A robot with autonomous spatial learning: A project overview

Christian Balkenius, Lars Kopp, Robert Pallbo

Lund University Cognitive Science
Kungshuset, Lundagård
222 22 LUND

E-mail:

`christian.balkenius@fil.lu.se`

`lars.kopp@fil.lu.se`

`robert.pallbo@fil.lu.se`

1. Introduction

To act in an unknown and continously changing environment, an autonomous robot must be able to react instantaneously on changes and unexpected events in order to avoid collisions and to update its maps. Successful navigation requires that the robot reacts primarily on its immediate sensory information and secondarily on its internal mapping of the spatial layout of the environment.

We have developed and constructed an experimental mobile robot equipped with a number of complementary sensory systems (Balkenius and Kopp 1994a). A video camera is mounted on a movable head that also contains a pair of microphones. Ultrasonic sensors are located around the body of the robot and a set of tactile sensors (whiskers) and a bumper are used to detect obstacles at a short range.

The project aims at developing the attention and navigation systems of the robot to include vision for spatial orientation. The choice of vision is natural since this modality contains the richest information for this tasks. The problems we are studying include automatic recognition of visual landmarks and reactions towards changes in the environment as well as the production of linguistic output on unexpected events. The solution to these problems are highly dependent on the *behaviour* of the robot and not only on its perceptual abilities. In this view, the main problem of visual navigation is not vision itself but rather the behavior that makes vision useful.

We have performed extensive computer simulations of reactive navigation and learning based on other modalities (Balkenius 1993a, 1994a) and developed algorithms for visual place recognition (Kopp 1994b) and motion detection (Balkenius & Kopp 1994b, Pallbo 1992a, b, 1993a, b, 1994a, b). A simple form of visually based obstacle avoidance has already been implemented successfully on the robot together with a tactile reactive control system. We have also studied the connection between visual input and linguistic output (Balkenius 1992a, 1994b, Kopp 1994a), and developed a neural network based system that is able to learn spatial relations between objects and produce elementary linguistic output (Kopp 1994a).

Our theoretical aim is to develop learning methods for autonomous agents that can construct control strategies based directly on their sensory and locomotor abilities. Instead of using a prespecified map, like a CAD design, our goal is to let the agent construct its own map from its sensory inputs in a form suitable for its own actions. Furthermore, the maps constructed should not depend on a specially prepared environment or artificial landmarks. We are using natural visual input from the video camera. Since there exists a large database on spatial orientation in biological systems (cf. Balkenius 1994c, Ellen & Thinus-Blanc 1987, Pallbo 1992c), this research field is one of the most promising areas for cognitive technology inspired by biological systems. We have previsouly studied biologically inspired architectural principles for the construction of autonomous agents (Balkenius 1993a, 1994c, Gärdenfors and Balkenius 1993). This work included a study of goal-gradients as a general representational tool for spatial programs and plans.

In a longer perspective, autonomous robots equipped with spatial learning have immense potentials for industrial applications. A system developed along the lines of our project will be useful in new types of automatic industries (e.g. in auto carriers). Furthermore, such systems can be used in applications for the physically disabled since autonomous mobile robots can function in a home environment which is not specially designed for robots. In this area, we are cooperating with CERTEC (Center for Rehabilitation Technology), Lund Technical University, and HADAR, Malmö.

The project is highly interdisciplinary and combines cognitive, language and neural network technology with autonomous systems.

1.1 Demands of spatial orientation

In order to realize an autonomous system capable of spatial learing and spatial orientation, one must solve a number of problems. During exploration and learning, the system must accomplish the following tasks:

- *Visual Place Categorization* Place representations must be constructed which can be used at later stages to recognize the location of the agent. The place categories must also make it possible for the agent to approach any location in the environment.
- *Map Learning* The agent must construct structures which can later be used to guide its locomotion from one location to another. This learning should be accelerated by using earlier spatial knowledge and requires an appropriate exploratory behaviour.

Our research focuses on the following abilities, which the agent must be able to perform when using the constructed map:

- *Place Recognition* The agent must be able to figure out where it is located on the basis of visual information only. This process involves the recognition of visual views or landmarks and a potential updating of the spatial map.

- *Action Selection* The agent must determine what action to perform in order to move closer to the goal. This mechanism is closely connected to dynamic task selection and goal-priorization.
- *Stable Approach* The agent must be able to approach its goal from any position within a region around the optimal path. If changes of the environment or imperfections in the motor system gets the agent off course, it should automatically try to approach the correct path again.
- *Reactive Obstacle Avoidance* The agent must be able to avoid obstacles in a reactive manner without too much computational overhead. Once the object is negotiated, the agent should continue on its way towards the goal. This ability rests on a combination of different sensory systems.

For each of these abilities, there exists several kinds of models, e.g. within control theory, pattern recognition, animal learning theory and cognitive science. Our goal is to develop these models in a way that makes them possible to combine in a unified system.

1.2 The Traditional Approach to Spatial Orientation

Spatial learning has traditionally been tackled with the same learning mechanisms as other areas of AI and robotics without much concern for the special requirements of this domain. As with many other techniques within AI, sensory learning and motor control have been considered problems distinct from map learning and path planning.

The traditional architecture of autonomous agents can be divided into three modules, each with its set of problems: Perceptual categorization, planning and reasoning engine, and execution interface to motor functions.

A modularization of this kind is usually referred to as a horizontal decomposition of the agent (Brooks 1991). In brief, the main problems with this approach are: (1) the computational complexity; (2) the interface between the plan and motor control; (3) the delayed feedback caused by the complexity of perception and planning; (4) the instability of the locomotor control as a consequence of delayed feedback. These problems have been approached during the last years by moving away from the traditional architecture in different ways. Our view of spatial orentation and learning have much in common with these newer research directions.

1.3 Behaviour Based Reactive Control

A number of investigations have shown that it is possible to attack the problem of spatial learning and motor control in a different way. We refer to what is usually called a vertical decomposition of the system where the whole chain from sensory signals to motor control is considered continuously (Brooks 1991, Maes 1990). These studies suggest new ways of controlling autonomous agents based on a close coupling between sensors and effectors which can avoid the problems outlined above.

According to the alternative principles, the construction of a system should progress from simple connections between sensors and effectors that control fundamental actions, such as moving forwards without colliding with obstacles, towards more complex behaviour that may be controlled by global maps of the environment. The alternative architecture emphasizes reactive control and making the path from sensors to effectors as short as possible.

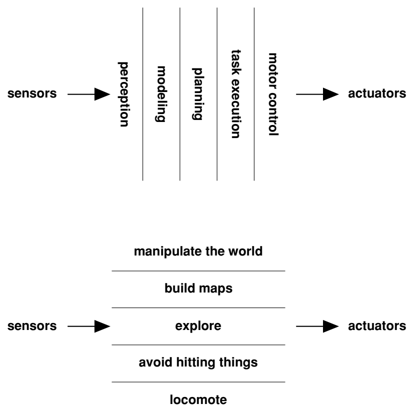

Figure 1. TOP. Traditional AI decomposition of intelligent control systems. **BOTTOM**. Behaviour based decomposition. (After Brooks 1991).

If the map controlling the agent is constructed using the actual sensory and locomotor equipment of the agent, it is possible to construct plans that can be executed reactively in a stable manner. This has the advantages of the reactive approach in that sensory signals are almost directly converted into motor commands.

Purely reactive control of behaviour as well as exploratory behaviour based on simple sensory systems such as IR-sensors and sonars have recently been intensively studied. Some systems using more complex inputs such as a laser scanner (Connell 1990) and vision (Horswill 1993) also exist. These systems have in common that they are not very computationally demanding since they build on certain invariants of the environment. This feature thus makes them cheap to manufacture and hence attractive for practical applications.

So far, none of these systems uses *visual* input to learn global maps of their environment. This is probably because vision has traditionally been considered a very computationally demanding process. One of our main goals is to extend the reactive types of robot architectures with spatial learning abilities based on visual input.

1.4 Cognition and Behaviour as a Hierarchic Control Process

The above approach lends itself naturally to the view of cognition as a *hierarchical adaptive control process*. The view that behaviour should be based on control theoretical notions, and not on planning and deduction, was pioneered by Powers (1973) and has recently been repeated by Klopf, Morgan and Weaver (1993). In the spatial domain, this means that the goal of the agent is to achieve a certain value for its spatial location. Its current location is considered a deviation from this desired value.

We believe that this view of spatial orientation makes it possible to construct, in a unified manner, control mechanisms that combine sensory processing and motor control with spatial learning. Analysis of the constructed maps in terms of stability and optimality can also be made in a direct way while still retaining a more classical type of analysis in terms of soundness and completeness (cf. Kartam and Wilkins 1989).

1.5 Potential Fields Methods

Another attempt to achieve stable locomotor control is by using potential fields methods (Arkin 1987). In this approach, goals and obstacles in the environment are given positive or negative potentials that generate gradient fields in space. By following these gradients, an agent will reach a specific goal object without running into walls and other obstacles.

The representation of the environment as a potential field, and more general as vector fields (Payton 1990), are powerful ways to understand the spatial representations constructed by an autonomous agents. They directly address two of the problems with the traditional approaches (i.e. the interface between the plan and motor control and the stability of the control scheme). We use potential fields as way to globally analyse the behaviour of an agent while keeping the local analysis in an agent-centered representation.

1.6 Reinforcement Learning and Stochastic Learning Automata

Reinforcement learning and Q-learning in particular can be used to link actions together in such a way that the execution of the action sequence results in a maximal pay-off from the environment (Barto, Sutton and Watkins 1990). This learning method is closely connected to other approaches such as dynamic programming (Barto 1992) and stochastic learning automata (Tsetlin 1973, Narendra and Thathachar 1974).

The learning progresses by testing a set of actions in a number of situations and collecting an immediate pay-off (or reward) from the environment. The sequence of actions that results after learning is the one that returns the maximal reward when it is executed. A crucial advantage of reinforcement learning compared to other learning approaches is that it requires no information about the environment except for the reinforcement signal. It also combines sequential learning with optimization in a simple way.

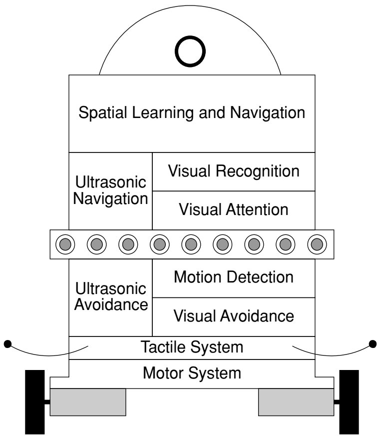

Figure 2. The goal of the project is to develop an autonomous mobile robot that can navigate using sensory information from touch, ultrasonic sensors and vision. Reactive control, in combination with spatial learning, allows the robot to optimize its performance in a familiar environment while retaining the ability to act in unknown situations.

Unfortunately, the method is very slow for most applications since every action must be tested a large number of times at each location. This has made reinforcement learning techniques impossible to use in mobile robots since the learning time would be too large for any practical use. However, recent progress in this field has shown that if reinforcement learning is combined with an adaptive forward model of the environment, fast learning is possible (Peng and Williams 1993).

Our approach to spatial mapping is based on a combination of reinforcement learning and potential fields methods and is much faster than earlier methods since it exploits a number of properties of the spatial domain (cf. Balkenius 1993a). It is different from potential fields methods in that it utilizes agent centered action based representations.

2. Central Project Goals

Our central goal is to develop a robot that is able to solve various problems of spatial navigation. The sensory inputs of the robot will be a combination of tactile, ultrasonic and visual information. We strive for a robot that can solve the following problem types, in increasing order of difficulty: Reactive obstacle avoidance using tactile and ultrasonic sensors; place recognition based on ultrasonic information only; exploratory behavior; visual obstacle avoidance; visual place recognition; goal-seeking behavior using boh ultrasonic and visual information; attention focusing on changes in the environment; linguistic production of information concerning such changes, using either speech synthesis, or written output on a monitor.However, in order that the robot be capable of performing these tasks, we are developing theoretical models for the following three areas.

2.1 Visual Place Detectors

Vision is the only sensory system that works at all distances while combining distance sensitivity with object recognition. Tuned visual detectors are place recognition devices that can be taught to generate a local generalisation surface around any goal point in an environment. In a region around the goal point, the output of a tuned detector generates a stable control strategy for the approach of the goal (Schmajuk & Blair 1993). We have developed a new type of visual place detector the output of which can be tuned to produce a maximal response at any location in space (Kopp 1994b). By moving the agent towards the maximum of this output, an agent can be made to locally approach any location in space.

The visual algorithm is based on a new type of unsupervised neural network that can associate between the visual input and a corresponding place category as well as to similar visual views. As a consequence, the network forms an adjacency net of visual views. This approach to visual representation is very different from the traditional view (Marr 1982) since it does not try to construct an object centered representation of the visual scene. The network design is inspired by a range of neurophysiological findings (e.g. Edelman and Bülthoff 1992, Tanaka 1993).

The algorithm is quite fast and does not require any complex and time consuming visual preprocessing such as segmentation or object recognition. Thus, our analysis of the visual scene is very shallow compared to other approaches (e.g. Suburo and Shigang 1993), but is sufficient for spatial navigation. This part of our work has progressed to a point where a demonstration of the algorithm in an unknown environment is in principle already possible.

2.2 Mapping of Spatial Locations

Tuned visual detectors can successfully control the approach to an arbitrary goal location. But since this method applies only to local regions around the goal, we intend to extend the mapping process to the entire environment. To do this, the whole environment is mapped, using tuned detectors, into a set of approach regions that cover the entire space. It is necessary to cover the environment with a large number of these regions. The regions are then linked together in such a way that the goal can be reached via a succession of subgoal locations.

We have developed a learning method that can be used to link locations in space into maps of the environment based solely on unanalyzed sensory information and the locomotor repertoire of the agent (Balkenius 1993a). In this work, however, the sensory system was based on simulated 'olfaction' and not on vision. In the project we will continue to develop new architectures for neural networks for dynamic updating of spatial maps.

One goal in the project is to combine this method with the tuned detectors described above in order to implement the whole chain from visual signals to locomotor control. We are currently evaluating the mapping system using ultrasonic sensors and active ultrasonic landmarks. In the second step, visual detectors will be used for this task.

This mapping technique is based on reinforcement learning. The whole process can be viewed as a stochastic learning automata that establishes a goal gradient for the environment (Barto, Sutton and Watkins 1990). Goal gradients are similar to plans in the traditional approaches except that they include a stable control strategy. Balkenius (1993a, b) presents mechanisms that can be used to dynamically select between a number of goal gradients depending on the current goal of the agent.

During learning and exploration, the reactive strategies of the agent plays a role similar to search heuristics in the traditional approaches. If the autonomous exploratory behaviour of the agent is replaced by explicit manual control, the agent can learn by instruction as well as by exploration. In this case, the automatic mapping features are only used to keep track of changes in the environment.

Reinforcement learning is usually considered too slow for use in path learning. In order to achieve efficiency, it must be complemented with some preprocessing in the form of place recognition or establishment of location in space. This learning algorithm, described in Balkenius (1993a), shows many similarities with Q-learning (Watkins 1992) but is much faster since it exploits a number of aspects of the spatial domain as well as using a reactive control system as a search heuristic. This learning method has also been thoroughly studied in computer simulations (Balkenius 1993a, 1994a).

The mapping process is highly dependent on the attentional system and the exploratory behavior used by the agent. The study of such behaviors is thus one of the central goals of the project (Balkenius 1993c).

2.3 Advanced Learning and Generalization

Another project goal is to develop the learning methods to make generalisations possible without giving up the control view of behaviour. New places have many similarities with old situations and the agent should generalize from previously encountered situations. The central problem here is to represent actions and situations on multiple levels (Pallbo 1993b). Such an ability is in many respects similar to chunking in SOAR (Newell 1990). But like in most traditional systems using chunking, the operators in SOAR do not constitute control strategies and can thus not be directly used to control the actions of an agent. This is something we hope to accomplish using selectionist learning methods in neural networks (cf. Edelman 1987).

Another use for higher level representations is that they aid the construction of a forward model of the environment (cf. Nguyen and Widrow 1989, Jordan and Rumelhart 1992, Sutton 1992). This gives the agent the ability to train and retrain its reinforcement learning system faster than without such a model (Peng and Williams 1993). The agent constructs an inner world where actions can be tried out before committing them to the unforgiving reality (Gulz 1991, Gärdenfors 1993). This is especially important when the environment has changed and a large number of updates are necessary to establish the new goal gradient.

3. Long-Range Plans

There are two long-term goals of the project. The first is to develop our models into a complete and self-contained system that can be used as the basis for a product in the form of a robot.

Our second long-range goal is theoretical and involves the extention of our model to more complex situation and learning tasks. In the design of a more general model one must take into account that not only the environment is unknown, but also the nature of the problems that the agent encounters. Consequently, it must create a problem space before it can search for a solution. This task would be impossible to handle if the agent were not able to reapply parts of it knowledge from other problem spaces. By being able to generalize from one familiar situation to an unknown one, the agent does not need to construct the problem space from scratch.

To accomplish such a dynamic learning, one cannot describe the process of knowledge acquisition in a static (meta)system. If this were possible, the acquisition would be restricted to the frames of that system. Any event that cannot be interpreted by the metasystem can simply not be taken into account. It follows that there will be a predetermined limitation on the knowledge that the agent can acquire.

The solution to this problem is likely to be found in approaches other than the functionalistic. Darwin machines (Calvin 1987) is one potential candidate that we wish to explore. Lately, some selectionist models have been proposed in the field of cognitive science (e.g. Edelman 1987, Changeux et al. 1982).

Our approach, which is primarily characterized by the usage of spontaneous activity as a source of variation, can be integrated into the adaptive system described in the previous sections. This will be made by modifying the current learning algorithms rather than interfacing a new subsystem on top of the rest. Our first effort to study how this can be achieved can be found in Pallbo (1993b, 1994b).

References

Arkin, R. C. (1987) "Motor schema based navigation for a mobile robot: an approach to programming by behavior", Proceedings of the 1987 IEEE international conference on robotics and automation, 264-271.

Balkenius, C. (1992a), "Neural mechanisms for self-organization of emergent schemata, dynamical schema processing, and semantic constraint satisfaction", *Lund University Cognitive Studies*, *14*.

Balkenius, C. (1993a), "Natural intelligence for autonomous agents". In B. Svensson (ed.), *Proceedings of the international workshop on mechatronical computer systems for perception and action*, 181-196.

Balkenius, C. (1993b), "The roots of motivation". In J.-A. Mayer, H. L. Roitblat and S. W. Wilson (eds.), From animals to animats II, 513, MIT Press, Cambridge, MA.

Balkenius, C. (1993c), "Motivation and attention in an autonomous agent", paper presented at the Workshop on Architectures Underlying Motvation and Emotion – WAUME 93, University of Birmingham.

Balkenius, C. (1994a) "All you never wanted to know about the BERRY III environment", *LUCS Minor*, *2*, ISSN 1104-1609.

Balkenius, C. (1994b), "Some properties of neural representations". In M. Bodén and L. Niklasson (eds.), *Connectionism in a broad perspective,* 79-88, Ellis Horwood, New York.

Balkenius, C. (1994c), "Biological learning and artificial intelligence", *Lund University Cognitive Studies*, *30*.

Balkenius, C. & Kopp, L. (1994a), "My first robot", *manuscript*, Lund University Cognitive Science.

Balkenius, C. & Kopp, L. (1994b), "A simple object tracking algorithm", *manuscript*, Lund University Cognitive Science.

Barto, A.G. (1992), "Connectionist learning for control: an overview". In Miller, W.T., Sutton, R.S. & Werbos, P.J. (eds.) *Neural Networks for Control*, Cambridge, MA: MIT Press.

Barto, A. G., Sutton, R. S. & Watkins, C. J. C. H. (1990), "Learning and sequential decision making". In M. Gabriel and J. Moore (eds.) *Learning and computational neuroscience: foundations of adaptive networks*, 539-602, Cambridge, MA: MIT Press.

Brooks, R. (1991), "New approaches to robotics". *Science1*, **253**, 1227-1232.

Calvin, W. (1987), "The brain as a Darwin machine", Nature, 330, 33-34.

Changeux, J.-P., Heidmann, T. and Patte, P. (1982), "Learning by selection", in Marler, P. and Terrace, H. S., The Biology of Learning, 115-133, Berlin: Springer-Verlag.

Connell, J. H. (1990), *Minimalist Mobile Robots*, Boston, MA: Academic Press.

Edelman, G. M. (1987), Neural Darwinism: The Theory of Neuronal Group Selection, Cambridge, MA: Basic Books.

Edelman, S. & Bülthoff, H. H. (1992), "Orientation dependence in the recognition of familiar and novel views of three-dimensional objects", *Vision Research*, *32*, 12, 2385-2400.

Ellen, P. & Thinus- Blanc, C. (1987), Cognitive processes and spatial oreintation in animal and man, Dordrecht: Nijhof.

Gulz, A. (1991), "The planning of action as a cognitive and biological phenomenon", *Lund University Cognitive Studies*, *2*.

Gärdenfors, P. (1993), "Människans medvetande – och maskinens" (The consciousness of man – and of machines), to appear in *Om själen*, ed. by A. Ellegård, Natur och Kultur, Stockholm.

Gärdenfors, P. & Balkenius, C. (1993), “Varför finns det inga riktiga robotar?”, *Framtider*, *12*, 11-14.

Horswill, I. (1993), “A simple, cheap, and robust visual navigation system”. In J.-A. Mayer, H. L. Roitblat and S. W. Wilson (eds.) *From animals to animats II*, Cambridge, MA: MIT Press.

Jordan, M. I. & Rumelhart, D. E., (1992), “Forward models: supervised learning with a distal teacher”, *Cognitive Science, in press*, *16*, 307-354.

Kartam, N. A. & Wilkins, D. E.,(1989), “Toward a foundation for evaluating AI planners”, Report no 471, SRI International.

Klopf, A. H., Morgan, J. S. & Weaver, S. E. (1993), “A hierarchical network of control systems that learn: modelling nervous system function during classical and instrumental conditioning”, *Adaptive Behavior*, *1*, 263-319.

Kopp, L. (1994a) “A neural network for spatial relations: Connecting visual scenes to linguistic descriptions”, *manuscript*.

Kopp, L. (1994b) “Place recognition”, in preparation.

Maes, P. (1990), *Designing autonomous agents*, Cambridge, MA: MIT Press.

Marr, D. (1982), *Vision*, San Fransisco: Freeman.

Narendra , K. S. and Thathachar, M. A. L. (1974), “Learning automata”, IEEE Transactions on systems, man, and cybernetics, 4, 323-334.

Newell, A.,(1990), *Unified theories of cognition*, Cambridge, MA: Harvard University Press.

Nguyen, D. & Widrow, B. (1989), “The truck backer-upper: an example of self-learning in neural networks”. *Proceedings of the international joint conference on neural networks*, 357-363, Piscataway, NJ: IEEE Press.

Pallbo, R. (1992a), “Neuronal selectivity without intermediate cells”, *Lund University Cognitive Studies*, *13*.

Pallbo, R. (1992b), “Simple direction sensitive neurons”, in Astor, E. and Davidsson, P., Projekt i Tillämpad AI 1992, Del1, LU-CS-IR:92-2, Lund University Dept. of Computer Science.

Pallbo, R. (1992c), “Natural and artificial planning systems”, *manuscript,* Lund University Cognitive Science.

Pallbo, R. (1993a), “Visual motion detection based on a cooperative neural network architecture”. In E. Sandewall and C. G. Jansson (eds.) *Scandinavian conference on artificial intelligence - 93*, 193-201, Amsterdam: IOS Press.

Pallbo, R. (1993b), “Ontogenesis in neural networks”, *Lund University Cognitive Studies*, *24*.

Pallbo, R. (1994a), “Motion Detection — A neural model and its implementation”, *LUCS Minor*, *1*, ISSN 1104-1609.

Pallbo, R., (1994b), “Motion detection by cooperation: A case study of the time dimension in neural networks”, *LU-CS-TR: 94-125* (Licentiate thesis).Payton, D. W. (1990), “Internalized plans: a representation for action resources”, in Maes, P., *Designing autonomous agents*, Cambridge, MA: MIT Press.

Peng, J. & Williams, R. J. (1993), “Efficient learning and planning within the Dyna framework”, *Adaptive Behavior*, *1*, 4, 437-454.

Powers, W. T. (1973), *Behavior: the control of perception*, Chicago: Aldine Publishing Company.

Schmajuk, N. A. & Blair, H. T. (1993), “Place learning and the dynamics of spatial navigation: a neural network approach”, Adaptive Behavior, 1, 353-385.

Suburo, T. & Shigang, L. (1993), “Memorizing and representing route scenes”. In J.-A. Mayer, H. L. Roitblat and S. W. Wilson (ed.) *From animals to animats II*, Cambridge, MA: MIT Press.

Sutton, R. S. (1992), “Reinforcement learning architectures for animats”. In J.-A. Meyer and S. W. Wilson (ed.) *From animals to animats*, 288-296, Cambridge, MA: MIT Press.

Tanaka, K., (1993), “Neuronal mechanisms of object recognition”, *Science*, *262*, 685-688.

Tsetlin, M. L.,(1973), Automaton Theory and Modeling of Biological Systems, Academic Press.

Watkins, C. J. (1992), “Q-learning”, *Machine Learning*, *8*, 279-292.