Challenges for Complete Creature Architectures

Rodney A. Brooks
MIT Artificial Intelligence Laboratory
545 Technology Square
Cambridge, MA 02139, USA
brooks@ai.mit.edu

Abstract

In recent years there has been a move within the artificial intelligence and robotics communities towards building complete autonomous creatures that operate in the physical world. Certain approaches have proven quite successful, and have caused a re-analysis within the field of artificial intelligence of what components are necessary in the intellectual architecture of such creatures. However nothing built thus far yet comes close the dreams that many people hold dearly. Furthermore there has been quite some criticism of the new approaches for lacking adequate theoretical justification. In this paper we outline some of the more obvious challenges that remain for these new approaches, and suggest new ways of thinking about the tasks ahead, in order to decompose the field into a number of manageable sub-areas that can be used to shape further research.

1 Introduction

There is a growing interest in building artificial creatures of some sort. One example is the recent boom in a field known as Artificial Life (see [Langston 87] and [Langston 90]). While much of the emphasis is on building forms resident in computers, which are agents acting in an information domain, there has also been some interest in physical embodiments of artificial creatures.

This author, at the MIT Al Lab, introduced the subsumption architecture ([Brooks 86] and extended in [Brooks 90]) with the explicit goal of building mobile robots with long term autonomy. Later the word creature crept into the language of the MIT group (e.g., [Connell 87]). The goal is to build autonomous mobile robots which operate over long periods of time, completely autonomously, in dynamic worlds. It is envisioned that these worlds are worlds which already exist for some other purpose-not worlds specially built to house the robots. Further, it is envisioned that these robots carry out some task which has some utility for whoever wanted the robots to exist and live in this world.

As [Flynn 87] points out, there are many components to such creatures, including sensors, actuators, power sources, and intelligence. Over the last five years we have found that all these components are intimately related as we have tried to build prototype creatures ([Flynn and Brooks 89]). Choices in any part of the system architecture (e.g., sensor characteristics) have major impacts upon other parts of the system. In general it is very dangerous to think that any one component (such as intelligence) can be isolated and studied by itself.

Our experience with the subtleties of such interactions has led us to our current construction of a very complex robot, named Attila ([Angle and Brooks 90]) Pictured in figure 1 (in fact we are building multiple copies of Attila). It has six legs, each with three degrees of freedom. an active whisker, a gyro stabilized pan-tilt head carrying a range finder and a CCD camera, 10 onboard processors, and over 150 sensors. We built an earlier six legged robot named Genghis ([Angle 89], [Brooks 89], but its complexity pales in comparison to that of Attila. Many of the issues raised in this paper were brought to our attention as we have tried to work out how to program this complex robot Attila to be an artificial creature.

The bulk of this paper is devoted to the problems and challenges in designing and building the computational architectures for such creatures. However, the reader should not forget that the other aspects of a creature's architecture cannot be considered in isolation from intelligence. in a complete design, all aspects greatly influence each of the others.

We first argue that there are multiple levels of analysis or abstraction with which we must be concerned in designing and building complete creature architectures. There can be no single magic bullet or theory which will tell us all we need to know. Some problems within these levels are well circumscribed and so can be worked on in isolation. However, in order to build complete creatures we need to bridge the gaps between these levels also.

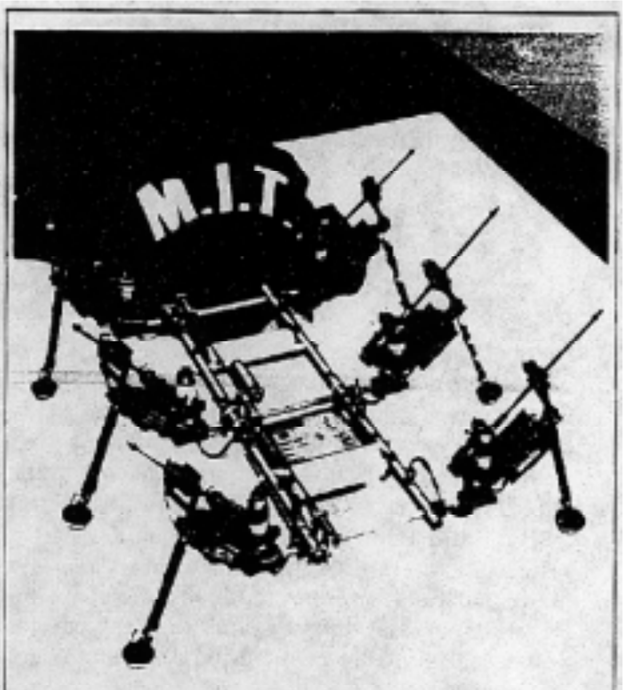

Figure 1: A partially constructed Attila robot. It has six legs, 24 motors, 10 processors, and over 150 sensors. Still missing at the time of this photograph are a pan-tilt camera mount, gyroscopes, and an active steerable whisker.

The bulk of the paper then goes on to examine each level of analysis from three points of view: theory, relations, and practice In terms of challenges which must be faced in building complete creature architectures, these can be described as:

Theory: What theoretical tools can be developed, and how will they be useful?

Relations: What existing bodies of work in other disciplines can be drawn upon to assist with analysis and synthesis at this particular level of abstraction?

Practice: How can we go about practically building systems, with or without the help of future developments in relevant theory, or related work?

Finally the paper outlines a tentative agenda for development of complete creature architectures.

2 Levels of Analysis

Traditional science has found it necessary to relate complex biological systems to fundamental properties of matter in terms of a number of levels abstraction. In the coarsest terms these levels correspond to physics, chemistry, and biology. Traditional science tends to work within these boundaries. It is impossible to do good science without having an appreciation for the problems and concepts in the other levels of abstraction (at least in the direction from biology towards physics), but there are whole sets of tools, methods of analysis, theories and explanations within each discipline which do not cross those boundaries.

While there are newer disciplines such as physical chemistry, bio-chemistry, and molecular biology which try to bridge the gaps between the major disciplines there is still no definitive crossing of the major boundaries except in the simplest of cases, and then only very recently. For instance, the best that has been-achieved in terms of explaining a chemical reaction in terms of the physics of mass. force and quanta, is (as reported by [Pool 90]) a fairly complete understanding of the reaction $D + H_2 \rightarrow HD + H$ (where D is deuterium). This is just about the simplest possible reaction in terms of the number of fundamental particles involved.

As science progresses we can expect a more unified theory to emerge with more and more bridges between the traditional disciplines being built to the point that the abstraction barriers are completely understood. But that is certainly not the case at this point in history, despite the enormous body of scientific and philosophical work that has been undertaken within the realm of modern science.

As we try to build artificial creatures there is no reason to expect that we should instantly be able to come up with a uniform theory or set of tools of analysis and synthesis that cover all problems which need to be examined In fact, there seems a quite natural decomposition of the task into three somewhat distinct levels of abstraction:

Micro: The study of how distinct couplings between the creature and its environment are maintained for some particular task. Example tasks we temperature regulation, wall following, climbing, etc. Primary concerns here are what forms of perception are necessary, and what relationships exist between perception, internal state, and action (i.e., how behavior is specified or described).

Macro: The study of how the many micro perceptions and behaviors are integrated into a complete individual creature. Primary concerns here are how independent various perceptions and behaviors can be, how much they must rely on and interfere with each other, how a competent complete creature can be built in such a way as to accommodate all the required individual behaviors, and to what extent appearingly complex behaviors can emerge from simple reflexes.

Multitude: The study of how a multitude of individual creatures interact as they go about their business. Primary concerns here are the relationships between individuals' behaviors, the amount and type of communication between creatures. the way the environment reacts to multiple individuals, and the resulting patterns of behavior and their impacts upon the environment (which would not occur in the case of isolated individuals).

These levels or fields by no means correspond to physics, chemistry, and biology, but there is a certain metaphorical connection in terms of them being somewhat distinct levels.

3 Micro Level

Any artificial creature which we build must be able to interact with its environment. At any particular time, an observer might say that a creature is doing some particular task. or doing many particular tasks in parallel. Examples of such tasks in the robot domain include wandering, avoiding obstacles, wall following, looking for a recharge station, delivering some object, cleaning the floor, following someone, etc. We will refer to what the robot is doing when it is achieving some task as carrying out a behavior.

At the micro level we will be concerned with one particular such task at a time, although the creature may really be carrying out many in parallel, The particular challenges at the micro level are as follows.

Convergence: Demonstrate or prove that a creature is programmed in such a way that its external behavior will indeed achieve a particular task successfully. For instance, we may want to give some set of initial conditions for a creature, and some limitations on possible worlds in which it is placed, and show that under those conditions, the creature is guaranteed to follow a particular wall, rather than diverge and get lost.

Complexity: Deal with the complexity of real world environments, and sift out the relevant aspects of received sensations rather than being overwhelmed with multitudes of data.

Synthesis: Given a particular task, automatically derive a program for the creature so that it carries out that task in a way which has clearly demonstrable convergence.

3.1 Micro Theory

Suppose we are to have some theory about how a creature interacts with the world. The creature both senses the world and acts, to change its own position, and perhaps to change the world itself. The world may also change for other reasons. As in figure 2, any observer of the system must sense both the creature and the world. A theory may be couched in terms of objective truths, but in order for us to test a theory we must compare it with observations made by an observer. In fact, no theory can be given complete objective truth. it must be given approximations to the true situation supplied by some observer.

The creature itself acts as an observer of the world. It has sensors which provide the only inputs it can have concerning the world. The first question is what should the creature do with the sensor inputs? How should they be organized and manipulated?

We argue below that the traditional idea of building a world model, or a representation of the state of the world is the wrong idea. Instead the creature needs to process only aspects of the world that are relevant to its task. Furthermore, we argue that it may be better to construct theoretical tools which instead of using the state of the world as their central formal notion, instead use the aspects that the creature is sensing as the primary formal, notion.

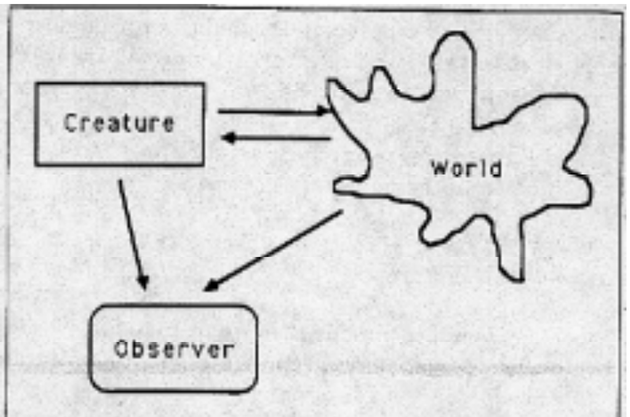

Figure 2: The creature and the world interact. The observer or modeler watches them both.

3.1.1 Uncertainty

In the early days of robotics it was noticed that uncertainty is inherently present when sensors are used to glean information to model the real world (e.g., [Roberts 63]). The early approaches to this problem were either to ignore it (e.g., [Lozano-Pérez 76]) or to engineer the situation so that it could be. ignored (e.g., [Nilsson 84] and [Giralt, Chatila and Vaisset 84]; both are reports on earlier work).

Later attempts were made to explicitly model the uncertainty in the sensors (e.g., [Moravec 80]), and hence in the world models built. In the manipulator domain, others (e.g., [Taylor 76] and [Brooks 82]) tried to reason over these uncertainties, and eventually a whole sub discipline grew up wound planning for manipulator motions in the presence of uncertainty ([Lozano-Pérez, Mason, and Taylor 84] provides a foundational paper in this area).

But why is uncertainty a problem, and what is the problem if there is one?

> *The key problem is in trying to build a model of the world.*

By building a model, we mean that the system tries to have some internal representation of an external objective reality. The problem is in correlating the current readings from the sensors with the existing (partial) world model. The original readings used to build the existing model were noisy and introduced uncertainties in the representation of the world. The new readings also include noise. Furthermore, if the robot has moved between sensor readings then there is uncertainty in how the coordinate systems of the two (or more) sets of sensor readings are related. [Moravec 80] and [Chatila and Laumond 85] describe the two fundamental approaches to these problems. If more than one type of sensor is used, there is also the problem of fusing the different classes of data into a single representation on.

If there are no models built, the problem of uncertainty is inherently reduced. This alternative is to operate in a tight coupling with the world through a sensingacting feedback loop. Instead of relying on inaccurate values returned by noisy sensors, we can rely on the time averaged derivative of these signals as the creature actively changes its state within the world in a way which forces larger changes in the sensor readings than those contributed by noise.

This approach has actually been used by people working in Artificial Intelligence for many years in order to avoid the issues of uncertainty.

The first computer controlled mechanical hand [Ernst 61] was programmed in a language (THI for Teleological Hand Interpreter) which lets the programmer specify how sensor readings (Ernst calls sensors “sense organs”) should change. The interpreter then executes motion commands to make them happen. Ernst distinguishes this *outside-in* mode of control. where the position of a block is indicated by the block itself, from an *inside-out* mode where the the position of a block must be stored inside the machine as a set of coordinates[1].

The robot Shakey [Nilsson 84] at Stanford Research Institute in the late sixties is usually viewed as a *tour de force* in the use of explicit internal world models, and indeed was the basis for a whole generation of planning research. The planner planned in terms of atomic actions, without worrying about how they were implemented in the world. Atomic actions were simple commands that just "worked". But in fact, the robot worked as it did because the environment was carefully engineered so only a few possible classes of situations could be encountered, and the atomic actions were actually implemented as tight sensing-action feedback loops which did not use explicit world models at all. They were called *intermediate level actions* (or ILAs) and enabled the robot to handle such things as the differences in location of objects from the world model, and to handle such uncertain and unsensed things such as the exact way a large block would slide when pushed by the robot.

More recent work (e.g., [Brooks 86, 89], [Horswill and Brooks 88], [Connell 89], and [Ballard 89]) has dispensed with the higher level planner altogether, but has been able to get robots to operate well in the world without having to explicitly worry about uncertainty.

3.1.2 The State of the World*

Can a robot, or creature, be anything but reactive if it does not build world models? We answer this question affirmatively below, but first we need to understand the following.

It turns out that there are more problems with world models than just uncertainty.

Creature researchers who work in simulated worlds (e.g., [Pollack and Ringuette 90] is an example; [Chapman 90] is an exception) are usually not even aware of the problem of uncertainty. But worse, they are not aware of the current technological impossibility of using sensors and perception systems, to build models of the world which reflect an objective reality. Such an objective reality is

---

[1] Admittedly, a rather confusing (to this author) discussion of the nature of analog versus digital control follows, with comparisons to animal control systems. From this discussion it is fair to assume that Ernst, at that time, would not necessarily have agreed with all the arguments being put forth here, in spite of the way he actually made things work.

usually assumed by such researchers (see the collection [Langston 90]). In these simulated 9 systems, objective reality is rather simple. and can usually be described by a short vector of bits. The vector encodes the state of the world, and is an implementation of the notion of a logical description of the state of the world (e.g., [McCarthy and Haves 69]). The real world is not so simple, however.

The way we humans, as observers. describe the world is very much biased by our cultural. morphological, and functional knowledge. Extracting such descriptions of the world from sensors. in terms of things like "chairs", "living rooms" and "crowds" is certainly beyond the state of technology. In very controlled situations a small library of CAD-like models can be matched to grey-level or depth images (e.g., [Grimson 90]) In less controlled situations, some regions of an image can be labeled as belonging to a certain class (e.g., [Thorpe, Hebert, Kanade and Shafer 88]).

The idea that such complete objective models of the world could he built within a creature is somewhat problematic in any case. What does the creature do with such models? It seems that it simply reduces the problem of building a creature to one of building a homunculus that is to live within the creature and operate on complete world models.

What about simpler objective descriptions, such as surface models of what is in the world Computational vision research, as outlined by [Marr 82], is predicated on the assumption that such descriptions of the world can be *recovered* from retinal images. These descriptions are known as 2 1/2-D sketches. Recently, there has been a movement away from the belief that such general purpose representations can be computed. [Ballard 89] argues that such representations are not available to humans.

Without such representations what are we left with? Recent work suggests that we can still have creatures do many things that have traditionally been thought to require representations such as we are rejecting. For instance, despite having no such representations, [Mataric 90] has demonstrated a robot that is much more than reactive, incorporating an ability to modify its behavior based on its experiences within the world. It builds "maps" of the world (topological with some metric annotation, rather than purely metric) as it wanders about indoor environments. But the features in its maps are not objective things in the world. Rather they are the records of certain time averaging detectors having fired while the robot moves about. These detectors are based on very noisy sonar readings, but they have the property that they are both positively and negatively repeatable with respect to any particular place in the world over a wide variance in the robot's actual position and trajectory through that place. Such representations of the world we called *deictic* by [Agre 88] and [Chapman 90].

3.1.3 Analyzing Behaviors

The discussion above can lead us to a rather simpler way of analyzing the behavior of a creature than if we adopted a stance of having objective world models. In particular this method will let us deal with uncertainty entirely at analysis time (analogously to dealing with, say, type checking at compile time), rather than have the creature deal with it at run time.

Given a state of the world and the position of a robot creature within the world, the creature's sensors may deliver different values at different times—i.e., there may be some dynamic noise in the sensors. Worse than that, however, is the system noise, which makes it impossible for the observer of the robot/world system to predict what the sensor readings will be. For instance [Connell 90] found that the flux gate compass on his robot delivered different directions for north in different parts of its domain. The actual values were consistent over time for a given place, but without using that sensor itself, or a similar one, at the exact places where the readings were to be taken, there was no way in which the experimenter or observer would know the actual sensor reading. However it was relatively easy to determine empirically the range of possible errors in the sensor reading. This range of errors, then, is the uncertainty we must deal with for this particular sensor. Similar analyses apply-to other sensors, such as infrared proximity sensors, or sonar sensors.

Now suppose we characterize a behavior at analysis time in terms of an allowable range of sensor readings for the creature (or more generally allowable ranges on some functions computed on the sensor readings).

For instance, instead of characterizing wall following as staying within some distance of a wall, and within some orientation range, we could characterize it as maintaining some range of values on three sonars pointing north-east, east, and south-east, relative to the robot's orientation being north. We might even say that only readings two had to be within their ranges at any particular time.

Or consider approaching a recharging outlet visually. We could characterize that behavior as forcing the visual image of the outlet to appear at every frame taken by a forward looking camera, and for that image size to be non-decreasing between successive images, and to be bigger than some monotonically increasing sequence.

Now our analysis does not have to be concerned with the creature having uncertain sensor readings. Indeed they will be uncertain with respect to the actual world. But what we have to show is that given those sensor readings, the action computed by the creature is guaranteed to produce a next set of sensor readings that lie within our allowable ranges. To do this reasoning at analysis time, and formally prove that a creature will follow a particular behavior, we may need to consider all possible worlds which could have given rise to the initial set of sensor readings, and take into account uncertainty there, and then consider all possible results of the computed action on those worlds and all possible new sensor readings. But the creature itself does not have to maintain any runtime measure of uncertainty, nor be confused or concerned by it.

The essence of what we are suggesting here is to invert the sensor (or more generally a function of the sensors) readings. Instead of thinking in terms of a particular world state and then analyzing all possible sensor readings that could be generated we are analyzing using an inverse method. Namely, given a sensor reading, consider which possible worlds could have given rise to that reading. Recall that this is done at analysis time.

3.2 Micro Relations

There are two obvious fields which are quite related to the micro level of creature architectures: control theory and animal ethology.

Control theory (e.g., [Doyle and Stein 81]), at first glance, seems an attractive tool of analysis for investigating couplings of creatures to the environment. There is a long and rigorous tradition of such analysis of the interaction of mechanical, electrical, and fluid systems with their environments. If a single behavioral mode of the creature can be sufficiently isolated from other aspects of its interaction with the environment, and if the sensors and actuators are well enough understood, then control theory may help to analyze convergence of the behavior. It should be understood, however, that control theory can only be one tool or component of the theories that will be needed for full mastery of engineering artificial creatures.

[McFarland 87] provides a useful introduction to all aspects of animal behavior. Ethology is an approach to studying animal behavior which combines both functional and causal explanations. and in the main concentrates on behavior of animals in their natural environments. To a large extent however, ethology per se does not quite get down to the level of analysis we need at the micro level. It does have some interesting observations about fixed action patterns, essentially feedbackfree (at the gross level in any case) patterns of action that are quite stereotyped. Such patterns may be of use to our artificial creatures in well prescribed environmental niches.

Some literature, closer to neuroscience, does deal with the sensor to action connection we have been discussing at the micro level (e.g., [Wehner 87], [Cruse 90], [Götz and Wenking 73]). This literature can inspire us by the robustness of behavior that is generated by relatively simple systems.

3.3 Micro Practice

The practical problems at the micro level in building complete creatures are still immense.

Primary amongst these is perception-how to map from sensor inputs to aspects of the world. There has been a huge amount of research focused on computer vision over the last twenty years. Unfortunately most of it has been aimed, however indirectly, towards object recognition. Recently there has been a move towards vision that operates on sequences of images, but even there, there are strong traditions of regarding the purpose of vision as something to "recover" properties of objects. Only recently ([Ballard 89], [Chapman 90], and [Horswill and Brooks 88]) has there been a move towards visual processing for a creature that operates continuously in the world, in such a way that appropriate aspects of the world are extracted directly for the particular tasks with which the creature is concerned.

Another approach to useful perception is to make careful choices of sensors, which are well adapted to the task at hand. [Ernst 61] chose sensors that were close to the task-touch and proximity sensors to determine when his mechanical hand should grasp. Such sensor choice is also used extensively in the biological world, as described by [Wenner 87]—in fact the presentation there has great resonance with the thrust of this complete section.

At this point in time, for actually specifying a single behavior for a creature, our programming tools are rather primitive. We are forced to individually craft programs to generate particular behaviors. It would be good to have almost any form of synthesis tools which could take a goal specification (note that [Kaelbling 88] is trying to do this, but at the macro level rather than the micro level), and produce some code or rules to generate the appropriate behavior. In some sense, this is what [Lozano-Pérez, Mason and Taylor 84] are trying to do, but only in very well characterized situations.

4 Macro Level

Once we have imbued our creatures with multiple behaviors to handle a variety of circumstances and achieve a variety of tasks, we are faced with the problem of deciding which behavior or behaviors should be active at any particular time. There are two components to this decision, which behaviors are potentially correct in the circumstances, and how to resolve conflicts between behaviors. In more detail we need to consider the following issues:

Coherence: Even though many behaviors may be active at once, or are being actively switched on or off, the creature should still appear to an observer to have coherence of action and goals. It should not be rapidly, switching between inconsistent behaviors, nor should two behaviors be active simultaneously, if they interfere with each other to the point that neither operates successfully.

Salience: The behaviors that are active should be salient to the situation the creature finds itself init should recharge itself when the batteries are low, not when they are full.

Adequacy: The behavior selection mechanism must operate in such a way that the long term goals that the creature designer has for the creature are met. i.e., a floor cleaning robot should successfully clean the floor in normal circumstances, besides doing all the ancillary tasks that are necessary for it to be successful at that.

4.1 Macro Theory

There are at least two fundamental questions of organization for a behavior selection mechanism at the implementation or algorithmic level.

- Should the behavior selection mechanism be centralized or decentralized?

- Should the conflict resolution scheme be fixed priority or dynamically reconfigurable?

At this point in time we have no theoretical way of linking these questions to the issues of coherence, salience and adequacy raised earlier. Later, we will point to work which has been done in these areas however in terms of practical implementations.

One point of theoretical importance, however, concerns hierarchical planning. Centralized behavior selection mechanisms have a long tradition in the form of AI planning as strongly introduced in the late sixties [Nilsson 84]. Recently classical planning has been shown to be intractable even in quite limited circumstances ([Chapman 87]), and has fallen from favor amongst creature builders.

4.2 Macro Relations

Ethology (again, see [McFarland 87]) has struggled with the problems of behavior selection for many years. Theories of drive have come and gone to explain motivation for particular behaviors. but there are still no definitive answers, as to exactly what leads to a particular behavior being activated when many are at least not inappropriate in certain circumstances.

Ethologists have also observed and described displacement activities—inappropriate behaviors that operate when two conflicting but appropriate other behaviors could reasonably be triggered. It seems that real creatures are by no means bug free.

From the neuroscience side there has been some recent progress in explaining animal behavior selection, at least in simpler animals. [Kravitz 88] describes a model of hormonal control of behavior. Following Kravitz' own summary the major points of this model are:

1. Sensor inputs enhance the release of hormones.

2. A hormone may be released from tissues or neurons and may act in isolation or in concert with other substances.

3. The hormone finds receptors where it either stimulates a new behavior pattern or enhances or diminishes an existing pattern.

4. The systems affected by hormones may include sensory sory elements, groups of higher processing centers, and motor or hormonal output systems

5. The method of stimulation is a gain setting mechanism, biasing the output of the organism in particular directions.

6. The change is apparent in that the organism then responds to particular sensory inputs with altered outputs.

7. Individual organisms may respond to particular hormones with different levels of effectiveness determined by genetic or environmental influences.

8. The affected circuit may itself further enhance or diminish the release of the hormonal substance.

Kravitz uses the lobster as his primary instantiation of this model.

In [Brooks and Viola 90] we demonstrated a robotic instantiation of Kravitz' model (programmed in the "Behavior Language" [Brooks 90]) which successfully demonstrated seven of the listed eight properties (number seven being the only one not accounted for).

4.3 Macro Practice

If we restrict ourselves to implementations on actual robots there has been very little actual work in behavior selection.

[Brooks 86] used a distributed system, called the subsumption architecture, with an almost fixed priority scheme to arbitrate between behaviors. There was a small amount of internal communication between the behaviors, but it was not directly related to priorities in conflict resolution or behavior activation. Indeed the behaviors were all active at all times. [Connell 89] later purified this scheme so that there was no internal communication between behaviors. A major criticism of the scheme used is that there is no explicit mention of behavior selection or priorities in the programming language for expressing behaviors.

[Kaelbling 88] remedies that shortfall by having an explicit programming language for behavior selection. The code generated by her scheme was again decentralized and used fixed priorities determined at compile time.

[Maes 89] suggested another decentralized scheme but with more explicit references to behavior activation. She develops the scheme in [Maes 90] and [Brooks 90] produced a superset language. the Behavior Language, which could be used to implement the scheme on mobile robots. This language was used by [Mataric 90] in a system where some individual behaviors represented landmarks in the world, and by [Maes and Brooks 90] where a learning mechanism was introduced to learn arbitration between behaviors (enabling a six legged robot to learn to walk from negative feedback when it fell down). An interesting thing about the Behavior Language is that it actually compiles into subsumption architecture programs with a fixed priority arbitration scheme. The user of the Behavior Language does not see this restriction however, as it is hidden in an abstraction mechanism.

The Behavior Language also contains features inspired by the work of [Kravitz 88] with the hormone system of the lobster. [Brooks and Viola 90] show how these features enable almost a full implementation of the Kravitz model.

Two problems stand out with these schemes:

- We do not know how well they will scale. We certainly do not have anything like a Turing equivalence theorem that tells us that at least in principle these schemes can do anything we might ever want them to do.

- We do not have any analytic tools for understanding in advance what sort of conflicts and other unexpected interactions might arise from the ways we paste together behaviors using these methodologies.

These two problems are key research areas, and progress in either would fundamentally help the field.

5 Multitude Level

Once we have physical creatures successfully operating in the world, handling its dynamics and surprises, there naturally arises the question of how such creatures will interact with each other.

This question can be posed in terms of how just two creatures will interact, or inspired by large colonies of social insects such as ants and bees, we can ask whether and how interesting behavior of that class might arise when we have large collections of artificial creatures in contact with each other. It is this latter question that we will concentrate upon for the rest of this paper.

Some of the major issues for social interactions of large numbers of artificial creatures me:

Emergence: Given a set of behaviors programmed into a set of creatures we would like to be able to predict what the global behavior of the system will be, and as a consequence determine the differential effects of small changes to the individual creatures on the global behavior.

Synthesis: As at the micro level, given a particular task, automatically derive a program for the set of creatures so that they carry out the task.

Individuality: Robustness can be achieved if all creatures are interchangeable. A fixed number of classes of creatures, where all creatures within a class are identical, is also robust, but somewhat less so. The issue then is given a task, decide how many classes of creatures are necessary

Cooperation: In some circumstances creatures should be able to achieve more by cooperating—the form and specification of such possible cooperations need to be understood.

Interference: Creatures may interfere with one another. Protocols for avoiding this when it is undesirable must be included in the design of the creatures' instructions.

Density dependence: The global behavior of the system may be dependent on the density of the creatures and the resources they consume within the world. A characterization of this dependence is desirable. At the two ends of the spectrum it may be the case that (a) a single creature given *n* units of time performs identically to *n* robots each given I1unit of time, and (2) the global task might not be achieved at all if there are fewer than, say, *m* creatures.

Communication: Performance may be increased by increasing the amount of explicit communication between creatures, but the relationship between the amount of communication increase and performance increase needs to be understood.

This is an area where almost no work has been done with actual physical robots. so the following remarks are rather brief Some work has been done with a small number of robots (e.g., [Premvuti and Yuta 90]) but none with large numbers of robots.

5.1 Multitude Theory

In many ways the aim of controlling multiple creatures to generate a global behavior is similar to the problem of controlling a single creature to generate a single behavior. The difference is in the large amount of communication that goes on through the world. in the case of multiple creatures, where their very presence. besides any changes they may make to the world itself, are signals for changes in behavior of other creature.

For suitably abstracted creatures it may be possible to describe their group behavior as differential equations and again use classical control theory to understand the conditions under which they will produce a certain global behavior. This sort of analysis will become more difficult as the creatures modify the world and may turn out only to be useful for understanding grouping and flocking behaviors.

5.2 Multitude Relations

There are a number of disciplines which may shed light on organizations of many creatures.

From a descriptive and mechanistic viewpoint their is much interesting work on the organization of social insects [Deneubourg et al 87] is just one example of how complex patterns of behavior can emerge from simple systems of biological creatures.

The theories on dynamical systems. and self organization and the construction of dissipative structures ([Nicolis and Prigogine 77]) may have some relevance to the field also. See [Steels 89] for one approach to investigating the behavior of many simple creatures acting in a single environment.

None of this work at this point gives us any handle on how to synthesize creatures which will carry out global tasks that we desire.

5.3 Multitude Practice

There is no real work yet reported on multiple real (physical) robot creatures interacting in non-trivial ways. There is a lot of work on simulations (e.g., see [Langston 87] and [Langston 90]). It remains to be seen how relevant all this work is to actual physical systems interacting in the world.

6 Conclusion

Building complete creatures is a complex endeavor. No single theory will be adequate to guide us in either synthesis or analysis. We can expect many partial theories to be developed over time, some quite limited in their scope. Likewise, we can expect many programming methods to be developed, again limited in scope. By combining the pieces we will be able to build practical systems. In the meantime, more global theories will emerge, as we have more experimental systems on which we can gain experience.

7 Acknowledgements

Anita Flynn and Maja Mataric provided many valuable comments on earlier drafts of this paper.

Support for this research was provided in part by the University Research Initiative under Office of Naval Research contract N00014-86-K-0685, in part by the Advanced Research Projects Agency under Office of Naval Research contract N00014-85-K-0124.

8 References

[Agre 88] “The Dynamic Structure of Everyday Life”, Philip E. Agre, *MIT A.I. Lab Technical Report 1085*, Dec.

[Angle 89] “Genghis, a Six Legged Autonomous Walking Robot”, Colin M. Angle, *MIT S.B. Thesis in Electrical Engineering and Computer Science*, March.

[Angle and Brooks 90] “Small Planetary Rovers”, Colin M. Angle and Rodney A. Brooks, *IEEE/RSJ International Workshop on Intelligent Robots and Systems IROS ‘90*, July, Ikabara, Japan, 383-388.

[Ballard 89] “Reference Frames for Active Vision”, Dana H. Ballard. *IJCAI-89*, Detroit, MI, 1635-1641.

[Brooks 82] “Symbolic Error Analysis and Robot Planning”, Rodney A. Brooks, *The International Journal of Robotics Research* 1:4, 29-68.

[Brooks 86] “A Robust Layered Control System for a Mobile Robot”, Rodney A. Brooks, *IEEE Journal of Robotics and Automation, RA-2*, April, 14-23.

[Brooks 89] “A Robot that Walks: Emergent Behavior from a Carefully Evolved Network”, Rodney A. Brooks, *Neural Computation*, 1:2, Summer, 253-262.

[Brooks 90] “The Behavior Language; User’s Guide”, Rodney A. Brooks, *MIT A.I. Lab Memo 1227*, April.

[Brooks and Viola 90] “Network Based Autonomous Robot Motor Control: from Hormones to Learning”, Rodney A. Brooks and Paul A. Viola, *Advanced Neural Computers*, Rolf Eckmiller (ed), Elsevier, 341-348.

[Chapman 87] “Planning for Conjunctive Goals”, David Chapman, *Artificial Intelligence*, 32, 333-377.

[Chapman 90] “Vision, Instruction, and Action”, David Chapman, *MIT A.I. Lab Technical Report 1204*, Aug.

(Chatila and Laumond 85] “Position Referencing and Consistent World Modeling for Mobile Robots”, R. Chatila and J. Laumond, *IEEE International Conference on Robotics and Automation*, March, 138-145.

[Connell 87] “Creature Building with the Subsumption Architecture”, Jonathan H. Connell, *IJCAI-87*, Milan, August, 1124-1126.

[Connell 90] “A Colony Architecture for an Artificial Creature”, Jonathan H. Connell, *MIT A.I. Lab Technical Report 1151*, June.

[Cruse 90] “What Mechanisms Coordinate Leg Movement in Walking Arthropods?”, Holk Cruse, *Trends in Neurosciences*, 13:1, 15-21.

[Deneubourg et al 87] “Self-Organization Mechanisms in Ant Societies, II: Learning in Foraging and Division of Labor”, J.-L. Deneubourg, S. Goss, J. M. Pasteels, D. Fresneau, and J.-P. Lachaud, *For Individual to Collective Behavior in Social Insects (Experientia Supplement, vol 54)*, 177-196.

(Doyle and Stein 81] “Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis”, J. C. Doyle and G. Stein, *IEEE Transactions on Automatic Control*, AC-26(l), 4-16.

[Ernst 61] “MH-1, A Computer-Operated Mechanical Hand”, Heinrich A. Ernst, *MIT Ph.D. Thesis*, Dec.

[Flynn 87] “Gnat Robots (And How They Will Change Robotics)”, Anita M. Flynn, IEEE Micro Robots and Teleoperators Workshop, Hyannis. MA, Nov.

[Flynn and Brooks 89] “Building Robots: Expectations and Experiences”, Anita M. Flynn and Rodney A. Brooks, *IEEE/RSJ International Workshop on Intelligent Robots and Systems IROS '89*, Sept, Tsukuba, Japan, 236-243.

[Giralt, Chatila and Vaisset 84] “An Integrated Navigation and Motion Control System for Multisensory Robots”, Georges Giralt, Raja Chatila and Marc Vaisset, *Robotics Research 1*, Brady and Paul (eds), MIT Press, 191-214.

[Götz and Wenking 73] “Visual Control of Locomotion in the Walking Fruitfly Drosophilia”, Karl Georg Götz and Hans Wenking, *J. Comp. Physiol.* 85, 235-266.

[Grimson 90] “Object Recognition by Computer: The Role of Geometric Constraints”, W. Eric L. Grimson, MIT Press, Sept.

[Horswill and Brooks 88] “Situated Vision in a Dynamic World: Chasing Objects”, Ian D. Horswill and Rodney A. Brooks, *AAAI-88*, St Paul, MN, August, 796-800.

[Kaelbling 88] “Goals as Parallel Program Specifications”, Leslie Pack Kaelbling. *AAAI-88*, St Paul, MN. August, 60-65.

[Kravitz 88] “Hormonal Control of Behavior: Amines and the Biasing of Behavioral Output in Lobsters”, Edward A. Kravitz, *Science*, Vol 241. September 30, 1775-1781

[Langston 87] “Proceedings of Artificial Life”, Christopher Langston (ed), Addison-Wesley, appeared 1989.

[Langston 90] “Proceedings of Artificial Life, II”, Christopher Langston (ed), Addison-Wesley, to appear.

[Lozano-Pérez 76] “The Design of a Mechanical Assembly System”, Tomás Lozano-Pérez, *MIT A.I. Lab Technical Report 397*, Dec.

[Lozano-Pérez Mason and Taylor 84] "Automatic Synthesis of Fine-Motion Strategies for Robots", Tomás Lozano-Pérez, Matthew T. Mason and Russell H. Taylor, *Robotics Research 1*, Brady and Paul (eds), MIT Press, 65-96.

[Maes 89] "The Dynamics of Action Selection", Pattie Maes, *IJCAI-89*, Detroit, MI. 991-997.

[Maes 90] "Situated Agents Can Have Goals", Pattie Maes, *Robotics and Autonomous Systems*, 6.

[Maes and Brooks 90] "Learning to Coordinate Behaviors", Pattie Maes and Rodney A. Brooks, *AAAI-90*, Boston, MA, August, 796-802.

[Marr 82] "Vision", David Marr, W. H. Freeman and Co., San Francisco.

[Mataric 90] "A Distributed Model for Mobile Robot Environment—Learning and Navigation", Maja J. Mataric, *MIT A.I. Lab Technical Report 1228*, May.

[McCarthy and Hayes 69] "Some Philosophical Problems from the Standpoint of Artificial Intelligence", John McCarthy and Patrick Hayes, *Machine Intelligence 4*, Edinburgh University Press.

[McFarland 87] "The Oxford Companion to Animal Behavior", David McFarland (ed), Oxford University Press, Oxford.

[Moravec 80] "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover", Hans P. Moravec, *Stanford A.I. Memo 340*, Sept.

[Nicolis and Prigogine 77] "Self-Organization in Nonequilibrium Systems", G. Nicolis and I. Prigogine, John Wiley and Sons, New York.

[Nilsson 84] "Shakey the Robot", Nils J. Nilsson (ed). *SRI A.I. Center Technical Note 323*, April.

[Pollack and Ringuette 90] "Introducing the Tileworld: Experimentally Evaluating Agent Architectures", Martha E. Pollack and Marc Ringuette, *AAAI-90*, Boston, MA. August, 183-189.

[Pool 90] "Understanding the Simplest Reaction", Robert Pool, *Science*, Vol 247, January 26, 413.

[Premvuti and Yuta 90] "Consideration on the Cooperation of Multiple Autonomous Mobile Robots", Suparerk Premvuti, and Shin'ichi Yuta, *IEEE/RSJ International Workshop on Intelligent Robots and Systems IROS '90*, July, Ikabara, Japan, 59-63.

[Roberts 63] "Machine Perception of Three-Dimensional Solids", Larry G. Roberts, *MIT Lincoln Laboratory Technical Report No. 315*, May.

[Steels 89] "Cooperation Between Distributed Agents Through Self-Organization", Luc Steels, *Proceedings of the Workshop on Multi-Agent Cooperation*, Cambridge, UK, August, North Holland.

[Taylor 76] "The Synthesis of Manipulator Control Programs from Task-Level Specifications", Russell H. Taylor, *Stanford A.I. Memo 282*, July.

[Thorpe, Hebert. Kanade and Shafer 88] "Vision and Navigation for the Carnegie Mellon Navlab", Charles Thorpe, Martial Hebert, Takeo Kanade and Steven A. Shafer, *IEEE Trans. PAMI*, 10(3), May, 362-371.

[Wehner 87] "Matched Filters - Neural Models of the External World", Rüdiger Wehner, *J. Comp. Physiol. A.*, 161, 511-531.