Surviving in a Hostile Multi-Agent Environment: How Simple Affective States Can Aid in the Competition for Resources

Matthias Scheutz

Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
mscheutz@cse.nd.edu

Abstract. In this paper, I will argue that agents with simple affective inner states (that can be interpreted as "hunger" and "mood") can have an advantage over agents without such states if these states are used to modulate the agents' behavior in specific ways. The claim will be confirmed using results from experiments done in a simulation of a multi-agent environment, in which agents have to compete for resources in order to survive.

1. Introduction

Emotions, despite their importance for natural cognitive systems, have been ignored in artificial intelligence (AI) research for a long time, possibly because classical AI was focused primarily on different ways of processing symbols to achieve cognitive tasks.[1] Up to the early 80ies, the emphasis was on representations and how they can be manipulated.[2] Only after a shift in attention from "disembodied" and "disconnected" to "embodied" and "situated" AI, that is, from focussing on the mere abstract processing to focussing on the actual performing of actions, has AI started to recognize what animal behaviorists a ethologists (e.g. see [11]) kept stressing again and again, that emotions play a crucial role in the processing of information and the regulation of behavior.[3]

There is quite some evidence that many animals seem to have "moods" or at least certain "modes" related to important survival functions, and these moods or modes can be conditioned to environmental stimuli.[4] Actually, much evidence indicates that

[1] See [14] for example, for a discussion of the "sense-think-act" cycle in the field of autonomous agents.

[2] Only a few researcher have stressed the importance of emotions, even to the point that they were willing to attribute emotions to simple behaviors (e.g., see [4]).

[3] I still think that emotions have not yet received the attention in AI that they deserve. E.g., Arkin's great overview of behavior-based robotics dedicates only two pages to emotions (see ch. 10.1 of [1]).

[4] E.g., see [17] for feeding behavior, [8] for defense behavior, [7] for sexual behavior, or [9] for a review of some of the issues related to moods and learning. All of this research is in line

the power of these moods is often great enough to produce extremely *maladaptive* behaviors, when the animal is taken away from its natural habitat (e.g., in a laboratory setting, see [9], or [3]).

Although there are some recent investigations studying agents with emotions (e.g., the KISMET project at MIT, or [12]), not much is known about how emotional states can be employed to make agents more adaptive. Note that "adaptive" does not automatically imply "learning", as an agent might well have a fixed, unalterable architecture, yet might be able to alter its behavioral responses to the *same* environmental stimulus according to the status of some *inner* states (where these states are also sometimes called "proprioceptive" states, e.g., see [15]). These inner states provide additional input to modules that act upon stimuli, thus allowing for more complex decisions to be made without having the need of more sense organs. Furthermore, some of these states are quite simple to maintain, yet can serve a wide range of functions in the control of an agent (e.g., a state corresponding to the "hunger"-level of an agent can have tremendous influence on its behavior; at the same time its activation is indirectly proportional to the charge level of a battery in an autonomous robot, say).

It is the purpose of this paper to show that simple emotional states can facilitate adaptive behavior (which otherwise might be difficult to accomplish, probably only at the expense of adding additional behavioral modules). I will refer to these states as "affective states", as I do not want to get involved in a discussion as to whether robots can *experience* emotions.[5] The states I will describe do, however, come close to what is sometimes called *primary emotions* (such as startled, terrified, sexually stimulated, etc. see, for example, [16]).

In the following I will first spell out my claim in more detail and motivate the experimental setup, i.e., the computer simulation used to argue it. Then I shall describe the design of the simulation, in particular, that of the involved agents in detail, concluding with a summary of the results and their implications for AI.

2. The Power of Simple Affective States

There are various ways of elucidating how inner states can influence the behavior of agents. In fact, it does not take much to see that agents with inner states are in some sense "more powerful" than agents without inner states: the latter can keep track of

---

with the views of the classical ethologists [10] and [18], in particular, the views that the overall behavior of an agent can be understood as the interplay of many less complex behaviors that are arranged hierarchically.

[5] I am not only reluctant to call these states "emotional states", because they do not correspond to any of the widely accepted emotional states such as love, anger, faith, distrust, etc., but also because I think that one has to distinguish between "having these states" and "experiencing them". Usually, emotional states are associated with experiencing them as such, yet, in order to experience them, they have to be the experience of *somebody* and I do not believe that *there is somebody present in these primitive agents for whom they could be experiences* (interestingly, there are people who are willing to ascribe genuine emotions to even very "unemotional" behaviors, see [13], for example).

states of affairs in the environment, whereas the former cannot. Unfortunately, inner states have mostly been used to keep track of states of affairs that are *external* to the agent (and most of the examples in standard AI textbooks use such external factors). Much less attention has been devoted to inner states keeping track of *internal states of affairs*, that is, of information that is directly available *inside* the agent and often times *most relevant* to its proper functioning (or call it "survival", if you will). For example, consider a simple reactive agent, which consists solely of a sensor and a motor system that are connected in a certain way so as to exhibit a certain behavior (a reflex, for example). The sensor-motor-system *implicitly* divides the agent's world into two categories: stimuli that trigger the reflex, and stimuli that do not trigger it. Suppose, the reflex is thought to move the agent into a direction different from the current (e.g., away from an obstacle it might bump into). Because the purpose of the reflex (assuming the agent is an natural kind and the reflex is the product of evolution, or if the agent is an artifact that the reflex has been purposefully designed to prevent damage to the agent) is to prolong the life of the agent, it is "good" for the agent if the behavior is not evoked by an external stimulus. In other words, not having to react is good, having to react is bad. Having to react many times within a short period of time is *very bad* (as it implies that the environment is cluttered with obstacles or that the agent is making stupid moves). So, an internal state that integrates the stimulation of the sensor over time might serve as a measure of how much the agent "likes" the current environment. Or putting it less intentionally, such a state, connected to the motor system with inhibitory connections, could make the agent more reluctant to leave a "safe" environment. Note that this kind of state turns an implicit goodness measure into an explicit one. It allows the agent to *explicitly* monitor what otherwise is only *implicitly* given. Furthermore, such a state might be very easy to realize. In the above example a simple neuron with its axon connected via an inhibitory synapse to the motor system, performing temporal integration will suffice.

In the following, I will describe an experimental setup that is intended to demonstrate and confirm the above reasoning.

3. The Simulation

To make that case that agents with affective states outperform agents without such states, a setup is required, in which agents of both kinds can interact. Ideally, I would have liked to use real robots to test this claim as I believe that "the world is its own best model" [5]. Unfortunately, this was not possible under resource constraints. So, a simulation of an artificial multi-agent environment had to serve as the touchstone instead.

3.1 The Environment

The simulated environment (the "world") consists of a 40 x 40 grid. Each cells can host at most one of four kinds of objects at any given time: agents (with or without affective states), energy sources, moving, and static obstacles. While agents and energy sources can occupy only one square at any time, obstacles may occupy an arbitrary number of squares. Moving obstacles move at a constant speed in a predetermined direction (and "wrap around" the confines of the environment). Static obstacles and energy sources are stationary. Energy sources store a fixed amount of energy. They are created spontaneously at random locations within the world and stay there for a pre-determined period of time, unless consumed by agents. A parameter of the simulation determines the frequency at which energy sources appear.

3.2 The Agents

There are, as indicated before, two kinds of agents, those with affective states, call them "affective agents" and those without such states, called "reactive agents". Each agent, affective and reactive, is equipped with two sensors of limited reach: a "sonar" sensor to detects obstacles (including other agents and the boundaries of the environment, but not energy sources), and a "smell sensor" to detect energy sources (i.e., "food"). Sensors can be viewed as generating a vector-field for each location in the environment, where the orientation of the vector at each position corresponds to the direction the agent should be heading in to either come closer or to avoid a particular region (see [2]). In the case of energy-sources, the vector indicates attraction, whereas in the case of obstacles the vector indicates repulsion, the amounts of which are indicated by the magnitude of the vector.

Agents can use a combination of these vector fields to obtain the direction, in which they should be heading to find energy—this is where the difference between affective and reactive agents comes to bear: while reactive agents use a fixed weighted sum of both vector-fields (i.e., the gain in the resepctive schemas is given, see [1]), the affective states in affective agents can influence the combination of the vector-fields, thus leading to different combinations at different times depending on the status of the affective states. Two simple affective states, which I will call "hunger" and "mood", are implemented in affective agents. These states receive input from an additional sensor, the "energy level sensor". Note that this sensor, as opposed to the other two, is an *internal sensor* monitoring the state of the internal energy store (the batteries, say).

Agents can only move straight one square at a time in the direction they are heading (i.e., in one of the eight possible directions as determined by the eight surrounding squares). In order to move to a surrounding square that is not straight ahead, an agent must first turn in the new direction, then it can perform the move into the square. By design, agents can never occupy a square of the border of their world.

Each agent looses energy as time passes by and thus has to find energy sources in order to survive. While turning their head (i.e., changing the direction of movement) does not cost any more energy than staying still, moving results in an additional loss of three times the amount. Agents can obtain new energy by moving over energy sources, which will store the amount of energy provided by the energy source in the agent and cause the depleted energy source to be removed from the world. When agents come too close to (static or moving) obstacles, other agents, or the border of their world (as determined by a preset parameter) in the direction they are heading, a "reflex"-like behavior will attempt to turn the agent into a direction, in which there are no obstacles. If such a heading can be found, reactive agents will make a “reflex”-motion either always at normal speed or always twice as fast, whereas affective agents can choose between both speeds depending on their inner states.

3.3 The Task

Agents can be destructed in three different ways, in which case they are removed from the world: (1) they run out of energy, (2) they bump in other agents or obstacles, and (3) another agents or moving obstacles bump into them. The task of each agent is to avoid any of these circumstances and “survive” as long as possible. Simulations start with a certain number of agents of each kind and end after a predetermined number of updates, after which the number of remaining agents of each kind is counted.

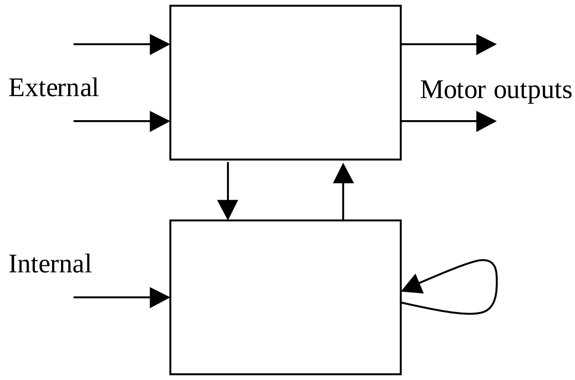

Fig. 1. The reactive layer shared by both kinds of agents and the additional affective layer of affective agents. Inputs to the reactive layer come from the external sonar and smell sensors, motor output goes to the moving and head-turning system. Inputs to the affective layer are provided by the energy level sensor, the reactive layer as well as the previous state of the affective layer, while outputs affect only both layers (they are not connected to any effectors of the agent).

4. The Agents’ Design and Its Justification

Both kinds of agents share the so-called “reactive layer”, their basic control structure. In addition, affective agents possess what I call the “affective layer”, which exerts influence on the reactive layer depending on affective states (see figure 1).

4. 1 The Reactive Layer

The reactive layer is realized as a set of finite state machines, which run in parallel and can influence each other (closely related to the style of Brooks’ subsumption architecture, [6]). It consists of the following four finite state machines (see figure 2):

- SONAR
- SMELL
- COLLIDE
- FORCE
- FORWARD

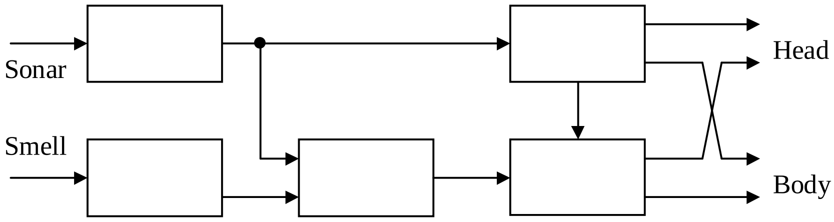

Fig. 2. The internal structure of the reactive layer. Arrows indicate information flow between systems. "Head" indicates output to the agent's head (which is used to change the current heading), "body" indicates output to the effectors that move the agent in the direction of its heading.

Very briefly, the SONAR and SMELL systems monitor the sonar and smell sensors, respectively, and store direction and magnitude of their current vectors. FORCE combines the information from SONAR and SMELL, computes the weighted sum of both vectors according to some predetermined weight scheme, and provides the actual direction the agent should be heading in. FORWARD uses this information to move the agent in that direction (possibly after first reorienting the agent's heading). COLLIDE is the most complex subsystem in that it continuously monitors the agent's heading and checks (using input from SONAR), if something is right ahead of the agent (within some predetermined range). If so, it obtains control over the motor outputs by inhibiting FORWARD and stops the agent. It then initiates a random sequence of reorienting the agent. The agent remains stationary until a heading has been found, in which there are no immediate obstacles (using again SONAR). Then a "reflex" moving the agent one step in this very direction is executed, after which control over the outputs is released and passed back to FORWARD. The rationale for the reflex should be clear: despite the agent's attempt to avoid obstacles (by integrating sonar readings in the FORCE module), its attraction to energy sources might still lead it on a collision course with an obstacle, if only the attraction to food is large enough. Also, a moving obstacle might cross the agent's trajectory causing an equilibrium between the attraction to the energy source and the repulsion caused by the obstacle, thus leaving the agent "undecided", i.e., stationary. Consequently, if the obstacle happens to be adjacent to the agent, the agent will get run over by the obstacle. Finally, if there is little energy available in the environment at any given time, agents tend to move towards the boundaries of their environment (by which they are least repelled) and the reflex prevents them from wandering off the confines of their world. Note that the reflex does not prevent agents from getting run over by obstacles or being "bumping into" by other agents (since there are again cases, where an agent—being cornered by other agents, for example—has reached an equilibrium point, is thus unable to move, and gets run over by a moving obstacle).

4. 2 The Affective Layer

The affective layer, as opposed to the reactive layer, is realized as a very simple recurrent neural network, which runs in parallel with the modules of the reactive layer (although nothing essential hinges upon the fact that it is a neural network).[6] The output of each unit is the weighted sum of its input units. The network *proper* consists of only two genuine (hidden) units—the two affective states—, the other units serve merely as inputs ("stopped" and "energy level") and outputs ("reflex strength" and "food attraction"), respectively—see figure 3).

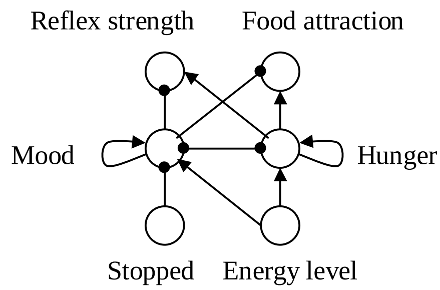

Fig. 3. The neural network implementation of the affective layer. Lines with arrows indicate excitatory connections, lines with bullets indicate inhibitory connections.

The input "stopped" is derived from the link between the COLLIDE and the FORWARD module in the reactive layer: if COLLIDE inhibits FORWARD, then this inhibitory input becomes active (expressing the agent's preference to move out of threatening situations). The other input reflects the current energy level of the agent, where the actual value is fed into the "mood" unit (reflecting the fact that the agent is in a better mood, when its energy level is higher), while *the reciprocal value* "1/(energy level)" is fed into the "hunger" unit (reflecting the fact that the "hunger" level increases as the energy level decreases).

Both, the "hunger" as well as the "mood" unit have self-excitatory connections, reflecting the idea that "being hungry" makes one even more hungry and "being in a good mood" tends to be self-sustaining (regardless of *minor* irritations). By the same token, "being in a bad mood" alone tends to even worsen the mood without further ado. The rational for the mutual inhibition of both units is that hunger contributes to one's bad mood, whereas being happy tends to let one forget about hunger.

Both hidden units have excitatory connections to the "reflex strength" parameter in the COLLIDE module, which determines, if the agent should move one or two fields at a time when performing a reflex action (recall from section 3.2 that this value in reactive agents is either set to 1 or 2 or is chosen at random each time a reflex is performed). The possibility of changing this value, i.e., of controlling the strength of the reflex depending on the agent's mood (as opposed to having it fixed), allows the agent to react more forcefully if it is in a "bad mood". Since being in a bad mood is directly related to hunger, an agent being caught between different repulsive forces

[6] The weights of the connections were obtained experimentally. It seems to me that it should be possible to "learn" using some version of reinforcement learning.

will turn further away from obstacles when it is more hungry, thus increasing the likelihood of “picking up a scent” (that is, a trajectory that will lead it to an energy source) instead of the being forced to remain “blocked”. At the same time, the agent will not *waste* energy on reflexes unnecessarily if it is happy (which means, as a consequence, that there is no immediate need for food), thus conserving energy. Note that an ongoing “block” situation *alone* can rapidly worsen an agent’s mood (without having the agent necessarily be hungry), and thus can also lead to more forceful reflexes.

The other output is connected to the parameter “food attraction” in the FORCE module. This parameter is the weight associated with the vector from the smell sensor, which is used by the FORCE module to compute the weighted sum of the sonar and smell vector fields. Changing this value while keeping the positions of energy sources and obstacles constant, can change the trajectories of the combined vector field significantly: what was attractive before, might not be attractive any longer (and vice versa). It is, in my view, to a large part this possibility of controlling their trajectories (and also to some degree the above-mentioned control of the reflex strength) that accounts for the advantage of affective agents over reactive agents in the experiments described in the next section.

5. Experiments and Results

Before starting the actual experiments, various parameters of the environment had to be experimentally determined such as the number and sizes of stationary obstacles, the number, sizes and speeds of moving obstacles, the number of energy sources together with their energy capacities, frequency of appearance, and life time, etc. Once the various degrees of freedom of the environment had been fixed, it was possible to determine an appropriate number of agents to inhabit the environment as well as an appropriate time span for which to run the simulation (in order to see any effects of the agents’ interactions at all). The final figures for the various parameters used in all the subsequent simulations are:[7]

- two stationary obstacles (of dimensions 3 x 5 and 5 x 6, respectively)
- two moving obstacles (of dimensions 3 x 3 and 4 x 3, respectively), the first moving vertically, the second diagonally at a constant speed of 1/10-th the speed of agents
- one energy source in the beginning
- the capacity of energy sources fixed at 250 with life time approximately 400 time steps
- new energy sources appear on the average every 30 time steps at random locations
- agents start out with an energy level of 500
- the duration of the simulation fixed at 2000 time steps

[7] Note that there are many possible, appropriate parameter settings and that nothing essential hinges on the above choice of parameter values.

It is worth pointing out that only reactive agents were used in the above simulations and that only in only 10 percent of all simulations did some reactive agents survive the end of a simulation.

Using the above configuration, three experiments were run: in *experiment A* four reactive agents and one affective agent were placed at random safe locations in the environment (that is, at least one field away from any obstacle), in *experiment B* instead only one reactive and four affective agents were used. Finally, in experiment C, five affective agents had to compete with each other, with the main difference that they did not use predefined weights in the affective layer, but rather random connections, which they could alter using Hebbian learning. In this setup, agents that survived for 1000 time steps would reproduce at that point and generate one identical offspring.

The purpose of experiment A was to test whether an affective agent could survive in an environment dominated by reactive agents, and furthermore, if the affective agent would do better than the average reactive agent. The purpose of experiment B was to see whether a small number of reactive agents could survive in an environment dominated by affective agents. The purpose of experiment C was to check whether agents with the same computational resources would do equally well or equally badly, and furthermore, whether associative learning would lead to a configuration in the associative layer similar to the "hard-wired" setup of the affective agents.

Simulations typically start out with some agent consuming the only energy source present, followed by a migration of most agents to the borders of the environment (out of lack of energy sources). Affective agents, being "content and full" in the beginning do not go after food at all, rather they *avoid* energy sources, thereby also avoiding contact with other agents, and thus reducing the risk of bumping into each them. After a while affective agents will get moody and hungry, thus becoming more attracted to food, which eventually leads to very aggressive "scent following" behavior (if an affective agent does not find food in time or has been stuck for quite some time in a particular location). Sometimes this behavior will worsen an agent's situation (and eventually lead to its destruction). By and large, however, affective agents will only compete for food if they are "in a bad mood and hungry", otherwise they will not participate in the competition for energy resources. Reactive agents, on the other hand, maneuver themselves regularly into predicaments because of their constant interest in food, which often forces them into situations where contact with other agents and/or obstacles is inevitable, eventually resulting in the destruction of at the involved agents.

All experiments consisted of 20 simulations. In experiment A, the affective agent survived in about half of the simulations and so did the reactive agent in experiment B. However, only one reactive agent survived in experiment A on the average, whereas *more than two affective agents* survived on the average in experiment B. This is what we should expect: by reducing the risk of being forced into no-win situations, a larger population of affective agents than of mere reactive agents can survive in a given environment. The chances of survival are better for affective agents because of their being able to control their own trajectories through the environment, whereas reactive agents cannot help chasing after food (even if they have enough energy). Experiment C, finally confirmed that using the two hidden nodes as "affective states" (or at least as states that serve the role of affective states or can be interpreted as such) is advantageous—some agents, whose initial random weights in their affective layer reflected a setup similar to the hard-wired affective layer (with variations only in magnitudes and all the signs) would eventually "learn" through experience to avoid food if the energy level was high and seek food aggressively if energy was low. Only such agents eventually survived in experiment C, which implies that the additional computational means of the affective layer are of any use only if they alter the perception of food depending on the inner state. Interestingly, while one of the two hidden nodes always reflected something like the "hunger level", it was sometimes difficult to interpret the role and/or meaning of the other (actually, there were agents where this other state was not used at all, leading to the conclusion that in this particular setup one affective states, namely "hunger" would have sufficed).

6. Discussion

What the experiments confirm, is what seemed intuitively clear from the beginning, namely that avoiding confrontations, unless really necessary, is a better strategy for survival than ignoring the risk that competition for resources bears on its sleeves. Since it is the affective states that regulate the behavior of affective agents, the above experiments confirm the original claim that affective states are an advantage in the competition for resources, and eventually for survival. It was also shown that the additional computational power is only useful if it is used in a setup that resembles the hard-wired affective layer, and thus suggests that the hidden nodes serve the role of affective states.

Note, however, that it can be experimentally confirmed that a single reactive agent shows better performance than a single affective agent (in a slightly altered environment), since once the risk of competition is excluded, the negligence of food of affective agents (when they are not hungry) can haunt them: they might happen to look for at a time, when no food is available, while reactive agents take whatever they can get at any time (thus being implicitly "more foresighted" by storing energy in advance). This is, in my view, further evidence for the biological plausibility of the model: affective agents are best suited for competitive environments, where avoiding threatening competitions is advantageous in the long run; taken out of their "natural habitat" and put in a less favorable environment, their performance will decrease (see also [3]).

Finally, I would like to point to an interesting open problem: the experiment suggests that a large population of affective agents can support a small population of reactive agents—what is the maximal population size of reactive agents that can be supported by a given population of affective agents? My hunch is that a satisfactory answer to this problem might even be able to shed some light on the actual evolution of affective agents (out of mere reactive agents), that is, *on the evolution of affective states.*

Acknowledgements

I would like to thank Norman Pecoraro for discussing with me at length the various findings in animal behavior research on emotional/affective states and their influence on the overall behavior of animals.

References

1. Arkin, R. C.: Behavior-Based Robotics. MIT Press, Cambridge, MA (1998)
2. Arkin, R. C.: Motor Schema-Based Mobile Robot Navigation. International Journal of Robotic Research, Vol. 8, No. 4, (1989) 92-112
3. Breland, K., and Breland, M.: The misbehavior of organisms. American Psychologist, 16, (1961) 681-684
4. Braitenberg, V.: Vehicle: Experiments in Synthetic Psychology. MIT Press, Cambridge, MA (1984)
5. Brooks, R.: Intelligence Without Reason, A.I. Memo No. 1227, MIT AI Laboratory, April (1990)
6. Brooks, R.: A Robust Layered Control System for a Mobile Robot. IEEE Journal of Robotics and Automation, Vol. RA-2, 1, (1986) 14-23
7. Domjan, M.: Formulation of a behavior system for sexual conditioning. Psychonomic Bulletin, 1(4), (1994) 439-450
8. Fanselow, M.S.: Neural organization of the defensive behavior system responsible for fear. Psychonomic Bulletin, 1(4), (1994) 429-438
9. Gardner, R.A. and Gardner, B.T.: Feedforward versus feedback: An ethological alternative to the law of effect. Behavioral Brain Sciences, 11, (1988) 429-447
10. Lorenz, K.: The foundations of ethology. New York: Springer Verlag (1981)
11. Lorenz, K. and Leyhausen, P.: Motivation and Animal Behavior: An Ethological View. Van Nostrand Co., New York (1973)
12. Mochida, T., Ishiguro, A., Aoki, T., and Uchikawa, Y.: Behavior Arbitration for Autonomous Mobile Robots Using Emotion Mechanism. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '95), Pittsburgh, PA, (1995) 516-21
13. Moravec, H.: Mind Children: The Future of Robot and Human Intelligence. Harvard University Press, Cambridge, MA (1988)
14. Pfeiffer, R. and Scheier, Ch.: From Perception to Action: The Right Direction. In: Gaussier, P. and Nicoud, J.-D. (eds.): From Perception to Action, IEEE Computer Society Press, Los Alamos, CA, (1994) 1-11
15. Salomon, R.: Self-organizing Neural Network Controllers for Mobile Robots (1996) (unpublished)
16. Sloman, A.: Architectural Requirements for Human-like Agents Both Natural and Artificial. (1999) (to appear in Human Cognition and Social Agent Technology, Ed. Kerstin Dautenhahn)
17. Timberlake, W.D. and Lucas, G.A.: Behavior systems and learning: From Misbehavior to general principles. In: Klein & Mowrer (eds.): Contemporary Learning Theories: Instrumental Conditioning and the Impact of Biological Constraints on Learning, Lawrence Earlbaum Assoc., Hillsdale, NJ (1989)
18. Tinbergen, N.: The study of instinct. Oxford: Oxford University Press, Clarendon Press (1951)