916

Studying Emotions: Fungus Eaters

Rolf Pfeifer
AI Lab, Institute for Informatics
University of Zurich-Irchel
Winterthurerstrasse 190, CH-8057 Zurich
Switzerland
e-mail: pfeifer@ifi.unizh.ch

April 13, 1993

Abstract

A topic that persistently re-emerges in the context of animats and adaptive behavior in general, is the one of emotion. Although the area has a history which is nearly as long as the one of artificial intelligence and robotics progress has been relatively modest. Many researchers in the animats field share the conviction that at some point emotions will be important, but nobody has been capable of pinning down exactly where this "point" would be. In this paper, based on a synopsis of the field, it is argued that one of the reasons for the lack of progress is the fact that emotions are investigated "directly". Using a case study of autonomous robots it will be demonstrated that if we are to make progress a better strategy is to view emotions as emergent and to employ an incremental approach which we have dubbed the "New Fungus Eater Approach". There is a striking similarity of the problems of the psychology of emotion and those of artificial intelligence. It is concluded that designing and building animats may be a better road to understanding emotion than traditional modeling approaches. This requires, however, that the approach be based on *value systems* and *self-organization* rather than traditional modeling concepts from artificial intelligence. It is shown how some of the unresolved controversies of the field of emotion "disappear" if the basic issues are approached in the right way. We conclude the paper by speculating about the utility of emotional considerations in designing animats.

Keywords: emotion, Fungus Eaters, self-organization, value systems

1 Introduction

Many researchers in the animat field[1] are convinced that emotions are important for understanding self-organization and living beings (humans and animals) and that at some point we will have to take emotions into account when designing animats. The importance of the topic for understanding intelligent (adaptive) behavior has been pointed out early on by Herbert A. Simon in his seminal paper on "Motivational and emotional controls of cognition" (Simon, 1967). However, there is little agreement on *how* ideas about emotion should influence designs and what would be the exact benefits of including them. This uncertainty is also reflected in the relatively sparse density of such papers at conferences and in pertinent journals.

In this paper we explore two questions, namely how designing and building animats can enhance our understanding of emotions, and how considerations about emotions can guide our research into designing better robots. The answers to the former are well-grounded in empirical evidence as well as theoretical arguments, whereas the treatment of the latter will be more speculative.

2 Motivation

In the context of adaptive behavior there are two main reasons for being interested in emotion: (a) to better understand emotion since this might be helpful in understanding adaptive behavior, and (b) since understanding emotion and applying pertinent insights might lead to better designs. Point (a) is obvious: building models of any sort has always been an excellent tool in the empirical sciences to enhance understanding. Animats are particular examples of models which are well-suited to studying adaptive behavior. Point (b) involves two aspects.

First, task performance might be enhanced, i.e. robots which include "emotions" one way or other in their designs might perform better on their tasks. Second, communication between the robot and humans might be improved by endowing the robot with knowledge of emotion. This would enable the robot to *reason* about emotions which is why we call such programs or robots "reasoners". Knowing that John is angry sets up expectations about his potential behavior and those expectations are very different if John is happy. One might well imagine that highly sophisticated robots (which we currently do not have) could benefit from being able to reason about emotions (of themselves, other robots, or humans). If they could recognize emotions in humans and could also display appropriate emotions in human-robot interaction the interaction between the two might be improved.

3 The concept of emotion

Since this paper is about emotion we should somehow delineate the sort of thing we are talking about. However, the large number of definitions which can be found in the literature is convincing evidence for the ill-definedness of the concept of emotion. Kleinginna and Kleinginna (1981), for example, list nearly a hundred different definitions and we don't think we should add another one to the list. This definition problem may also be one of the reasons why there are quite a few important controversial issues around in the field of emotion. It is symptomatic for some underlying

[1]The term animat is used in the sense of Wilson (1992).

problems which will be explored in the following sections. Here are a few examples of largely unresolved controversies:

- components: can emotions be decomposed into "components" (like cognitive-evaluative, physiological, expressive-communicative, subjective experience, action)?
- basic emotions: is there a (relatively small) set of basic emotions?
- patterning: is there a general level of (physiological) arousal or are there emotion-specific patterns?
- universality: are emotions universal or culturally specific?
- function: what is the function of emotions?
- generation: how do emotions come about?
- intensity: what are the factors influencing the intensity of emotions?

and so on, and so forth. Most researchers take positions on these issues and these positions are largely incompatible with each other. This is the reason why there are so many different "schools of thought". Given the current developments in the field of emotion psychology (and physiology) consensus does not seem to be emerging. We will argue later that in fact these controversies *need not be resolved:* if the approach is appropriate they largely "disappear" or are resolved automatically.

One of the reasons why consensus is diffucult to obtain is that most existing theories are phenomenon-based. This needs some elaboration. Sloman (1991) makes a distinction between three types of theories, semantic-based, pheonomenon-based, and design-based. Semantic-based theories focus on how people use emotion *terms* from natural language. Phenomenon-based theories start from the (tacit) assumption that there is agreement on when emotion "occurs" and then cognitive antecedents, correlates (physiological, facial expressions, cognitive, subjective experience) and consequences are measured. While this may be a starting point it is clearly not sufficient for a scientific theory. Design-based theories, in essence, focus on how systems must be designed in order to produce certain behaviors. In other words, the appropriate mechanisms have to be worked out. Within the latter there are again different "sub-cultures" (top-down, bottom-up, and middle-out). We feel — and we will argue accordingly — that the focus must be on design- based theories, otherwise the above-mentioned controversies cannot be resolved.

So we have not defined emotion for the reasons just given. But we can use our prescientific understanding — however vague this may be — to delineate the kinds of of behaviors we are interested in and which we are trying to explain.

4 Models of emotion

The goal of this section is to provide some background on existing modeling efforts pertaining to emotion. This will serve as a starting point for our criticism of (most of) these approaches and provide the motivation for suggesting a radically new approach. Since there are to date only very few models which concern directly the animat domain we discuss some models in an immediately related field, namely AI to make things concrete.

emotion	target	pos/neg	causality	control
pride/self-esteem	self	+	self	by self
anger	other	-	other	by other
gratitude	other	+	other	by other
guilt	self	-	self	by self
pity	self	-	other/env	uncontr.

Figure 1: Taxonomy of emotions. "+" stands for positive, "-" for negative evaluation.

We have already mentioned the "reasoners". They are typically based on so-called emotion taxonomies, i.e. schemas for classifying emotions with respect to emotion-eliciting situations. These situations are characterized along certain general dimensions. A simple taxonomy is given in Figure 1. Let us pick out just one example, anger. The situation is negative for the agent, the target (who is the focus of the emotion?) is the "other", the locus of causality (who caused the undesirable state?) is the "other", and the locus of control (was the situation controllable and who had control over it?) is also the "other". Of course, we are not dealing with the "real" situation but always with the agent's beliefs.

Programs capable of reasoning about emotions have proved quite successful, in particular in the area of natural language processing. But more important than reasoners are approaches which focus on mechanisms underlying emotion (or better what we would call emotional behavior). An early example is Abelson's "Hot Cognition" (Abelson, 1963), a computer model of attitude change. Emotional factors were introduced because of the difficulties of explaining the phenomena involved in purely cognitive terms. And these are precisely the situations in which one would sensibly resort to emotional explanations: whenever no other construct would suffice to explain the behavior under investigation.

A more recent model is FEELER (Pfeifer, 1982; Pfeifer, 1988; Pfeifer and Nicholas, 1985). It's purpose was to specifically model emotional processes. The simulations worked essentially as follows. The model which can be conceived of as an agent, was presented a (precoded) situation like sitting in a taxi on the way to the airport. Then a particular event would be introduced such as the taxi having a flat tire. The agent (i.e. the model) would try to classify the situation according to the dimensions given in a taxonomy and then apply the appropriate emotion identification rules. This would lead to a particular representation of the situation with associated emotional labels, physiological patterns, and causal links to the eliciting event or events. The classification or analysis process was triggered by a so-called interrupt (Simon, 1967; Mandler, 1985). In other words, emotional processes were initiated when some organized set of actions of the agent was interrupted.

Although there are numerous other models of emotion that differ significantly — e.g. they implement different theories of emotion (i.e. they are from a different school of thought) — they all suffer from some of the same underlying problems as illustrated in the next section. So, for the purposes of the present discussion these examples should suffice. Readers interested in more detail on the topic are referred to the review by Pfeifer (1988).

5 Problems with models of emotion

Emotion modeling has had some beneficial effects on AI. For example it has drawn our attention to behaviors which are normally not investigated in AI. But the successes have been rather limited. This limited success is largely due to some fundamental unresolved issues in the study of emotion. Very briefly they have to do with the fact that in virtually all of the models *emotion is engineered into the model* whereas emotion ought to be viewed as *emergent.* Here are some more details.

The study of emotion as an isolable faculty and the problem of reduction: There are many different definitions of emotion and there is little evidence that there will be any convergence of ideas in the near future and that the controversies mentioned earlier will ever be resolved. Emotion, whatever it is, seems to be elusive, multifacted, and extremely complex. This leads one to suspect that it is in fact not a "thing", a "faculty" or system component which can be clearly delineated and studied as such but rather that it is *emergent.* If a phenomenon is emergent its investigation must go below this level of description in order to find explanatory accounts. A similar point has been made by Steels (1990) in the context of robotics. It applies also to much of AI where the attempt to model intelligence directly has also been met with unexpected difficulties (e.g. Brooks, 1991).

To illustrate the point let us look at how anger is dealt with in FEELER. Anger is a behavioral phenomenon, i.e. it concerns an interaction of an agent with its environment. There is some agreement on when it occurs. In the model there are rules for identifying anger and nodes representing situations involving anger. In other words, it assumes a direct causal link between mental states (as described by rules and nodes) and behavioral states (the observation). This is an instance of what Clancey (1991) has termed the "frame-of-reference problem". The main purpose of Clancey's discussion is to conceptualize the relation between observer, observed agent, artifact, and the environment. One implication of the "frame-of-reference problem" is that behavior cannot be reduced to internal processes. This requires some elaboration.

The emotional stance: In some respects emotions are like goals, beliefs, and knowledge. They are attributed to an agent by an external observer. Dennett (1971) has used the term "intentional stance". In that sense, emotions, just as goals and knowledge, are intentional-level constructs. One might also talk about an *emotional stance.*[2] While intentional-level constructs are extremely powerful in that they enable an observer to make sense of an agent's behavior and to perform counterfactual reasoning (e.g. Pylyshyn, 1984) we have to be careful not to confound observer-based attributions with mechanisms. Looking at the models discussed in the previous section, for example at FEELER, it seems that in most of them there is somewhat of a confusion between an intentional-attributional level (which is to do with behavior, i.e. an agent-environment interaction) and the underlying mechanisms: emotional terms, corresponding to the intentional or emotional stance, are reduced to internal processes which amounts to a category error (Clancey, 1991).

Connection with bodily activities: A particular problem with most models of emotion is their inability to deal appropriately with physiological activities. One of the few models making a serious attempt is FEELER. But most psychologists agree that physiology is essential to behavior that we tend to call emotional. Some researchers even see physiology as the primary factor in these types of phenomena.

[2] I owe this idea to Hugues Bersini of the Université libre de Bruxelles.

This reflects a general point about computer science, namely the difficulties in dealing with anything physical, anything having to do with the *real* world in contrast to a *virtual* one. While traditional AI can be considered successful in "virtual" domains (formal games, logic, abstract problem solving) it is less successful in dealing with the real world, although intelligence has a lot to do with interacting with the real world. In the field of "New AI" (or "Nouvelle AI", Brooks, 1991) this has been recognized and the research is focused on the physical instantiation of intelligent systems.

In conclusion, there seem to be a number of strong commonalities between the area of modeling intelligence and modeling emotion and it is therefore not really surprising that the problems encountered are also highly similar. This suggests that a similar methodological innovation may be needed. In AI this innovation could be along the lines of "New AI" (or animats) (e.g. Brooks, 1991; Pfeifer and Verschure, 1992b). We suggest to use the same approach for the study of emotion (or more precisely behavior that we tend to relate to emotion).

6 "The New Fungus Eater Approach"

The starting point for the strategy we propose for studying emotion is essentially not to study emotions *directly*. If a phenomenon is emergent its investigation must go below the level at which the phenomenon is described. More concretely the suggested strategy is to design and build "Fungus Eaters"[3], rather than trying to study emotions directly. "Fungus Eaters" are integrated autonomous systems capable of behaving in real-world environments. They are equipped with everything needed to survive and perform a particular task. They have an energy store and they feed on a special kind of fungus (hence the term "Fungus Eater"). They are sent to a distant planet to collect uranium ore. They have to collect as much ore as possible while maintaining sufficient energy store and avoiding predators. An energy store of zero means death. Any activity consumes energy, even just sitting still and deliberating. This creates a *need to act* for the "Fungus Eater". The idea to study simple but "complete" creatures has been around for a long time. An example are the turtles of Grey Walter (e.g. Walter, 1951). But for several reasons this approach has not been pursued to a significant extent.

The "New Fungus Eater Approach"[4] builds on an approach for developing autonomous agents, called "Distributed Adaptive Control", or DAC for short, which we developed in our group. The details of this approach have been discussed elsewhere (Pfeifer and Verschure, 1992a; Verschure, Kröse and Pfeifer, 1992) and will not be elaborated here. It consists in designing "Fungus Eaters" (robots), running experiments, analyzing the behavior of the robots, relating the experiments to the literature on emotion, and gradually increasing the complexity of the tasks and of the environment. [5]

[3]The term "Fungus Eater" was created around 1960 by the Japanese Psychologist Masanao Toda who suggested the study of "Fungus Eaters" as an alternative methodology to traditional cognitive psychology (see Toda, 1982, for a collection of papers on the topic).

[4]The approach is called "New" to distinguish it from the now 30 year-old "Fungus Eater" of Masanao Toda.

[5]The work reported here is based mainly on simulations, but we have implemented our architecture on a number of different physical robots.

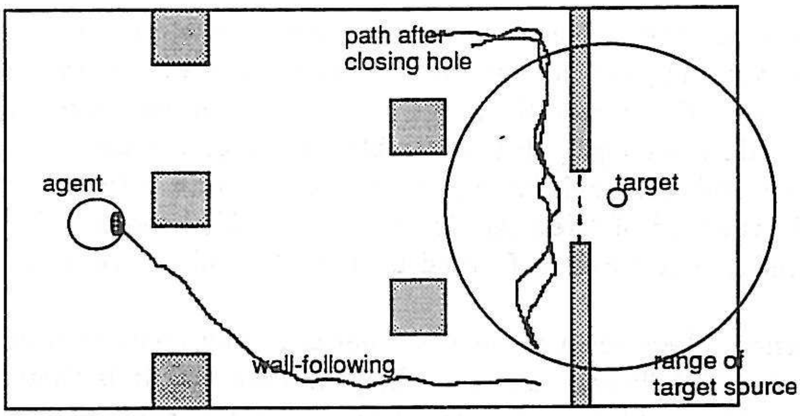

Figure 2: The "Fungus Eater" environment (traces redrawn manually for clarity of presentation).

6.1 Designing the robot and its environment

The robot is a fairly standard. It has a number of sensors (collision detector, range finder or some kind of visual sensor, target detector) and a means of locomotion. The environment shown in Figure 2 is also fairly standard. In addition to a number of objects there is an opening in the wall and behind this opening there is an attractive force (e.g. a food source) which can be detected by the target sensor. The traces will be explained later.

One of the assumptions underlying our approach is that biological systems are self-organizing and that the process of self-organization must be guided by a value system (e.g. Edelman, 1989).

Value system: The value system (or value scheme) encodes what matters to the robot. Examples of values of the "Fungus Eater" are: "being close to targets is better than being further away", "avoiding obstacles is better than hitting them", etc. Important values also pertain to keeping the essential variables within a certain range such as maintaining an energy level (or Fungus store).

In our "Fungus Eater" values are not explicitly encoded but rather *implicitly* in the reflexes, the architecture of the network, the properties of the sensors and effectors, and the learning mechanisms. The value system is predefined by the designer and is considered fixed for the period of the experiments.[6] It contains the following elements. There are a number of "sense-act-reflexes", e.g. "whenever a collision occurs, reverse and turn in the other direction", or "whenever a target is detected turn towards the target". There is also a default action which is simply to move straight ahead.

The signals from the sensors have to be translated into inputs for the neural architecture by transducer functions. Together with the morphology of the robot, i.e. its size, the location of the sensors and effectors, the transducer functions are needed to physically connect the control architecture with the environment.

Self-organization: The value system serves the purposes of providing a basic kind of coarse adaptivity of the organism to the environment and of directing the process of self-organization. The self-organizing process is implemented via a neural network architecture (Figure 3). Each sensor has its own layer. The connections between the collision and target layer and the motor layer are hard-wired: they implement the

[6]This does not imply that no higher-level values can develop on top of the predefined ones, but we have not yet conducted the pertinent experiments.

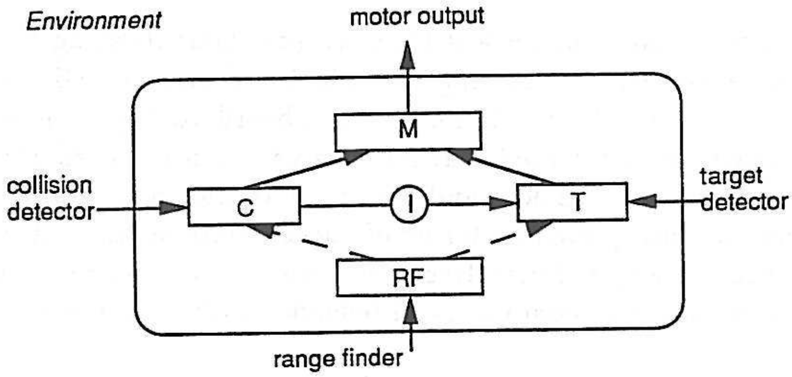

Figure 3: The control architecture

basic reflexes. The sensory layers are connected via modifiable connections (through Hebbian learning with active forgetting). Hebbian learning takes place whenever the nodes in the collision or the target layer are sufficiently active to trigger an action. In other words the learning is contingent on the actions. The inhibitory element which connects the collision and the target layers is also part of the value system: it encodes the preference of avoidance over approach.

6.2 Experiments

We have performed a number of simulation experiments which we have described in great detail elsewhere (e.g. Verschure *et al.*, 1992). What happens if the robot is released in the environment depicted in Figure 2 is the following: It will initially hit obstacles. Every action (reverse and turn) will enable Hebbian learning between the collision layer and the range finder (or visual) layer and thus the "Fungus Eater" eventually learns to avoid obstacles. If there is a target source behind an opening in the wall it will eventually evolve a wall-following behavior even in the absence of targets (see Figure 2). This behavior is emergent — it has not been programmed into the system; there is no "component" responsible for the wall-following. It is the result of an interaction between the avoidance and approach behaviors when the "Fungus Eater" is in the vicinity of the target source (the "Fungus Eater" can only sense the target if it is within its range, indicated by a circle in Figure 2).

6.3 Characterizing the behavior of the "New Fungus Eater"

There are several ways in which we can try to characterize the behavior of the "Fungus Eater". Here we focus on two, namely (a) in terms of internal processes, (b) in intentional-level terms, and (c) in emotional terms. This is compatible with points made in the context of the "frame-of-reference" problem.

(a) is used to account for the robot's behavior in terms of how the network mechanisms lead to particular behaviors through the dynamics of the system-environment interaction. We have briefly mentioned how obstacle avoidance comes about, namely through a Hebbian associative process. The constraints are provided on one hand by the value system (learning is triggered when reflex-related actions are triggered), and on the other by regularities in the environment: whenever the collision detector hits an obstacle there will also be systematically related activation in the range finder (or the visual sensor). Similarly the wall-following behavior can be explained (left as an exercise to the reader).

(b) and (c) are similar and we will focus on (c), characterizing the robot's behavior in emotional terms, i.e. adopting the "emotional stance". Given the extreme simplicity of our "Fungus Eater" it may seem absurd to apply emotional terms. However, it has been demonstrated that humans very quickly attribute emotions to objects. There is a movie by Helder and Simmel with moving squares and triangles which many viewers interpreted in terms of social relations and interactions, and emotion terms such as anger, frustration, etc. This means that we have to be careful not to confound *apparent* complexity of behavior with complexity of mechanism (Simon, 1969).

Let us do a little experiment. Its result is shown in Figure 2. Let the "Fungus Eater" explore its environment for a while until it starts following walls. Then close the hole in the wall so that it can no longer get near the center of the target (this is indicated by the dashed line). This is inspired by our (prescientific) ideas about frustration. We observe that the "Fungus Eater" follows the wall. Using a combination of intentional-level and emotional terms, we could say that it follows the wall because it *hopes* or *expects* to find a target since it "knows" that targets are behind openings in walls — remember that we are talking about behavior, not about internal mechanisms. As it has gone past the center it realizes that it is moving away from the target. It turns around, tries again, and finally leaves the target. We could use the term *resignation* to describe why it no longer stays near the target. But clearly there is neither a component for expectancy or hope, nor one for resignation in the system, as there would have to be (and actually was) in a system like FEELER. What accounts for the apparent frustration behavior in terms of the dynamics of the system-environment interaction is the sluggishness of the inhibitory element: if approach is inhibited often enough activation will accumulate in the inhibitory element which implies that approach will eventually be inhibited long enough so that avoidance can take over. This in turn enables the "Fungus Eater" to leave the target range.

In this example the striking difference between a characterization of behavior in emotional terms and the internal machinery employed to achieve the behavior becomes obvious: the description of the behavior might suggest something highly sophisticated or complex in terms of internal processes whereas the mechanisms employed are extremely simple. But does the "Fungus Eater" actually *have* emotions? The example leads one to think that this might not be a sensible question because it is by no means obvious what *having* and emotion would mean here. A more sensible question to ask is: given a certain behavior that we find interesting (for example because it reminds us of emotional behavior) what are the mechanisms that could have produced it?

6.4 Discussion: Studying emotions

How does this approach contribute to the study of emotions?

First, through the notion of the value system all the processes occurring within the agent are tightly integrated: processes taking place in the neural architecture only make sense with respect to the value system. In other words, there is no clear separation between "cognitive" and "physiological" processes. This is precisely what is needed to explain behaviors that we tend to call emotional. But it is interesting to note that the same holds for intelligent behavior in general.

Second, the integration of a value system provides the foundation on top of which higher-level values can develop. This idea captures the intuition that what is considered emotional behavior somehow relates to "values".

Third, there is no need to define emotion, in fact it would be a mistake to do so. By successively increasing the complexity of the "Fungus Eater" and its environment, for example, by introducing other "Fungus Eaters", we will eventually arrive at an understanding of the relation between behavior and internal processes that will enable us to clearly identify those processes which lead to behaviors that have traditionally been associated with emotions. In other words we will have a *theory* of "Fungus Eaters".

Fourth, the "New Fungus Eater" approach provides a convenient platform for exploration and systematic experimentation.

Fifth, some of the basic unresolved controversies are automatically resolved, others become irrelevant. For example physiological processes, i.e. processes concerned with the physical setup, are automatically part of all activities of the agent not only of those that we might want to call emotional. Moreover, the very fact that emotions are emergent implies that there can be no set of basic emotions. Also, because emotion is not considered a "component" some questions can no longer be asked. But they can be replaced by more constructive ones about mechanisms and control architectures. As a last example, we will be able to say under what circumstances processes will be triggered which one might want to call emotional. This point pertains to the issue of "emotion generation".

A caveat: While this may sound plausible it seems that the causal nature of emotions is somehow lost in this approach. After all, we do know that our behavior is influenced by emotions. But we have to be careful on the one hand not to confound our own conceptualizations of complex states with underling mechanisms and on the other not to resort to our prescientific understanding for explanation (see our discussion above).

6.5 Designing agents

We have now argued in detail why studying "Fungus Eaters" is a productive approach for studying emotions. This may be an interesting endeavor to a psychologist but to an engineer it is only relevant if it can be shown that knowledge of emotions will, at least potentially, lead to better designs. In what ways can we expect the "Fungus Eater" approach to lead to better designs?

The answer to this question is not straightforward. What we are interested in are self-organizing systems. If the goal is to eventually be able to account for behaviors that we want to call emotional the relation between control architecture and physical set-up will have to be carefully taken into consideration. Since value systems provide precisely this connection it is natural to use designs which are based on value systems and self-organization, rather than goal- or plan-based approaches as in more traditional robotics.

The question of endowing the agent with knowledge *about* emotion is really a different issue. As mentioned earlier this might make for a smoother interaction of humans with robots (or perhaps between robots). In fact there are severe doubts about that since this knowledge, if programmed by a human designer (and not acquired by the agent itself), would not be grounded and could therefore not be applied appropriately.

7 Conclusions

We have outlined a new approach to study emotions, the "New Fungus Eater Approach" and we have argued why we believe that it will be productive. However, when working with a bottom-up approach there is, of course, always the question of whether the approach will scale up to capture more complex behaviors. This is an empirical question which is completely open at the moment. But we prefer to live with this uncertainty rather than working with approaches that have been shown to be severely limited.

References

[1] Abelson, R.P. (1963). Computer simulation of "hot cognition". In: S.S. Tomkins, and S. Messick (eds.). *Computer simulation of personality.* New York: Wiley, 277-298.

[2] Brooks, R. (1991). Intelligence without reason. *Proc. IJCAI-91,* 569-595.

[3] Clancey, W.J. (1991). The frame of reference problem in the design of intelligent machines. In: K. van Lehn (ed.). *Architectures for intelligence. The 22nd Carnegie Mellon Symposium on Cognition.* 357-423.

[4] Dennett, D.C. (1971). Intentional systems. *Journal of Philosophy,* **8**, 87-106.

[5] Edelman, G.M. (1989). *The remebered present.* New York: Basic Books.

[6] Kleinginna, P.R., Jr., & Kleinginna, A.M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion,* **5**, 345-379.

[7] Mandler, G. (1985). *Mind and body. Psychology of emotion and stress.* New York, N.Y.: W.W. Norton.

[8] Pfeifer, R. (1982). Cognition and emotion: an information processing approach. Carnegie Mellon University, CIP Working Paper Nb. 436.

[9] Pfeifer, R. (1988). Artificial intelligence models of emotion. In: V. Hamilton, G. Bower, & N. Frijda (eds.). *Cognitive perspectives on emotion and motivation. Proc. of a NATO advanced research workshop.* Kluwer, 287-319.

[10] Pfeifer, R., and Nicholas, D.W. (1985). Toward computational models of emotion. In L. Steels, and J.A. Campbell (eds.), *Progress in Artificial Intelligence.* Chichester, U.K.: Ellis Horwood, 184-192.

[11] Pfeifer, R., and Verschure, P.F.M.J (1992a). Distributed adaptive control: a paradigm for designing autonomous agents. In F.J. Varela, and P. Bourgine (eds.); *Toward a practice of autonomous systems. Proc. of the First European Artificial Life Conference*, 21-30. Cambridge, Mass.: MIT Press/Bradford Books.

[12] Pfeifer, R., & Verschure, P.F.M.J. (1992b). Beyond rationalism: symbols, patterns, and behavior. *Connection Science,* 4, 313-325.

[13] Pylyshyn, Z.W. (1984). *Computation and cognition.* Cambridge, MA.: MIT Press-Bradford Books.

[14] Simon, H.A. (1967). Motivational and emotional controls of cognition. *Psychological Review,* **1**, 29-39.

[15] Simon, H.A. (1969). *The sciences of the artificial.* Cambridge, Mass.: MIT Press.

[16] Sloman, A. (1991). Prolegomena to a theory of communication and affect. In: A. Ortony, J. Slack, and O. Stock (eds.). *AI and Cognitive Science Perspectives on Communication.* Heidelberg: Springer, 1991/92.

[17] Steels, L. (1991). Towards a theory of emergent functionality. In J.-A. Meyer, & S.W. Wilson (eds.) *From animals to animats. Proceedings of the First International Conference on The Simulation of Adaptive Behavior.* Cambridge, MA: MIT-Press/Bradfort Books.

[18] Toda, M. (1982). *Man, robot, and society.* The Hague: Nijhoff.

[19] Verschure, P.F.M.J., Kröse, B.J.A., and Pfeifer, R. (1992). Distributed adaptive control: the self-organization of structured behavior. *Robotics and Autonomous Systems,* **9**, 181-196.

[20] Walter, W.G. (1951). A machine that learns. *Scientific American, 183,* 60-63.

[21] Wilson, S.W. (1992). The animat path to AI. In J.-A. Meyer, and S.W. Wilson (eds.) *From animals to animats. Proceedings of the First International Conference on The Simulation of Adaptive Behavior.* Cambridge, MA: MIT-Press-Bradford Books.