

State of the artificial rat Psikharpax

David Filliat*,*, Benoit Girard*,**, Agnès Guillot*, Mehdi Khamassi*,**,
Loïc Lachèze*, Jean-Arcady Meyer***

jean-arcady.meyer@lip6.fr

*AnimatLab/LIP6, 8 rue du capitaine Scott, 75015 Paris, France

** LPPA, Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France

*** DGA/Centre Technique d'Arcueil, 16 bis Av. Prieur de la Cote d'Or, 94114
Arcueil, France

Abstract

This paper describes the current state of advancement of the Psikharpax project, which aims at producing an artificial rat equipped with control architectures and mechanisms that reproduce as nearly as possible those that have been widely studied in the natural rat.

The article first describes the navigation system of Psikharpax, which is inspired from the anatomy and physiology of dedicated structures in the rat's brain, like the hippocampus and the postsubiculum. Then, it defines the animat's action-selection system, which aims at replicating other structures, the basal ganglia. It also explains how navigation and action-selection capacities have been combined thanks to the interconnection of two different loops in the basal ganglia: a ventral loop that selects the direction of motion, and a dorsal loop that selects other behaviors, like feeding or drinking. Finally, preliminary results on the implementation of learning mechanisms in these structures are also presented.

1. Introduction

As everybody knows, Psikharpax was the King of the Rats -- i.e., an intelligent and adaptive character -- in the *Batrachomyomachy*, a parody of *Iliad* written in Greek verses and (falsely) attributed to Homer. The name means "crumb robber". It has been given to a modern project that brings together several academic partners --

the Laboratoire de Physiologie de la Perception et de l'Action (LPPA), in Paris; the Laboratoire d'Informatique et de Microélectronique (LIRMM), in Montpellier; the Center for Neuromimetic Systems of the Swiss Federal Institute of Technology, in Lausanne; the Adaptive Behaviour Research Group, in Sheffield; and the AnimatLab, in Paris -- as well as two private companies -- BEV.S.A. and Wany.S.A. It aims at equipping an artificial rat with a bio-inspired behavioral control architecture that should hopefully afford it some of the capacities of autonomy and adaptation that characterize the natural rat. In particular, Psikharpax will be endowed with internal needs - such as hunger, rest, or curiosity - which it will try to satisfy in order to survive within the challenging environment of a laboratory populated with humans and, possibly, other robots. To this end, it will sense and act on its environment in pursuit of its own goals and in the service of its needs, without help or interpretation from outside the system.

Both fundamental and applied objectives are targeted by the project. In the first place, in the perspective of assessing the coherency and completeness of current knowledge about the rat's nervous system and about the mechanisms that contribute to its adaptive capacities, this project aims at integrating these mechanisms in a robot that may be confronted with the same situations as those a real rat may encounter in a laboratory or in nature. Secondly, this project should also help in assessing the operational value of these adaptive capacities in the situations where an artificial agent has to "survive" or fulfil its mission, without human assistance, and in a more or less unpredictable environment.

This paper is centred on the AnimatLab's contribution to this project and will describe the software developments that will contribute to endowing Psikharpax with integrated capacities for navigation, action selection and learning. It will capitalize on both simulations and robotic implementations. In the latter case, only commercial robots or simple home-made machines are concerned. However, the paper will end with a short presentation of parallel developments that will produce the final robotic platform on which the control software will ultimately be implemented.

2. General situation of the project

It is clear from recent reviews (Bar-Cohen and Breazeal, 2003; Holland and McFarland, 2001; Webb and Consi, 2001) that, although many research efforts have been devoted to the design of biomimetic sensors or effectors for robots, relatively little work has been done on control system architectures, and what has been done has focused primarily on invertebrate models. Only a few groups are currently building biomimetic robot control architectures modelled on mammalian nervous systems and, moreover, their efforts are often centred on isolated behaviors, like locomotion in cats (Patla et al., 1985) or feeding in mice (Guillot and Meyer, 1987), which are not dealt with in an integrated perspective.

To the best of our knowledge, the scope of the Psikharpax project is unique. First, it draws inspiration from a vertebrate instead of an invertebrate. Second, it aims at designing both biomimetic sensors and control architectures. Third, because it capitalizes on a dedicated robotic platform, it will integrate a variety of sensors, actuators and control systems making it possible to assess its adaptive capacities in much more challenging circumstances than those that characterize seemingly comparable biomimetic robotic approaches (Capi et al., 2002; Gaussier et al., 2000; Hafner et al., 2003; Montes-Gonzalez, 2000; Touretzky and Saksida, 1997) – a point to which we will return at the end of this paper.

3. Experimental results

To demonstrate the range of adaptive behaviors that a rodent-like control architecture affords a robot, we want to endow our artificial rat with the following capacities:

- to explore its environment and build a cognitive map of it;

- to use this map to localize both itself and places where rewards and punishments have been experienced;
- to learn which behaviors and objects in the environment generate which emotions and fulfil which motivations;
- to use its motivational system to select a current goal, like finding something to eat when it is hungry, or finding something to drink when it is thirsty;
- to switch to another behavior when specific emotions, like fear, are generated by specific events;
- to adjust its energy balance, notably by initiating rest periods.

Basically, this plan calls on two major control systems, respectively responsible for Psikharpax's navigation and action selection capacities. It also presupposes that these systems be integrated with various motivations and emotions in a coherent whole, and that it also exhibit appropriate learning mechanisms. The following paragraphs describe significant steps already made in these directions.

3.1. Navigation

Numerous simulation models – see Trullier et al. (1997) for a review – call upon so-called *place cells* and *head direction cells* to implement navigation systems that are inspired from the anatomy and physiology of dedicated structures in the rat's brain, like the *hippocampus* and the *postsubiculum*. The model currently used in Psikharpax's developments implements a *multiple-hypothesis tracking* navigation strategy, maintaining a set of hypotheses about the robot's position that are all updated in parallel (Filliat and Meyer, 2003; Meyer and Filliat, 2003).

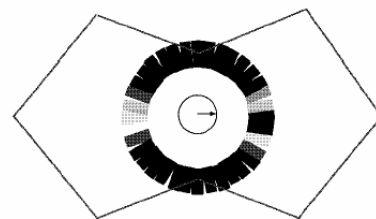


Figure 1. Schematics of allometric data used for navigation. The broken line joins the points detected by the robot's sonar sensors in eight absolute directions. The rectangles arranged in a circle indicate the mean grey-level perceived in the corresponding direction by the camera.

It serves to build a dense topological map (Filliat and Meyer, 2002), in which nodes store the allothetic data that the robot can perceive at the corresponding places in the environment. These data correspond to mean grey-levels perceived by a directional camera in each of 36 surrounding directions, and to sonar data providing distances to obstacles in eight directions (Figure 1).

A link between two nodes memorizes how far and in which direction the corresponding places are positioned relatively to each other, as measured by the robot's idiothetic sensors, i.e., by its odometry (Figure 2, left).

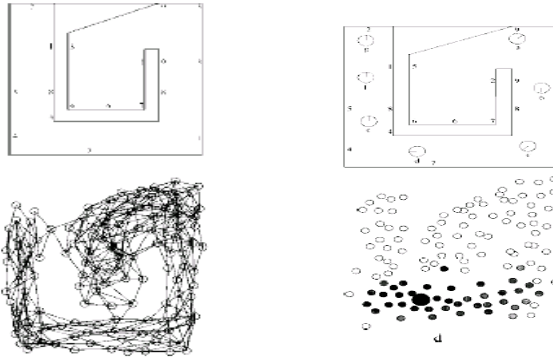


Figure 2. Left: The topological map (bottom) created by the robot when it explores an unknown environment (top). Right: This map may be used by the robot to localize itself because the activity distribution of the nodes in the map changes as the robot moves through successive places a, b,...g in the environment (top). Thus, when the robot is at place d, a blob of activity in the map surrounds the node that corresponds to this place (bottom). The grey level of each small node in the map indicates its activity, ranging from 0 for white nodes to 1 for black nodes. Larger black dots indicate the nodes which were successfully recognized.

The robot's position is represented by an activity distribution over the nodes, the activity level of a given node representing the probability that the robot is currently located at the corresponding position (Figure 2, right). This navigation model has been implemented on a Pioneer 2 mobile robot and proved to be efficient at exploring an unknown laboratory (Figure 3). From scratch, the robot succeeds in exploring its environment and in localizing itself accurately within it. Moreover, the system makes it possible for an external user to designate a goal-place in the map, from which a spreading activation algorithm computes the shortest path joining the goal to the robot's current position. When the robot follows this path, if it encounters an unexpected obstacle, the spreading activation algorithm may be triggered again to check if an alternative way to the goal exists. If so, the robot will follow the corresponding detour to fulfil its mission.

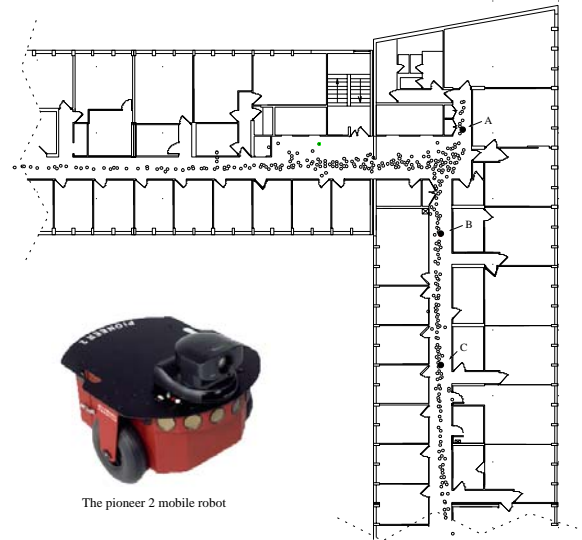


Figure 3. An example of a map created while a Pioneer 2 robot explored an unknown laboratory. The map is superimposed on an architectural sketch of the environment. When the robot is located in place A, if it is manually moved to place B, it will temporarily hesitate between two possible positions, B and C, that look very similar. However, after a few moves beyond place B, it will correctly identify its position.

3.2. Action selection

To survive, a rat must be able to solve the so-called *action selection problem* – i.e., it must be able to decide at every moment what to do next in service of its needs.

Some of the circuits that are involved in this task are known to be located in *basal ganglia-thalamus-cortex* loops and have inspired the *GPR model* (according to the authors Gurney, Prescott and Redgrave, 2001a & b) that is implemented in Psikharpax's control architecture (Girard et al., 2002; Girard, 2003). Basically, this model (Figure 4) assumes that the numerous segregated channels observed in the basal ganglia each correspond to a discrete motor action (the granularity of which is still not deciphered) that is inhibited by default and thus prevented from being executed. Inputs to these channels are so-called *salience*, a kind of *common currencies* for actions that take into account both internal and external perceptions to assess the relevance of each action with respect to the robot's needs. A positive feedback loop involving the thalamus serves to introduce some persistence in such assessments. Two parallel *selection* and *control* circuits within the basal ganglia act to modulate interactions between channels.

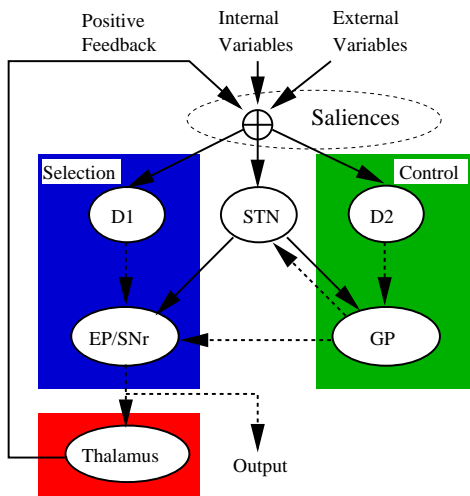


Figure 4. A single channel within the basal ganglia: the GPR model. D1 and D2 are striatal neurons with different dopamine receptors. STN is sub-thalamic nucleus. EP/SNr designates both the entopeduncular nucleus and the substantia nigra reticulata. GP is the globus pallidus. Solid arrows represent excitatory connections, dotted arrows represent inhibitory connections.

Finally, at the output of these circuits, the action that is less inhibited by others is selected and allowed to be executed by the motor system (Figure 5).

Experimental results demonstrate the model's ability to promote survival in the sense that it permanently keeps two *essential variables* (Ashby, 1952) above minimal levels: *Potential Energy* (obtained via “feeding”) and *Energy* (converted from Potential Energy via “resting”). Moreover, the model avoids dithering or interferences between actions - thanks to saliences' discrepancies enhanced by the control circuit - and demonstrates some advantages with respect to a simpler version in which the selected action would be the one with the greatest salience - thanks to the conjugated action of the control circuit and the thalamic loop.

This model has been implemented in a Lego robot whose task was to efficiently select between four actions – wandering, avoiding obstacles, “feeding” and “resting” – in order to “survive” in an environment where it could find “food” and “rest” places (Girard et al., 2003) (Figure 6).

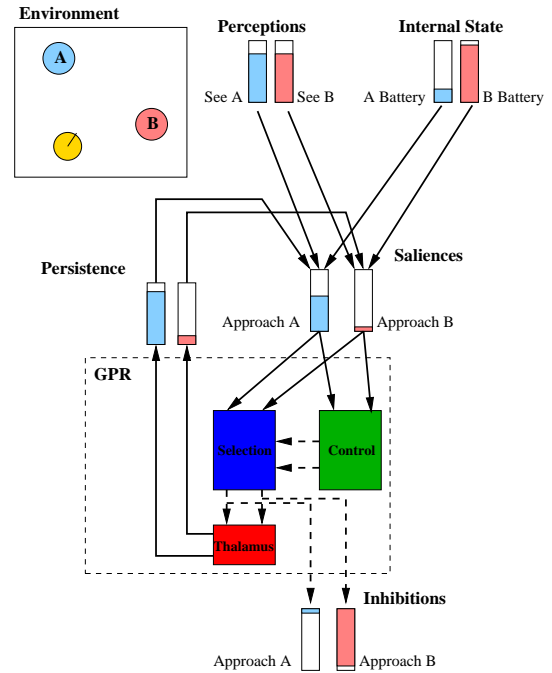


Figure 5. An illustration of the action selection circuit involving two possible actions: move-toward-A or move-toward-B. Solid arrows represent excitatory connections, dotted arrows represent inhibitory connections. The saliences associated with the two corresponding actions and channels depend upon both external and internal perceptions. In this case, although A and B objects are both perfectly perceived, because the A-battery level is lower than that of the B-battery, the action move-toward-A gets less inhibited than the other, and the robot accordingly decides to move towards the A object.

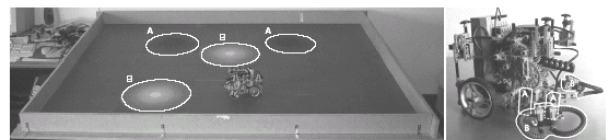


Figure 6. Left: The environment showing “food” (A) and “rest” (B) places. Right: A Lego robot equipped with light sensors (A) and bumpers (B).

However, the robot's survival depended on its chances of getting to the right place at the right moment, i.e., to a food place when its *Potential Energy* level was low, or to a rest place when it lacked *Energy*. Obviously, additional adaptive capacities would depend on the robot's capacity to record the position of such places on its map and to use this map to reach such places when needed. This has been made possible thanks to a model combining navigation and action selection capacities.

3.3. Combining navigation and action selection

The connection of the previously-described navigation and action selection models and their implementation on a simulated robot were inspired by recent hypotheses concerning the role of dedicated structures within the basal ganglia – like the *nucleus accumbens* in particular – and the interplay of *basal ganglia-thalamus-cortex* loops in the rat's brain (Girard, 2003; Girard et al., in press). The corresponding model (Figure 7) basically involves two such loops: a *ventral loop* that selects locomotor actions, like moving to the north or the east, and a *dorsal loop* that selects non-locomotor actions, like feeding or resting. Each of these loops has been modeled as a GPR system like the one previously described. The STN of the dorsal loop provides the interconnection between them because it sends excitatory projections to the output of the ventral loop. As a consequence, when the dorsal loop is active and triggers some non-locomotor action, the excitatory signal that is sent towards the ventral loop raises the inhibition level of every locomotor action and prevents it from being selected. Hence the robot cannot move and eat at the same time.

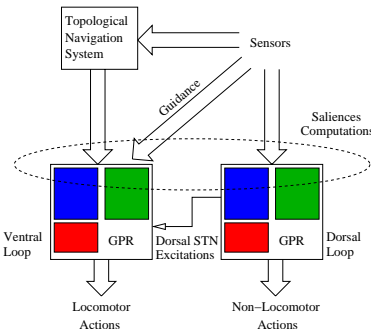


Figure 7. Interconnection of the ventral and dorsal loops in the basal ganglia. The ventral loop selects locomotor actions, the dorsal loop selects non locomotor actions. The latter subsumes the former.

Salience in the ventral loop depend upon direction profiles that are generated by two different navigation strategies, i.e., a simple *guidance* strategy and a more elaborate *topological navigation* strategy (Trullier et al., 1997). This makes it possible for the robot to be attracted either by an object that it directly perceives or to move towards a region where such an object is located in its map. The latter possibility puts some constraints on action selection because the robot is committed to regularly returning to previously mapped areas in its environment in order to check the accuracy of the current map. This need is expressed by a *Disorientation*

variable managed by the model, which increases when the robot enters unexplored areas, decreases when it returns to known areas, and affects the computation of saliences. Saliences in the dorsal loop depend upon both internal and external perceptions (Figure 8).

This model has been implemented in a version that manages 36 locomotor actions – i.e., moving in each of the 36 possible directions – and two non-locomotor actions – i.e., reloading actions that change the robot's *Energy* and *Potential Energy* levels.

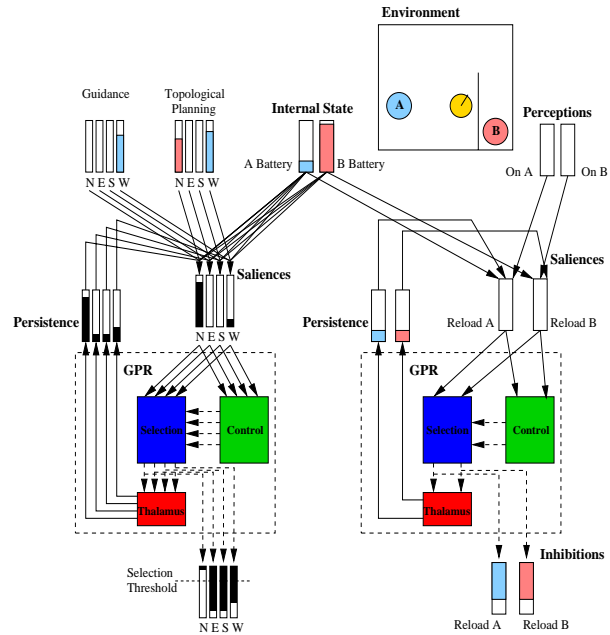


Figure 8. The model integrating navigation and action selection calls upon two basal ganglia-thalamus-cortex loops. Each loop is managed by a GPR model, and the coordination between loops is provided by the subthalamic nucleus of the dorsal loop, which is connected to the ventral loop (connection not shown here). The dorsal loop selects one of the two possible reloading actions, the ventral loop selects one of the 36 directions of movement (simplified here to four cardinal directions). Inhibitory connections are represented by dotted arrows, excitatory connections by solid arrows. In the current situation, only locomotor actions are selected because external perceptions are not strong enough to suppress the inhibition of reloading actions. When the robot gets close to the B object, this fact will be detected by its sensors, a reloading action will be triggered and an excitatory signal will be sent to the ventral loop in order to inhibit further locomotor actions.

The robot simulated in the environment on the left of Figure 9 survives successfully because it uses its map to navigate between places *E* and *Ep* where it can reload its

Energy and *Potential Energy* levels. Likewise, in the environment on the right of Figure 9, assuming that place *Ep1* is the only one that the robot has previously encountered and recorded on its map, if it decides to move towards that place to reload its *Potential Energy* and if it detects on its way the close presence of another food place like *Ep2*, it will give up navigating towards *Ep1* and will opportunistically divert to *Ep2*. Then, having consumed the corresponding “food”, it will record the position of *Ep2* on its map. Thus, next time it needs to reload its *Potential Energy*, it will have the choice of navigating towards *Ep1* or *Ep2*.

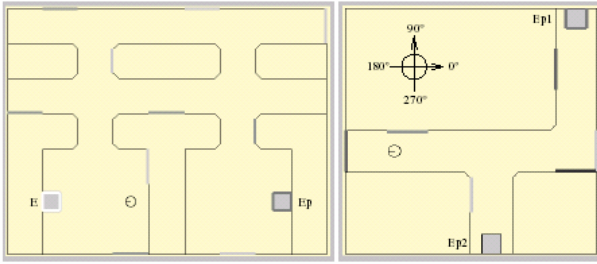


Figure 9. Two environments used to test the connection of navigation and action selection models. E: “rest” place, Ep: “food” place.

In the environment of Figure 10, the robot has the choice between two trajectories leading to a “food” place. The first one is shorter but entails passing through a “dangerous” place. The second one is longer, but safer. The robot is able to decide to navigate through the longer path when its *Potential Energy* level is not low enough to compromise its survival by a long journey, but it chooses the shorter path in the opposite case, at the risk of facing the potential danger recorded on its map.

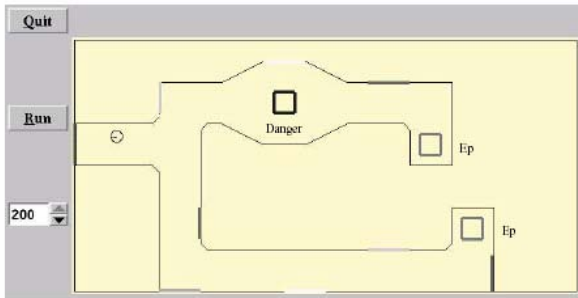


Figure 10. Two trajectories leading to “food” places (Ep). One is shorter than the other but entails passing through a dangerous place (Danger).

In the complex and challenging environment of Figure 11, the simulated robot autonomously survives, thanks to the numerous adaptive mechanisms and behaviors it has been endowed with (Girard, 2003).

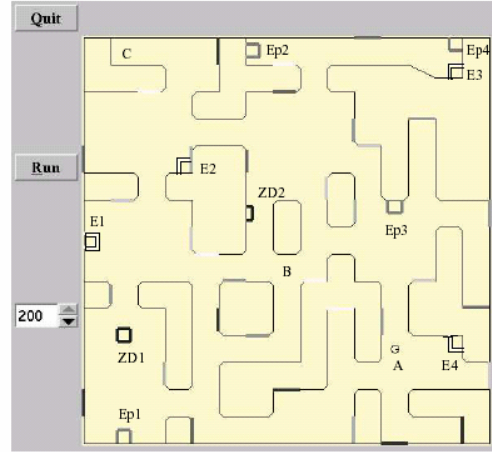


Figure 11. A complex environment with four “rest” places (E), four “food” places (Ep) and two dangerous places (ZD).

3.4. Learning

In an unknown environment, a rat is able to explore it and to incrementally build a map that describes the topology of this environment. Such *associative learning*, which combines both allothetic and idiothetic data, has been implemented in the navigation model described above.

However, a rat is also able to improve its behavior over time through *reinforcement learning*, i.e., thanks to adaptive mechanisms that raise its chances of exhibiting behaviors leading to rewards and that lower those of behaviors leading to punishments. Concerning action selection, a recently debated hypothesis (Barto, 1995; Houk et al., 1995; Schultz et al., 1997) postulates that such mechanisms could be mediated by dopamine signals within so-called *actor-critic* architectures (Figure 12).

Within such architectures, an action-selection module plays the role of an actor, while a critic module calls upon both the episodic reinforcement signal r_t occasionally generated by the robot’s actions and a dopamine signal that is assumed to evaluate the difference $gP_t - P_{t-1}$ between currently expected and future

rewards. This estimate is used in the actor module to adapt the way saliences are computed and used to select the most appropriate action, i.e., the action the most likely to maximize the reward that it will generate.

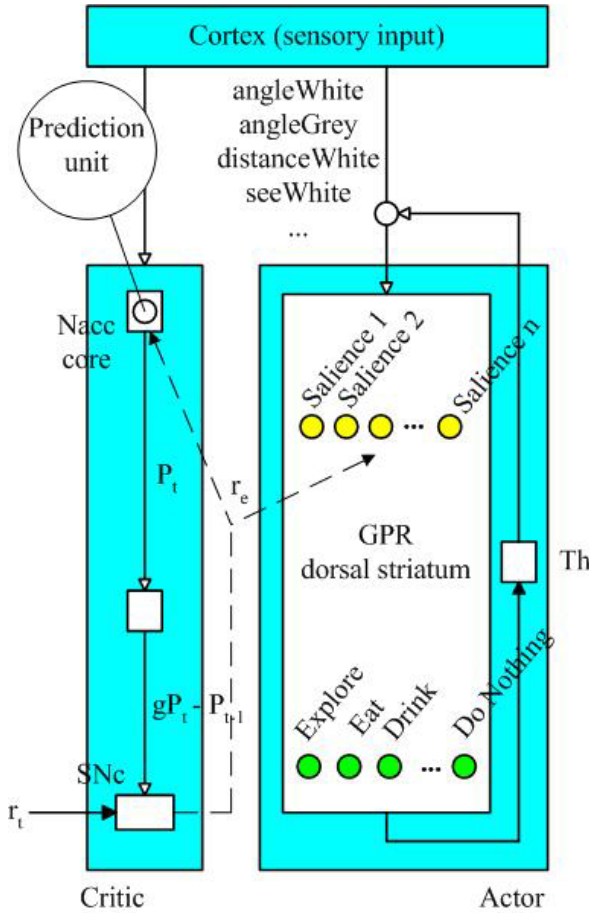


Figure 12. The actor-critic model of reinforcement learning. The actor module is a GPR model that is segregated in different channels, with saliences as inputs and actions as outputs. The critic module (involving the nucleus accumbens core (Nacc core) and the substantia nigra compacta (SNc)) propagates towards the actor module an estimate r_e of the instantaneous reinforcement triggered by the selected action.

This type of model has been implemented in a simulated robot that must learn in a plus-maze, and through successive trials, which movement to perform in order to get to the end of an arm where a door may provide access to a reward - i.e., some water to drink (Figure 13). This kind of learning corresponds to a simple Stimulus-Response association. At every trial, one lamp out of four is lighted indicating the door behind which the

reward is accessible. When the robot succeeds in getting such reward, the corresponding lamp is turned off and the robot must learn to return to the center of the maze, where the lighting of another lamp will designate another reward place. This setting reproduces an experiment on real rats (Albertin et al., 2000) and helps to interpret the corresponding results.

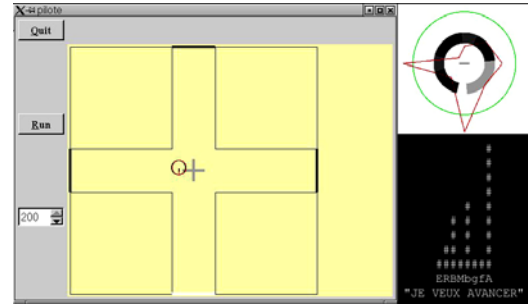


Figure 13. Left: the robot in the plus maze environment. A lighted lamp indicates which door (in white) leads to reward. The other doors do not lead to reward and are shown in black. Upper right: the robot's visual perceptions. Lower right: activation level of each channel in the actor module.

Two different critic modules, respectively adapted from Houk, Adams and Barto (1995) and from Baldassare (2002), have been implemented and connected to the same actor module, i.e., the action selection model described in Section III.2 above (Khamassi et al., in press). The main difference between these modules is that the first module (Critic 1) calls upon only one unit to predict an instantaneous reward, while the second (Critic 2) calls upon two units, on the one hand, and takes into account the time course of the rewarding stimulus, on the other hand.

The corresponding results were compared on the basis of two criteria: the zone in the environment where the robot has already learned something - i.e., where it is able to select the appropriate actions to get to the reward - and the way the prediction errors decrease at reward location along successive trials - information that is essential to propagate learning to other regions of the environment. It thus turns out that Critic 2 gives better results than Critic 1 (Figure 14) but that the corresponding module still has to be improved, for instance through the use of a greater number of prediction units, or the use of one controller per critic unit (Doya et al., 2000), to let learning extend to the whole experimental environment and to speed up the corresponding process.

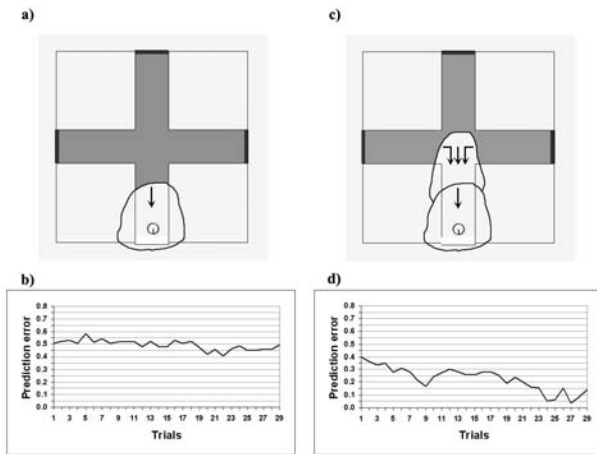


Figure 14. Reinforcement learning results obtained with an actor-critic architecture in which two different critic modules have been tested (left: Critic 1; right: Critic 2). Depending on which module is used, the correct actions to perform are learned in specific regions of the maze (a & c: white areas in the plus maze) and the improvement patterns of the prediction error at reward location over successive trials are different (b & d).

Referring to the beginning of Section 3 above, it thus appears that almost all the capacities that we wanted to endow our artificial rat with have started to be implemented, even if this animat currently exhibits very few motivations and emotions. In 2004, these capacities will be integrated on a Peeke platform that will make a first assessment of the Psikharpax endeavour possible.

4. The future robot

A more advanced version of Psikharpax is under construction. The morphology and sensori-motor equipment of this 50cm-long robot will be as closely inspired from the real rat as possible (Figure 15).

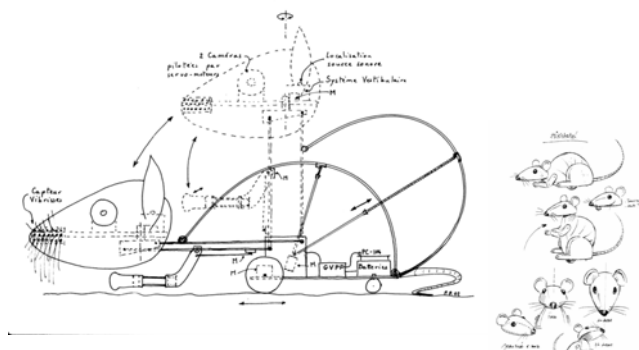


Figure 15. The overall design of the future Psikharpax.

In particular, it will be endowed with three sets of allothetic sensors: a two-eyed visual system, an auditory system calling upon two electronic cochleas, and a haptic system made of 50 whiskers on each side of its head.

Psikharpax will also be equipped with three sets of idiothetic sensors: a vestibular system reacting to linear and angular accelerations of its head, an odometry system monitoring the length of its displacements, and capacities to assess its current energy level. Sensor fusion will be accomplished through the use of GVPP, a biomimetic chip dedicated to low-level real-time signal processing that already serves robot vision (Gourichon et al., 2002).

The robot's capacity to move and act in the environment will be afforded by several motors and actuators. In particular - despite the fact that such a device is not really biomimetic - two wheels will allow Psikharpax to move at a maximum speed of a few meters per second. Although it will usually lie flat on the ground, it will also have the possibility of setting upright, as well as of seizing objects with two forelegs. Likewise, its head will be able to rotate, and three pairs of motors will actuate each of its eyes (Figure 16).



Figure 16. An eye equipped with a camera and a log-polar sensor, which is actuated by three motors. The whole device obeys the Listing's law (von Helmholtz, 1955).

Finally, several low-level reflexes will connect Psikharpax's sensors to its actuators, thus making it possible, for instance, to keep looking at an object even when its head is moving, and to avoid an obstacle detected by its whiskers or by its visual or auditory systems.

5. Directions of future work

Being able to integrate the past (through its recorded map), the present (through its sensors) and the future (through its planning capacities), Psikharpax is a nice example of a *motivationally autonomous animat* (McFarland and Bösner, 1993) whose survival may depend on a huge number of perceptions and actions that must be carefully integrated. Future work will be devoted to increasing this number as much as possible, in order to delineate the limits of the controllers described above, an attitude that is at the opposite of that of traditional robotics, when the chances of threatening the survival of a costly device are deliberately held as low as possible. Clearly, the control of such a complex sensori-motor system is a great challenge that goes beyond any other attempt at making an animat do something more than merely moving around. To successfully tackle this challenge will entail extending the capacities of the controllers described herein by letting them manage more perceptions -- notably a sense of smell -- more actions, more motivations and more emotions. It will also probably involve adding other biomimetic control structures, among which the cerebellum and the prefrontal cortex will probably prove to be of utmost importance, as well as other learning processes than the S-R one currently implemented, notably those that would afford capacities for goal-directed behaviors and behavioral sequence chunking (Dayan, 2001; Graybiel, 1998).

6. Conclusions

Considerable research effort has been already devoted to the Psikharpax project and the corresponding results have been summarized in this paper. In particular, several adaptive mechanisms have been designed that make navigation, action selection and learning possible. When a dedicated version of the Psikharpax platform becomes available, the adaptive mechanisms that have been studied so far will be implemented and integrated on it. This will probably generate new and interesting control challenges, from which useful contributions to biology, cognitive sciences and robotics are to be expected.

Acknowledgments

This research has been granted by the LIP6 and the Project *Robotique et Entités Artificielles* (ROBEA) of

the Centre National de la Recherche Scientifique, France.

References

- Albertin, S. V., Mulder, A. B., Tabuchi, E., Zugaro, M. B. & Wiener, S. I. (2000). Lesions of the Medial Shell of the Nucleus Accumbens Impair Rats in Finding Larger Rewards, but Spare Reward-Seeking Behavior. *Behavioral Brain Research*. 117(1-2):173-83.
- Ashby, W. (1952). *Design for a Brain*. Chapman and Hall.
- Baldassarre, G. (2002). A Modular Neural-Network Model of the Basal Ganglia's Role in Learning and Selecting Motor Behaviours. *Journal of Cognitive Systems Research*, 3(1):5-13.
- Bar-Cohen, Y. and Breazeal, C. (2003) (Eds). *Biologically Inspired Intelligent Robots*. SPIE Press.
- Barto, A.G. (1995). Adaptive Critics and the Basal Ganglia. In Houk et al., (Eds). *Models of Information Processing in the Basal Ganglia*, pp 215-232. The MIT Press.
- Capi G., Uchibe E., Doya K. (2002). Selection of Neural Architecture and the Environment Complexity. *The 3rd International Symposium on Human and Artificial Intelligence Systems: Dynamic Systems Approach for Embodiment and Sociality*, pp 231-237. Fukui, Japan.
- Dayan, P. (2001). Motivated Reinforcement Learning. *NIPS*, 14: 11-18. The MIT Press.
- Doya, K.; Samejima, K.; Katagiri, K.; Kawato, M. (2002) Multiple Model-based Reinforcement Learning. *Neural Computation*. 14, 6: 1347-69
- Filliat D. and Meyer, J.-A. (2003). Map-based Navigation in Mobile Robots - I. A review of localization strategies. *Journal of Cognitive Systems Research*. 4,4, 243-282.
- Filliat, D. and Meyer, J.-A. (2002). Global Localization and Topological Map Learning for Robot Navigation. In Hallam et al. (Eds). *From Animals to Animats 7. Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior*. pp 131-140. The MIT Press.
- Gaussier, P., Leprêtre, S., Quoy, M., Revel, A., Joulain, C., Banquet, J.P. (2000). Experiments and Models about Cognitive Map Learning for Motivated Navigation. In Demiris and Birk, (Eds). *Interdisciplinary approaches to robot learning. Robotics and Intelligent Systems*. 24 : 53-94. World Scientific.
- Girard, B. (2003) *Intégration de la navigation et de la sélection de l'action dans une architecture de contrôle*

- inspirée des ganglions de la base*. PhD. Thesis, University of Paris 6, France.
- Girard, B., Cuzin, V., Guillot, A., Gurney, K., Prescott, T. (2002). Comparing a Bio-inspired Robot Action Selection Mechanism with Winner-takes-all. In Hallam et al. (Eds). *From Animals to Animats 7. Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior* pp. 75-84. The MIT Press.
- 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(22), 179-200.
- Girard, B., Filliat, D., Meyer, J.-A., Berthoz A., and Guillot, A. (In Press). An Integration of two Control Architectures of Action Selection and Navigation inspired by Neural Circuits in the Vertebrates: the Basal Ganglia. *Proceedings of the 8th Neural Computation and Psychology Workshop. Progress in Neural Processing*. World Scientific.
- Gourichon, S., Meyer, J.-A. and Pirim, P. (2002). Using colored snapshots for short-range guidance in mobile robots. *International Journal of Robotics and Automation: Special Issue on Biologically Inspired Robotics*. 17, 4:154-162.
- Graybiel, A. M. (1998). The Basal Ganglia and Chunking of Action Repertoires. *Neurobiology of Learning and Memory*, 70:119-136.
- Gurney, K.N., Prescott, T.J. and Redgrave, P. (2001a). A Computational Model of Action Selection in the Basal Ganglia. I. A new functional anatomy. *Biological Cybernetics*. 84, 401-410.
- Gurney, K.N., Prescott, T.J. and Redgrave, P. (2001b). A Computational Model of Action Selection in the Basal Ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*. 84, 411-423.
- Hafner, V. V., Fend, M., Lungarella, M., Pfeifer, R., König, P., Körding, K. P. (2003), Optimal coding for naturally occurring whisker deflections. *Proceedings of the Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP)*. pp. 805-812. Springer Lecture Notes in Computer Science,.
- von Helmholtz, H. (1955). *Treatise on Physiological Optics*. The Thoemmes Press.
- Holland, O. and McFarland, D. (2001). *Artificial Ethology*. Oxford University Press.
- Houk, J.C., Adams, J.L. and Barto, A.G. (1995). A Model of how the Basal Ganglia generate and Use Neural Signals That Predict Reinforcement. In Houk et al., (Eds). *Models of Information Processing in the Basal Ganglia*. The MIT Press.
- Khamassi, M., Girard, B., Berthoz, A. and Guillot, A. (In Press). Comparing three Critic Models of Reinforcement Learning in the Basal Ganglia Connected to a Detailed Actor in a S-R Task. *Proceedings of the 8th Intelligent Autonomous Systems Conference*. IOS Press.
- McFarland, D.J. & Bösser, T. (1993) *Intelligent Behavior in Animals and Robots*. Cambridge: The MIT Press.
- Meyer, J.-A. and Filliat D. (2003). Map-based navigation in mobile robots - II. A review of map-learning and path-planning strategies. *Journal of Cognitive Systems Research*. 4, 4: 283-317.
- Montes Gonzales, F., Prescott, T.J., Gurney, K., Humphries, M. and Redgrave, P. (2000). An Embodied Model of Action Selection Mechanisms in the Vertebrate Brain. In Meyer et al. (Eds). *From Animals to Animats 6. Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior* pp. 157-166. The MIT Press.
- Schultz, W., Dayan, P. & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275:1593-1599.
- Touretzky, D. and Saksida. L. (1997). Operant Conditioning in Skinnerbots. *Adaptive Behavior*, 5, 3/4 :219-247.
- Trullier, O., Wiener, S., Berthoz, A. and Meyer, J.A. (1997). Biologically-based artificial navigation systems: review and prospects. *Progress in Neurobiology*. 51, 483-544.
- Webb, B. and Consi, T.R. (Eds) (2001). *Biorobotics: Methods and Applications*. The MIT Press.