurney and T. J. Prescott, Comparing a bio-inspired robot action selection mechanism with winner-takes-all, *SAB'02*, 75 (MIT Press, Cambridge, MA, 2002).

[6] A. Guillot and J.-A. Meyer, The Animat Contribution to Cognitive Systems Research, *J. Cog. Sys. Res.*, **2(2)**, 157 (2000).

[7] K. Gurney, T. J. Prescott and P. Redgrave, A computational model of action selection in the basal ganglia. I. A new functional anatomy, *Biol. Cyber.*, **84**, 401 (2001).

[8] K. Gurney, T. J. Prescott and P. Redgrave, A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour, *Biol. Cyber.*, **84**, 411 (2001).

[9] M. Khamassi, B. Girard, A. Guillot and A. Berthoz, Mécanismes neuromimétiques d'apprentissage par renforcement dans l'architecture de contrôle du rat artificiel Psikharpax, poster presented at *plate-forme AFIA*, (2003).

[10] B. P. Kolomiets, J.-M. Deniau, P. Mailly, A. Menetrey, J. Glowinski and A.-M. Thierry, Segregation and convergence of information flow through the cortico-subthalamic pathways, *J. Neurosci.*, **21(15)**, 5764 (2001).

[11] P. D. Martin and T. Ono, Effects of reward anticipation, reward presentation, and spatial parameters on the firing of single neurons recorded in subiculum and nucleus accumbens of freely moving rats, *Behav. Brain Res.*, **116**, 23 (2000).

[12] N. Maurice, J.-M. Deniau, J. Glowinski and A.-M. Thierry, Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the cortico-nigral circuits, *J. Neurosci.*, **19(11)**, 4674 (1999).

[13] O. Monchi, J. G. Taylor and A. Dagher, A neural model of working memory processes in normal subjects, Parkinson's disease and schizophrenia for fMRI design and prediction. *Neural Networks*, **13**, 953 (2000).

[14] J.-A. Meyer and D. Filliat, Map-based navigation in mobile robots - II. A review of map-learning and path-planning strategies, *J. Cog. Sys. Res.*, **4(4)**, 283 (2003).

[15] J. O'Keefe and L. Nadel, The Hippocampus as a Cognitive Map, *Oxford: Clarendon Press*, (1978).

[16] P. Pirjanian, Multiple Objective Action Selection Behavior Fusion using Voting, *PhD. Thesis, Aalborg Univ., Denmark*, (1998).

[17] J. K. Seamans, S. B. Floresco and A. G. Phillips, Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex, *Behav. Neurosci.*, **109**, 1063 (1995).

[18] J. K. Seamans and A. G. Phillips, Selective memory impairments produced by transient lidocaine-induced lesions of the nucleus accumbens in rats, *Behav. Neurosci.*, **108**, 456 (1994).