Psikharpax: An autonomous and adaptive artificial rat

Jean-Arcady Meyer *, Agnès Guillot *, Patrick Pirim **, Alain Berthoz***

*AnimatLab/LIP6, Université Paris 6 - CNRS, France
** BEV, France
***LPPA – CNRS, Collège de France, France

Abstract
The Psikharpax project aims at endowing a robot with a sensori-motor equipment and a neural control architecture that will afford some of the capacities of autonomy and adaptation that are exhibited by real rats. The paper summarizes the current state of achievement of the project. It successively describes the robot's future sensors and actuators, and several biomimetic models of the anatomy and physiology of structures in the rat's brain, like the hippocampus and the basal ganglia, that have already been put at work on various robots and that make navigation and action selection possible. Preliminary results on the implementation of learning mechanisms in these structures are also presented.

I. Introduction

Since the two-month workshop of Darmouth that founded the field of artificial intelligence in 1956, and since the enthusiastic comments on the prospects of the discipline that this event triggered (Simon, 1957; Feigenbaum and Feldman, 1963), serious doubts have been raised (e.g., Dreyfus, 1979, 1992) about the chances that an artificial system might compete in the near future with the amazing capacities exhibited by the human brain. In particular, several researchers consider that it is largely premature to try to understand and reproduce human intelligence – whatever this expression really means – and that one should first try to understand and reproduce the likely roots of this intelligence, i.e., the basic adaptive capacities of animals (Brooks, 1999). In other words, before aiming at reproducing unique capacities that characterize man, like logical reasoning or natural language understanding, it might be wise to concentrate first on simpler abilities that human beings share with other animals, like navigating, seeking food and avoiding dangers. In this spirit, several research efforts are devoted to the design of so-called *animats*, i.e., simulated animals or real robots whose sensors, actuators and control architectures are as closely inspired from those of animals as possible, and that are able to "survive" or fulfill their mission in changing and unpredictable environments (Guillot and Meyer, 2001).

This article describes one such endeavor, the *Psikharpax* project which aims at designing an artificial rat that will exhibit at least some of the capacities of autonomy and adaptation that characterize its natural counterpart. In particular, this robot 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.

This article summarizes the current state of this project. In particular, it describes the robot's future sensori-motor equipment and the major modules of its control architecture. It also describes the behaviors that the robot Psikharpax already exhibits in simulation.

II. Sensori-motor equipment

Psikharpax will be a 50cm-long robot (Figure 1) equipped 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. Sensor fusion will be realized 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).

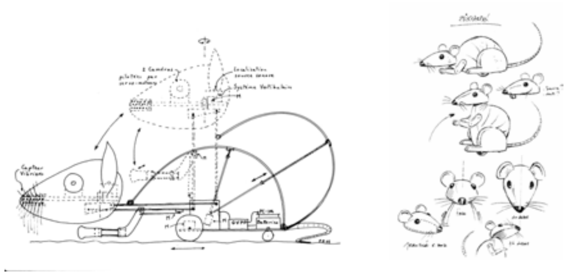

Figure 1. *The overall design of Psikharpax.*

Psikharpax will also be endowed 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.

Psikharpax will be equipped with several motors and actuators. In particular - despite the fact that such device is not really biomimetic - two wheels will allow the robot 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 2).

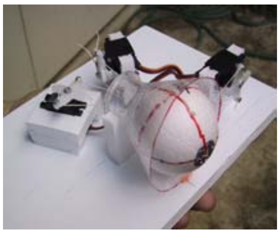

Figure 2. *An eye equipped with a camera and a log-polar sensor, which is actuated by three motors. The whole device obeys the Listing's law (von Helmholtz, 1955).*

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.

III. Control architecture

Likewise, several models of nervous circuits that contribute to the adaptive capacities of the rat are currently simulated or tested on real robots, and will be implemented in the final control architecture of Psikharpax. In particular, this artificial rat will be endowed with the capacity of effecting visual or auditory saccades towards salient objects, of relying on the optical flow to determine whether a given landmark is close or distant, of merging visual and vestibular information to permanently monitor its own orientation. Among such circuits, those that afford capacities for navigation and action selection have already been validated on preliminary versions of the future Psikharpax. The corresponding realizations will be now briefly described.

III.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 described here 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). 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. A link between two nodes memorizes at which distance and in which direction the corresponding places are positioned relatively to each other, as measured by the robot's idiothetic sensors (Figure 3, left). 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 3, right).

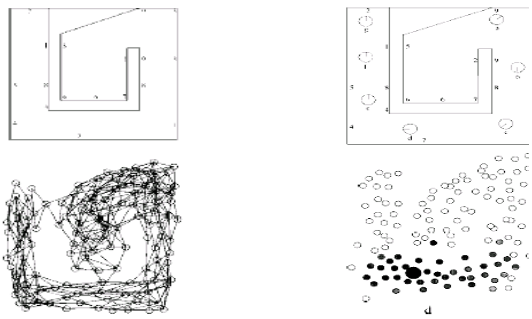

Figure 3. *Left: The topological map (bottom) created by the robot when it explores an unknown environment (top). Right: This map may be used by the robot to localize itself because the activity distribution of the nodes in the map changes as the robot moves through successive places a, b,…g in the environment (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 successfully recognized nodes.*

III.2. Action selection

To survive, the 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 the 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* designed by Gurney, Prescott and Redgrave (2001). 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 4). The inputs to the model are variables called *saliences* that are weighted functions computed from allothetic and idiothetic information monitoring the urgency associated with each possible act. The outputs of the model are inhibitions assigned to each possible action. At each time-step, the act which is the less inhibited is performed. 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 ensures clean and efficient switching between actions. However, the robot's survival depends on its chances of getting to the right place at the right moment, i.e., to a food place when it's *Potential Energy* level is low, or to a rest place when it lacks *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.

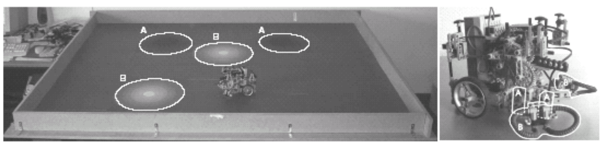

Figure 4. *Left: The environment showing "food" (A) and "rest" (B) places. Right: A Lego robot equipped with light sensors (A) and bumpers (B).*

III.3. 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 – like the *nucleus accumbens* in particular – and of several *basal ganglia-thalamus-cortex* loops in the rat's brain (Girard et al., 2003; Girard et al., In Press). The corresponding model (Figure 5) involves a *ventral loop* that selects directions of movement suggested by simple sensor processing – according to a simple *guidance* strategy – or by a more elaborated *topological navigation* strategy (Trullier et al., 1997). The latter puts some constraints on action selection because the robot is committed to regularly return to previously mapped areas in its environment in order to check the correctness of the current map. This need is expressed by a *Disorientation* variable managed by the model, which increases when the robot enters unexplored areas, which decreases when it returns to known areas, and which affects the computation of saliences. Likewise, the model takes into account a *dorsal loop* that controls non-locomotor behaviors, i.e., those corresponding to reloading actions that change the *Energy* and *Potential Energy* levels.

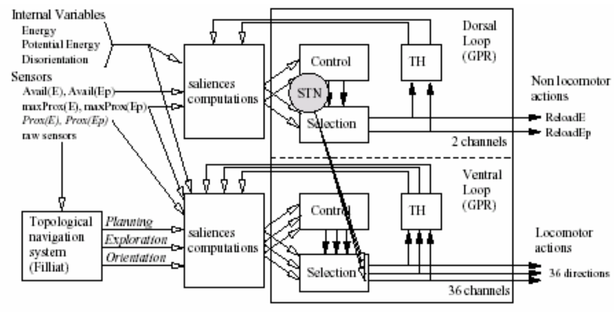

Figure 5. *The model integrating navigation and action selection calls upon two basal ganglia-thalamus-cortex loops. Each loop is managed by a GPR model and the coordination between loops is provided by the subthalamic nucleus (STN) of the dorsal loop, which is connected to the ventral loop. The dorsal loop selects a reloading action among two, the ventral loop selects a direction of movement among 36. Inhibitory connections are represented by black arrows, excitatory connections by white arrows.*

The robot simulated in the environment on the left of Figure 6 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 6, 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 via *Ep2*. Then, having consumed the corresponding "food", it will record the position of *Ep2* on its map. Thus, next time it will need to reload its *Potential Energy*, it will have the choice of navigating towards *Ep1* or *Ep2*.

In the environment of Figure 7, 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.

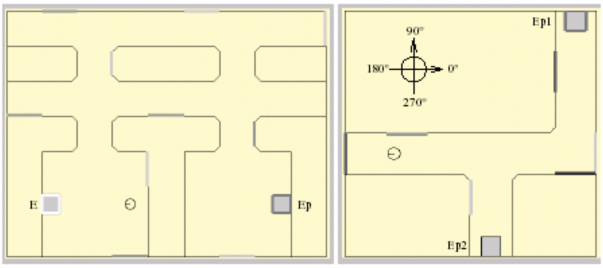

Figure 6. *Two environments used to test the connection of navigation and action selection models. E: "rest" place, Ep: "food" place.*

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.

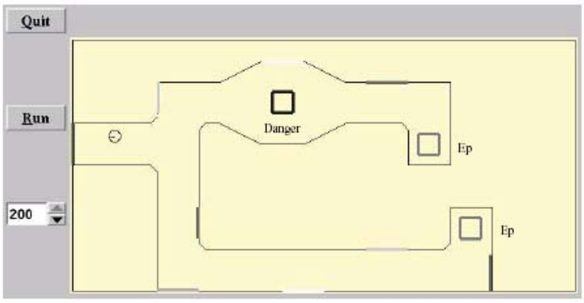

Figure 7. *Two trajectories leading to "food" places (Ep). One is shorter than the other but entails passing through a dangerous place (Danger).*

In the complex and challenging environment of Figure 8, the simulated robot autonomously survives, thanks to the numerous adaptive mechanisms and behaviors it has been endowed with (Girard, 2003).

III. 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 9).

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.

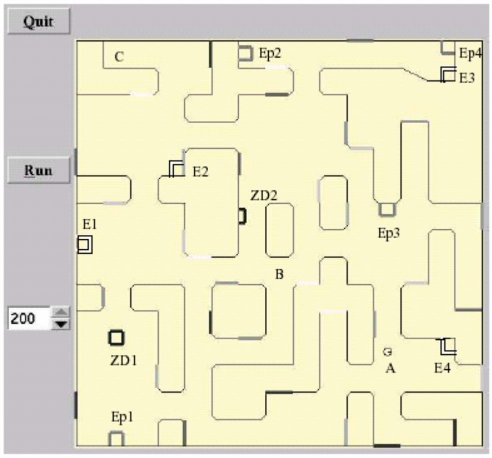

Figure 8. *A complex environment with four "rest" places (E), four "food" places (Ep) and two dangerous places (ZD).*

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 10). At every trial, one lamp out of four is lighted indicating behind which door the reward is accessible. When the robot succeeds to get such reward, the corresponding lamp is turned off and the robot has to learn to return to the center of the maze to trigger the lighting of another lamp designing another reward place. This setting reproduces an experiment on real rats (Albertin et al., 2000) and helps to interpret the corresponding results.

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 the reward, while the second (Critic 2) calls upon two. 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 - an information that is essential to propagate learning to other regions of the environment.

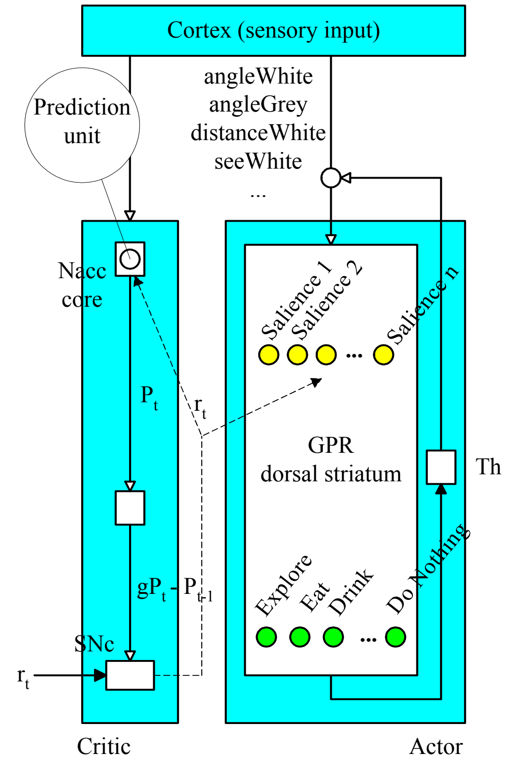

Figure 9. *The actor-critic model of reinforcement learning. The actor module is a GPR model that is segregated in different channels, with saliences as inputs and actions as outputs. The critic module propagates towards the actor module an estimate ř of the instantaneous reinforcement triggered by the selected action.*

It thus turns out that Critic 2 gives better results than Critic 1 (Figure 11) but that the corresponding module has still to be improved, for instance through the use of a greater number of prediction units, to let learning extend to the whole experimental environment and to speed up the corresponding process.

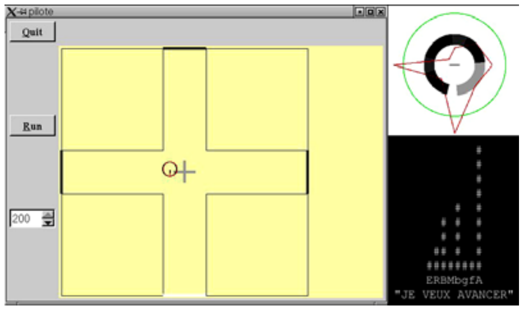

Figure 10. *Left: the robot in the plus maze environment. A lighted lamp indicates which door (in white) leads to reward. The other doors do not lead to reward and are shown in black. Upper right: the robot's visual perceptions. Lower right: activation level of each channel in the actor module.*

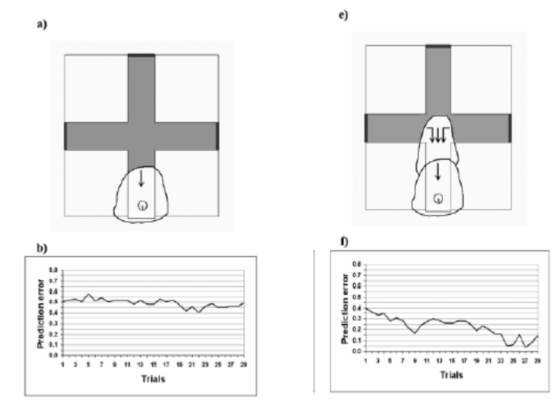

Figure 11. Reinforcement learning results obtained with an actor-critic architecture in which two different critic modules have been tested (left: Critic 1; right: Critic 2). Depending on which module is used, the correct actions to perform are learned in specific regions of the maze (top: white areas in the plus maze) and the improvement patterns of the prediction error at reward location over successive trials are different (bottom).

IV. Conclusions

The Psikharpax project aims at designing an artificial rat able to "survive" in a laboratory populated by humans and other robots. This animat will be endowed with numerous sensors and motors that are currently under development and that will serve to implement various reflexes. Its control architecture has been already tested in simulation and implemented on simpler versions of the future robot. In particular, models for navigation and action selection - which afford capacities of associative and reinforcement learning - have been successfully tested. It thus appears that Psikharpax will be able to explore an unknown environment, to build a topological map of it, and to plan trajectories to places where it will fulfill various internal needs, like “eating”, “resting”, “exploring” or “avoiding danger”. The first version of such an efficient robot is expected to be available at the end of year 2005: still a long way to the "whole rat" that Dennett (1978) - the spiritual father of the “whole iguana” - might have advocated; even a longer way to the intelligence of man.

Acknowledgments

This research has been granted by the LIP6 and the Project *Robotics and Artificial Entities* (ROBEA) of the Centre National de la Recherche Scientifique, France. The authors thank David Filliat, Benoît Girard, Mehdi Khamassi and Sidney Wiener for their valuable contribution to the project.

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.

Barto, A.G. (1995). Adaptive Critics and the Basal Ganglia. In Houk, Davis and Beiser (Eds). *Models of Information Processing in the Basal Ganglia.* The MIT Press.

Brooks, R. (1999). *Cambrian Intelligence. The Early History of the New AI.* The MIT Press.

Dennett, D.C. (1978). Why not the Whole Iguana? *Behavioral and Brain Sciences* 1:103-104.

Dreyfus, H.L. (1979). *What Computers Can’t Do: The Limits of Artificial Intelligence*. Harper and Row.

Dreyfus, H.L. (1992). *What Computers Still Can’t Do: A Critique of Artificial Reason*. The MIT Press.

Feigenbaum, E.A., and Feldman, J. (Eds.) (1963). *Computers and Thought*. McGraw-Hill.

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.

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. 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., J-A., Meyer, A., Berthoz 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.

Guillot, A. and Meyer, J.-A (2001). The Animat Contribution to Cognitive Systems Research. *Journal of Cognitive Systems Research.* 2(2), 157-165.

Gurney, K.N., Prescott, T.J. and Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*. 84, 401-410.

von Helmholtz, H. (1955). *Treatise on Physiological Optics*. The Thoemmes 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, Davis and Beiser (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.

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.

Schultz, W., Dayan, P. & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275:1593-1599.

Simon, H.A. (1957). *Models of man*. Wiley.

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.