Todd, P.M., and Wilson, S.W. (1993). Environment structure and adaptive behavior from the ground up. In J.-A. Meyer, H.L. Roitblat, and S.W. Wilson (Eds.), *From animals to animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 11–20). Cambridge, MA: MIT Press/Bradford Books.

Environment Structure and Adaptive Behavior From the Ground Up

Peter M. Todd **Stewart W. Wilson**
The Rowland Institute for Science
100 Edwin H. Land Boulevard
Cambridge, MA 02142 USA
ptodd@spo.rowland.org wilson@smith.rowland.org

Abstract

We describe a framework for exploring the evolution of adaptive behaviors in response to different physical environment structures. We focus here on the evolving behavior-generating mechanisms of individual creatures, and briefly mention some approaches to characterizing different environments in which various behaviors may prove adaptive. The environments are described initially as simple two-dimensional grids containing food arranged in some layout. The creatures in these worlds can have evolved sensors, internal states, and actions and action-triggering conditions. By allowing all three of these components to evolve, rather than prespecifying any of them, we can explore a wide range of behavior types, including "blind" and memoryless behaviors. Our system is simple and well-defined enough to allow complete specification of the range of possible action-types (including moving, eating, and reproducing) and their effects on the energy levels of the creature and the environment (the bioenergetics of the world). Useful and meaningful ways of characterizing the structures of environments in which different behaviors will emerge remain to be developed.

1 Introduction

We are interested in the effects that the environment can have on an organism's adaptive behavior. Since this question encompasses essentially the whole of psychology and biology, we must pare it down a bit before we can make much headway. We focus here on how the physical, spatial structure of the environment can foster the evolution of, and be in turn exploited by, particular adaptive behavior-generating mechanisms. We leave aside for now the fascinating questions that arise when the environment is considered not only in terms of the physical selectors at work in it, but also the biotic (e.g. parasites and hosts) and psychological (e.g. conspecifics and mimics) selectors (see Miller, 1992). Our work here may in fact be seen as complementary to explorations elsewhere attempting to characterize the effects of the social environment on adaptive behavior and evolution (Todd & Miller, 1991; Todd & Miller, submitted). As it turns out, even without this added complexity and realism, the sorts of behaviors that can evolve to take advantage of static spatial environments are still varied and interesting. This work is also intended to provide theoretical support for the behavior-based approach to AI described by Maes (these proceedings). By categorizing environments and describing the sorts of behaviors that are adaptive in them, we hope to provide not only insights into natural evolved systems, but also useful guidelines for the design of artificial agents existing in various application domains.

Creating simple simulated worlds and seeing what sorts of simulated behaving creatures will evolve in them is certainly not a new approach. Ackley and Littman (1992) and Werner and Dyer (these proceedings) have developed sophisticated simulations that embody many of the ideas described here, and serve as positive examples of what is possible with such an approach. But our goals here are different from those embodied in many similar research efforts. Rather than manipulating the environment with the specific aim of evolving creatures that can navigate, or communicate, or learn, we want to explore a more general question: what *will* creatures evolve to do, given certain environmental regularities or structures? What behaviors will prove adaptive in various types of environments? These questions require us to do two very intertwined things in our research program: both elucidate the sorts of environment-exploiting behavioral mechanisms that creatures might employ, and describe and characterize how environments can vary in ways that lead to the evolution of different forms of adaptive behavior. Obviously such goals are very ambitious, and perhaps still beyond our reach, but at least by beginning to formalize our thinking about these questions, we can hope to make some progress in our understanding of the complexities involved in the behavioral interactions between environments and organisms.

In a simulated 2-dimensional world across which food is distributed in some fashion, creatures attempting to find and eat that food may get along fine with no sensory systems, or without memories or internal states, or with very few motor commands; or the creatures may find it virtually essential to possess long-distance sensors, sophisticated internal world models, and finely-tuned motor sequences before they can achieve any adaptive advantage at all. What sensors, states, and actions prove adaptive depends on the environment in which the population of creatures evolves. But since it is exactly those three components (at least) which define an organism's adaptive behavior, in order to study the effects of the environment on adaptive behavior we must instantiate our study in a framework which allows the evolution of all three components, something not usually attempted in evolutionary simulations. Cariani (1990) and Pattee (1989) in particular have issued the call to consider the evolution of sensors and effectors in addition to the behavioral links in between, since it is only through the former two that creatures can ground themselves in connection to the outside world. We hope that the framework we present here will allow the exploration of the evolution of exactly those sensors and effectors, the process that Cariani calls *semantic-adaptation*, along with the *syntactic-adaptation* of evolving information-processing mechanisms, thereby yielding a more complete picture of the evolution of behavior in general.

To try to achieve these grand ends, we focus primarily in this paper on an open-ended scheme for the evolution of simulated creatures in terms of their sensors, internal states, and actions. Only after this framework has been laid down can we turn to the question of the structure of environments, and the effects they can have on these evolving creatures and their behaviors. In this combined context, we discuss briefly our planned explorations of the interactions of these two parts of our system, and what we can hope to learn from such an investigation.

2 The World and the Organism

The creature and its environment cannot be described separately; each is shaped by and shapes the other. Different types of environments call for different behavioral mechanisms to respond to them adaptively; and different behavioral mechanisms, including internal representations and sensory inputs, change the very structure of the world-lived-in for the creature living in it, or, via emitted actions, the structure of the world experienced by others. To understand the interactions of the physical world and behavior, then, we must consider the two in an intertwined fashion. In this section we first describe possible worlds and adaptive behaviors in them with an eye toward showing what components we'll need to create those behaviors, and then we discuss those components themselves, and how they can work together to create adaptive behavior in the world.

2.1 The Structure of the World

To begin with the simplest interesting (and easily visualizable) case, we use a 2-dimensional hexagonal-grid world consisting of N by M positions. (A hexagonal rather than a square grid is used to avoid the anisotropies in distances between orthogonally and diagonally adjacent squares in the latter.) Each position $W(x, y)$ in the world can contain food, one or more creatures, both, or nothing. Time in the world passes by in discrete clock-ticks, t. The currency in this world is simply energy; creatures live on it and use it up, and food contains it — the greater the amount of food in a particular location, the greater the amount of energy available to a grazing organism. (This is in contrast to Holland's ECHO system, which works on a flow of metabolic products equivalent to organic chemical compounds — see Holland, 1992.)

The exact layout of food in the world over time constitutes the spatio-temporal structure of the environment, in response to which the creature should behave adaptively. This layout is created primarily by the real-valued function, $\phi(x, y, t)$, which indicates the new food being added to each position in the world at each time-step (so that it can be thought of as specifying the "plant growth" in the world). Since creatures may, through their actions, cause the amount of food in the world to change, the contents of W at time t can be computed as the contents of W after time $t-1$ (including the effects of creature actions) plus $\phi(x, y, t)$, the new food appearance/growth that occurs.

2.2 The Structure of the Organisms

The world can contain one or more simulated organisms running around in it at a time. Each organism in the world has associated with it a 3-tuple (x, y, E) which codes its current position and internal energy level, and a behavior-generating component which defines its possible actions and triggering conditions, internal state variables, and sensors. The position-energy 3-tuple can change during a creature's lifetime; the beginning values for each creature are determined at the time of their creation, as will be described later. The behavioral portion of each creature is created by an evolutionary process, and is fixed throughout the creature's lifetime (thus we are for the moment ignoring learning processes). To introduce the subcomponents of the creature's behavioral mechanisms, we will first look at a few example environmental situations.

Consider four possible worlds, and the types of creatures that might do well in them. First, there could be food everywhere in the world, evenly and equally distributed (see Figure 1 — while the world is shown with squares in this figure, it is represented as hexes in the simulation). In this case a creature really does not need very sophisticated behavior to maximize its food input: it can basically just move about the environment in a straight line, perhaps turning every now and again, eating as it goes. Thus, if the creature merely has a small set of motor commands, such as "eat," "go straight," and "turn right," which it selects from stochastically in a fixed proportion, say 1:1:0.1, then this creature will do about as well as possible (in terms of gaining food

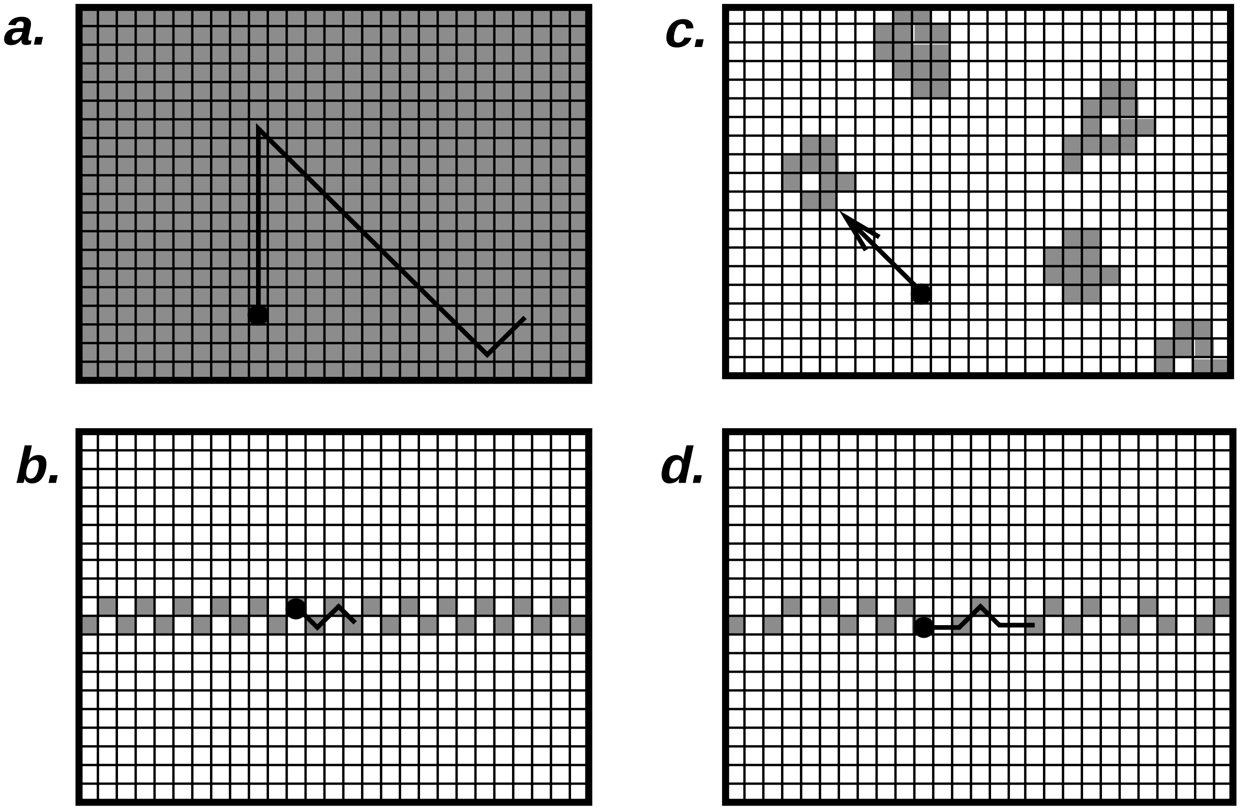

Figure 1. Environments with differing behavioral component requirements. a. An environment with food everywhere (dark grey), showing a creature (black circle) and the mostly-straight path it might take. b. A zig-zag environment, showing a creature and the regular path it might take. c. A clumpy environment, showing a creature heading toward one of the clumps. d. A zig-zag-with-gaps environment, showing a creature adjusting its regular path to take the missing food-squares into account.

energy) in this world. It has no need for any sensory input, nor for any memory of what it's already done or seen, that is, no internal state is necessary (or helpful). Thus in this case, evolution need only supply the creature with the appropriate motor commands, and appropriate rates of firing them stochastically.

If the world is laid out slightly differently, for instance with food in a regular "zig-zag" pattern stretching across the whole 2-d plain (see Figure 1b), then a slightly more sophisticated set of behavioral mechanisms could prove useful. A creature which performs the action sequence "eat, move forward, move left, eat, move forward, move right, eat..." over and over again (provided it starts off at an appropriate position!) will find as much food as possible. Here there is still no need for any sensory input, since the world is very regular and so evolved models of the world can mirror it exactly; but internal state *is* very important in this case. Essentially the creature must utilize an evolved "central pattern generator" (CPG) that produces this string of motor commands over and over. Since this CPG must keep track of where it is in the action-cycle, it must use some state information. In this situation, then, both appropriate actions triggered by certain internal states, and the means of properly generating those internal states, must evolve.

Next, consider a world in which food is distributed sparsely and randomly, in small clumps, as shown in Figure 1c. In this case, a creature would do well to have some sensory input. If it just moves in a fixed or random manner, as in the first situation, or in a blind preprogrammed fashion, as in the second, it is quite unlikely to run across much of the widely separated food in its lifetime. But if it has a sensory apparatus which indicates the direction of the nearest source of food, the creature can base it movements on this input and head in that direction, eating when it arrives at the food. Here, evolved actions can be based solely on the outputs of evolved sensors; there is no need for internal state to guide the actions.

Finally, imagine a modification of the "zig-zag" world of Figure 1b, in which there are occasional gaps in the zags, as shown in Figure 1d. In this case, if we drop one of our previously-evolved zig-zag creatures into this world, it will do fine for a while, but when it comes to one of the gaps, it will be thrown off by one position, and will miss all the food until the *next* gap comes. Obviously, it should be able to do better than this. What it needs so it can do better is again some sensory input, which this time modifies the action of the central pattern generator by adding an extra "move forward" whenever it detects a gap in the food pattern. So in this scenario, all three behavioral components will prove adaptive and could be expected to evolve: a set of actions including "eat," "move forward," "move left," and "move right"; a set of internal state variables which successively influence each other to trigger the proper actions in sequence; and a set of sensors which can detect gaps in the zig-zag sequence in the world, and modify the action of the internal state CPG loop to keep the sequence of generated behaviors in sync with the structure of the world. (This is similar to the sensor-modulated CPGs that Gallagher and Beer, these proceedings, have evolved to control artificial insect locomotion.)

We can see from these four examples that actions, internal states, and sensors can all combine independently in a variety of ways to create adaptive behavior in different environments. Thus to explore the evolution of such adaptive behaviors, we need a system which can evolve each of these components independently. Moreover, to be able to analyze the simulated creatures which result from our evolutionary scenarios in terms of these behavioral components, it is important that they be kept modularly distinct; if we were just to evolve big unstructured neural networks or Lisp code representations of our creatures, for instance, it would be very difficult to say when a network was using internal state, or whether or not a Lisp routine instantiates a central pattern generator. (It is typically easier to determine what actions and sensory inputs such systems are employing than what they're doing internally, but even that may not be so clear in more complex situations.) To help alleviate this difficulty of both evolution and analysis, we have developed a modular system of evolving lists of actions and triggering conditions, internal state variables, and sensory input systems. These lists of variables are restricted in how they can interact with each other, so that modularity and interpretability can be maintained. Evolution creates and builds up the entries in each of these lists within each individual creature. For now, we will remain agnostic in our choice of representation and implementation of the elements in these lists; they should merely be thought of as some sort of simple information-processing mechanisms. (We will return to the issue of implementation in a later section.) We now consider each of these three behavioral components in detail.

2.2.1 Sensors

The ultimate role of sensory systems is to tell creatures about what's out there in the world that might be important. Sensors are not for informing about spectral intensity, or airborne polypeptides, or modulated pitch formants; rather, they are for indicating the presence of a colorful ripe fruit, or a sexually-ready conspecific, or a snarling predator. That is, sensors should be thought of as signalling *fitness affordances*, things in the world which can have a positive or negative effect on the individual's ongoing fitness (see Miller & Todd, in preparation, for a more detailed account of this topic, patterned after Gibson's 1966 notion of perceptual affordances). As such, we will define the sensors in the present system as providing information about just those things that are out there in the world, that is, about the contents of $W(x, y)$.

Each individual has a list of sensors, $S[i]$, where i can go from 1 to some preset maximum number of sensors n_S. If n_S is preset to 0, then none of the creatures in this world can have any sensors — this is how we can restrict our experiments to the first two types of environmental situations presented earlier. Each $S[i]$ is defined as a function $f_{S[i]}$ of just the category-based contents of the world, $W_c(x, y)$, for a particular category c. That is, $S[i] \equiv f_{S[i]}(W_c(x, y))$, where c in the current simple worlds can be either food or (other) creatures, so that two general classes of sensors are possible. The categories c that sensors can respond to are just those classes of objects or entities that the creatures can act on — the granularity of actions and sensors is always equal. (There is no proprioception possible through the sensory inputs in S, but this can sneak in through the internal state variables described next.)

Each $S[i]$ can be as complicated or as simple as evolution deigns to make it. For instance, one creature might have $S[18] = \log(W(1, 1)) + 8.73$, a complex function of an absolute position in the world, and another might have $S[7] = W(x_p + 1, y_p) + W(x_p - 1, y_p)$, where x_p and y_p indicate the individual's current position, in which case this sensor looks at what's above and below the creature in the 2-d world. Sensors can also be noisy or inaccurate, delivering an only more-or-less realistic view of the objects in the world. If a particular ith sensor is not present (which can be signified by an "inactive" flag for that sensor in the evolving list of sensors), $S[i]$ is defined to be 0.

Such "direct-perception" sensors may not seem to match the sorts of indirect windows on the world our own sensors sometimes feel like, and one might object that there's nothing here like propagating chemical gradients or reflecting light or sound waves to carry information around in the environment. But again, it is the ultimate sources of these signals that we (and all living creatures) are really interested in, rather than the particular form of their proximal cues, and the sensory scheme presented here captures just this notion. In addition, by focussing on the objects in the world and their action consequences, we hope to make the behavioral orientation of this work clearer, and at the same time avoid the complicated (and important) problems of sensory transduction.

2.2.2 Internal States

The internal state variables available to an individual creature are designated in a manner similar to the sensors. They are indexed in a list $I[i]$, for i from 1 to n_I, the preset maximum number of internal state variables. Again, if n_I is preset to 0, all creatures will evolve without any internal state. Each $I[i]$ is defined as a function of (some of) the current sensory inputs, the previous internal states, and the previous action-states (indicating whether or not a particular action fired on the previous time-step); that is, $I[i] \equiv f_{I[i]}(S^t, I^{t-1}, A^{t-1})$. The inclusion of the previous action as one of the possible inputs to the internal state variables allows for limited proprioception, as mentioned in the previous section. Each $I[i]$ is set to 0 if not defined (inactive), just as for each $S[i]$.

Internal state variables allow further processing on the sensory inputs. For example, $I[3] = \textit{logistic}(S[7])$ would perform a gentle thresholding of the results of sensor 7 (looking up and down). They can also provide a memory for past events or actions, as in the pair $I[52] = A^{t-1}[3]$, which remembers whether or not action 3 was just performed, and $I[17] = I^{t-1}[52]$, which returns the performance-status of action 3 two time steps ago. More complicated systems, like the central pattern generators mentioned earlier, can also be evolved (though it may take quite a while for such things to be hit upon in such an open-ended framework). The outputs of the internal state variables, along with the current sensory values, are then made available to the final component of the behavioral triptych, the actions and their triggers.

2.2.3 Actions and their Triggers

Each action $A[i]$ an individual can perform is defined as a two-part structure consisting of an action-type and a triggering function. (Here again i can vary from 1 to n_A, the maximum number of different actions an evolved behavioral system can contain; in this case, n_A will never be set to 0, or else we'd end up with an uninteresting, unbehaving creature.) Action-types $A[i]_{\textit{act}}$ can be primitive motor commands such as "move forward," "turn," "eat," etc.; we will discuss these in more detail in the next section. These actions can be thought of as very simple "fixed action patterns" which are fired off and run to completion without any environmental influence either before or during their execution. Thus they are not functions of either the sensory inputs or the internal state variables, but rather are impenetrable units.

The way sensory inputs and internal states *are* able to affect emitted behavior is by controlling which behaviors actually get emitted at all, through the triggering functions. Triggering functions $A[i]_{\textit{trig}}$ are real-valued functions of the individual's current sensory inputs and current internal states (if there are any of each) which are used to determine when a particular action is invoked. That is, $A[i]_{\textit{trig}} \equiv f_{A[i]}(S^t, I^t)$. To determine which action an individual will perform at a given time-step, all of the creature's (active) trigger functions are evaluated, giving a real value representing the strength with which each possible (active) action is being triggered. From here, there are several ways that the final single action to be performed can be chosen, including selecting the action $A[i]$ with the highest trigger value, choosing an action stochastically based on the normalized trigger values interpreted as a discrete probability distribution function, or picking an action at random from among all those above some trigger threshold. Different selection schemes will result in different observed patterns of behavior, and therefore also in different evolved behavioral systems to achieve the same behavioral ends.

2.3 A Catalogue of Action-types

The things that an organism can do in its world depend on what there is in that world to act on and affect. In the system we've described here there is not all that much in the world for the creatures to alter: basically, there is only food/energy, and the creatures themselves. As a result, the number of action-type primitives that we can define for this world will also be limited. We view this as a plus, though, since it will keep our analysis and understanding of the evolutionary and behavioral dynamics of this system that much simpler.

First of all, creatures can change the distribution of food in the world, that is, change the food entries in $W(x, y)$. The most obvious action in this category is "eating" — lowering the food-value at a particular world location. (This will in turn result in an increase the creature's own energy level, E, and a decrease in the energy of the food, but we will consider these as energetic "side-effects" of the direct action, described in the next section.) Though this behavior seems simple, a variety of possible action-types are subsumed under this category. For instance, what (x, y) locations do we allow a creature to eat from? Only their own current location? Or neighboring locations? How big a "reach" should we allow creatures in that case? And how much of the food will the creatures eat? All of it at once? Or can they opt to eat a lesser amount? If there's a lot of food to be eaten at a certain location, should it take longer for the creature to do so? We plan to extend this system so that the eating command (and others) will itself be a 3-tuple, (x, y, E), specifying what location to eat from, and how much to eat (i.e., the amount of energy to absorb). This action will be generated as a function of the current internal state and sensory inputs, just as the triggering values now are. How difficult each such action is (i.e., how much effort, in terms of energy, it takes the creature to perform it), and what the energetic benefits and side-effects are, would again be specified in the bioenergetics of the world. But in the meantime, we can restrict this class of actions simply to "eat everything that's in your current hex, in one time-step."

Besides just eating the food in the world, though, there are other ways creatures can alter its distribution. For example, we might want to allow creatures to *increase* the amount of food at a given location, by "fertilizing" that position (certainly at a cost in energy terms to the creature itself, another side-effect). Or creatures could "farm" the food in the world, by "pushing" the food spread in some large area into a smaller region for easier later harvest or hoarding from other creatures. (This could also be achieved through a combination of "eating," "moving," and "fertilizing" actions.) Both of these actions could again be described via a 3-tuple (x, y, E) indicating the positions and amounts of food/energy involved. For now we will leave these possible actions out of the system, but for completeness later we anticipate including them in the evolutionary pot, to see whether or not any creatures find them adaptively useful. (It is important to try to begin with as few possible actions as we can, and let the selectable pool of them build up slowly, preferably also through a non-arbitrary evolutionary process.)

Another thing in the world that a creature can change is itself. At present, we do not allow a creature to alter its own behavior-generating mechanisms, that is, its list of usable sensors, internal state variables, and actions and triggers; in particular, as mentioned earlier, since the behavior-generators are fixed, no learning is possible yet. (Note that the *values* that the sensory, internal state, and trigger functions return can all be affected by what the creature does, since its actions can change its internal states and its view of the world, but the *functions themselves* cannot be altered, except by the action of evolution.) But a creature can change its own (x, y) location, moving about in the world. Just as for the eating action, questions arise as to how far and in what directions a creature can travel in a single time-step. Again for now, just to keep things simple, we will only allow single-hex movement to any of the positions currently adjacent to the creature, making six movement action-types.

With more than one organism in the world at a time, there comes the possibility for a wide range of other "social" (or anti-social!) behaviors, as one creature affects another: for instance, "move-other" or "eat-other." Once again, to start off simple, we are setting these types of creature interactions aside for the moment. But there is one final aspect of the world that the creatures should be able to influence: the *number* of the creatures themselves. To this end, we add a final action-type, "split." When a creature performs this action, it creates one or more copies of itself as new individuals in the world. In this case, we need to determine how many new individuals are created, how much energy each one gets, what location each one ends up at, and whether or not the new copies are exact or altered versions of the original. In our simple first instantiation of this system, we allow an individual creature to just split into two, with each getting an equal amount of the "parent's" energy, and each occupying the same hex as the original. Finally, one of the two new creatures that replace the original is kept identical to the original, and the other is mutated slightly, so that new behavioral mechanisms can enter the population. Without such mutation, the system would be more or less static, with no evolutionary change possible. In future versions of the system, we can imagine sexual reproduction through a "combine-with-other" action-type, as well as more morbid behaviors like "mutate-self," "mutate-other," and "split-other"...

We have described all the possible action-types in our simple world, in terms of all the things in the world that can be changed. As more types of objects are introduced into the world (more categories of entities, c, as described earlier in section 2.2.1), and more inter-organism interactions are allowed, correspondingly more action-types for changing each of those things will be created. By allowing *any* possible change in the world and its contents to be a potential action-type (subject to the bioenergetic constraints of the world), we hope to enable every possible type of action in these simple worlds to appear, whether or not they have analogs in the "real" world. Such generality is essential for the study of adaptive agency (Miller & Todd, 1990), construed as the investigation of adaptive behavior by *any* sort of behaving agent, whether real or hypothesized, terrestrial or extraterrestrial, biological or artificial.

2.4 Bioenergetics of the World

As we indicated in the previous section, every action a creature can perform in its simulated world may have energetic side-effects, in terms of changing the distribution of energy in the world (that is, transferring it from one entity to another). In particular in the current simple world, these effects will be manifested as raising or lowering the individual organism's own personal energy level, E, and changing the energy level in food-plants. How each action-type affects the world's energy-distribution is determined by the bioenergetics of the world. For now, we will just assign the energy side-effects of actions by fiat, but we will describe briefly at the end of this section a way in which the bioenergetics, like the behaviors themselves, can emerge during the evolutionary process.

The main rule governing the bioenergetics of our system is that *nothing any creature does can raise the energy in the world*. The only way the total world-energy can ever increase is via the food/growth function, $\phi(x, y, t)$. Energy can be transferred from one entity to another, and certainly can be lowered or lost, but never increased through creature actions. If we did not impose such a restriction, then obviously the best thing a creature could do in its lifetime (in terms of survival or ongoing existence) would be to increase its own energy level directly, resulting in essentially a perpetual-motion organism. The only other global rule is that once an individual's energy level E drops to (or below) zero, it dies and is removed from the world and the simulation.

In the previous section, we ended up allowing our initial creatures to perform only three main action-types: eat everything from the current hex, move to one of the adjacent hexes, and split into two. Thus for the energetics of this world, we need only specify the energy side-effects associated with each of these actions. (In general, the bioenergetics of a given world consists of a table or list of all the action-types possible in that world, along with their associated energetic consequences.) For eating, the energy effects will simply be to lower the individual's energy by an eating-exertion cost (for example, 0.5), and then increase the individual's E by the amount of energy absorbed from the food in its current hex. Furthermore, the food-energy in that hex is set to zero (since it is all eaten). The eating-exertion cost guarantees that creatures that always just try to eat, even when there is no food around, will eventually run out of energy and die. Thus the creatures must evolve to only try to eat when it's adaptive, that is, when there *is* food present, rather than attempting to constantly shovel everything around them down their gullets.

For movement, we will charge the creature (i.e., lower its energy E) a movement-exertion cost (e.g. 1.0) for each step in any direction. For splitting, we impose a splitting-exertion cost (e.g. 1.0) on the parent creature, and divide the remaining energy equally between the two resulting offspring. Finally, if the creature does nothing in a given time-step, we still impose a small energy sloth-cost on it (e.g. 0.3), so that completely useless sedentary organisms will eventually be weeded out, dying and being removed when their energy slowly leaks away to nothing.

All of the other action-types that we chose not to include in our initial system could in fact appear in the bioenergetic specification of the world. But they would have a prohibitively high energetic cost associated with them so that no creature could ever perform these actions. In general, this is how particular action-types are "disallowed" in our system, through unpayable energy penalties for performing them. Thus when we introduce the general location-energy 3-tuple (x, y, E) action-type form, we will be able to allow or disallow "reaching" in the world by how we assign energy costs to actions with $x \neq x_p$ and $y \neq y_p$; similarly for movement beyond adjacent hexes or multiple-offspring splitting, etc. (Note that there is no energy cost associated with either sensing or performing cognitive computations, e.g. in the internal state variables, but this could also be introduced at some point, no doubt with profound effects on the sensors and behaviors evolved.)

The energetic costs and benefits of the actions we have described here have all been set by hand, in what we hope are reasonable ways that will allow interesting organisms and behaviors to evolve. But we would be happier if this aspect of the world as well could emerge through evolution, rather than being predetermined in what might be an ultimately uninteresting way. One idea we have for how to allow such emergence of the basic bioenergetic laws of the world is to let every organism "vote" or "bid" for the energy costs and benefits that *it* would like to see associated with various action-types. If we have only roving, eating creatures evolving in the world as we have so far described here, we can imagine that they will all vote or bid to lower the movement cost to 0, so that they can roam about the world eating over as long and far a range as they want without incurring any cost. This would not be particularly interesting. However, if we also allowed the *food itself* to evolve, akin to say plants or sedentary or immobile prey organisms, and let *these* entities also vote on the bioenergetics of the world, they would no doubt call for very *high* movement costs, to cripple the roving eaters and keep them from getting close enough to devour the food organisms.

With both organism types voting, each with different self-interests, and with the bioenergetics coming out of this conflict-based democratic process, we expect that much more interesting dynamics would emerge. Perhaps some sort of energetic cost/benefit compromise would be settled upon, allowing an uneasy truce where individuals in both species could survive; or maybe there would be a continuous predator-prey cyclic pattern as first one faction and then another gathered enough votes to sway the bioenergetics temporarily in their own favor. In this way perhaps "evolution at the edge of chaos" (Langton, 1992) could be achieved and maintained, along with the ongoing interesting patterns that this entails.

3 Evolutionary Dynamics and a Day in the Life

Creatures in our simulations obviously have rather simple lives. To begin each run, we just create a random initial organism via abiogenesis (generation of life from inanimate substance), with random sensors, internal state variables, and actions and triggers, as well as a specific amount of energy to get it started. We then set it loose in the world and see what happens. We introduce new such random creatures by abiogenesis every so often (after a typically fixed interval of many time-steps), to keep up a constant trickle of potentially adaptive creatures into the world. Some of these creatures will move about in the world, but do little else. Others will sit and eat, but never move. Still others will split repeatedly, until their energy is too low to allow any further splitting. All of these will burn up or waste away their energy until they reach zero and die, with their ineffectual behavioral endowments being removed from the world. Some lucky few in any given run though may combine these actions in an appropriate fashion, moving, then eating, then splitting and spreading their successful behavior-generating mechanisms to new offspring. In turn, some of the new offspring, mutated slightly, will improve upon the behaviors of their parents, and more and more adaptive behavior will evolve over time. In any given time-frame, this emergence of surviving creatures may not occur; but as we continue to introduce new random creatures into the world, eventually an interesting (i.e. adaptive) one will show up, and the system will take off, bubbling with life.

There are two important subtleties to this evolutionary scheme. First, reproduction is not imposed on any individual. For evolution to work, some of the creatures must hit upon "reproductive behavior" (i.e., splitting) themselves. Otherwise, the best they can do is just be long-lived. An immortal creature that works its way around the world and eats enough to remain alive time-step after time-step may be impressive, but without such creatures splitting and having slightly mutated offspring, behavioral improvements will not be introduced into the population. Furthermore, our creature-controlled reproduction scheme allows a fitness measure to emerge naturally from the simulation. Fitness will be defined exactly in terms of those creatures who behave appropriately to produce the most offspring. (This is in sharp contrast to the usual fitness measures imposed from above by hand, e.g. by forcing any creature above a certain energy threshold to split, or by running the simulation in generations and giving each individual a number of offspring in the next generation proportional to its amassed energy.) Note that with our current method, we will also have (perhaps widely) fluctuating population sizes; however, this should not be seen as a disadvantage, but rather as an added measure of realism.

Second, one of the most important things that a creature can know about its world, namely, its own energy level E, is never directly available to it. An individual's energy level could be crucial for knowing when to split, when to head toward food, when to sit and wait for it to grow back, etc. However, we do not allow any creature direct access to this important piece of information. Rather, if creatures are to make use of this information, they must evolve an internal representation of it which is updated appropriately in response to the actions the individual organism emits. Thus, if an individual eats, it should increase its internal assessment of its energy level by an appropriate amount; if it moves or splits, it should likewise lower its energy-level internal variable. We expect the most adaptively-behaving individuals will make use of just such an internal constructed bit of "self-knowledge," so we are on the lookout for any such feature in our evolving scenarios.

4 Actual Implementation Methods

The behavioral evolution system we have presented here is intended to be implementable by a variety of methods; the description has been purposefully placed at a high level to be agnostic in terms of implementation details. We are focussing our own efforts though on two very different types of implementations which we will describe briefly in this section: evolving Lisp routines, and evolving neural network architectures.

The lists-of-functions representation we used for the actions and triggers, internal state variables, and sensors earlier is obviously readily translatable into Lisp code and s-expressions. In such a form, we can use a "genetic programming" method to try to evolve these functions, such as that of Koza (1992) or Sims (1991). Since our creatures currently just split and mutate, the latter method may be more suited to our purposes at present, because it emphasizes mutational changes to the Lisp code, rather than the recombination of s-expressions primarily used in the former. In either case, while the translation of our system into Lisp may be clear, the evolutionary process itself could still be quite extended and tortuous.

When compared to the obvious Lisp representation, it may seem very unclear how we could construct neural network architectures that will precisely embody the modularized lists of sensors, internal state variables, and actions and triggers that we want our creatures to be able to use. But there is also a natural interpretation in this case, in terms of a structured network with each layer of nonlinear units interpreted as one of the three categories just mentioned. First of all, we can have a set of possible output unit types, each of which is associated with one of the possible action-types. Every network will have an evolutionarily-determined selection of these. Each output unit in every creature has an evolved bias level associated with it, which forms part (or all) of its triggering function. Actions are selected from among the active output units based on one of the selection schemes listed in section 2.2.3. If the network has only such output units, as shown in Figure 2a, and uses the probability-based action selection scheme, we will end up with a stochastically-behaving creature like that described in the first scenario in section 2.2.

Now we can add the internal state variables as another layer of (hidden) units in the network, as shown in Figure 2b. These units receive recurrent connections from the output units, and from themselves, and pass their logistic-function activations forward in turn to the output units. In this case, the evolutionary mutation process would determine not only the number and type of output units, but also the number of internal state units, and the weights on all the connections between these two layers. With such a network, we would get the memory-guided creature described in the second scenario in section 2.2.

Finally, we can add sensors to this network as another layer of (input) units, as shown in Figure 2c, and end up with the full-blown sensory-guided behavior of the fourth scenario. As can be seen in the figure, each sensor unit (which is sensitive to one particular category) has direct connections to positions in the world, that is, $W(x, y)$. Sensors are in turn connected both to the internal state units, and to the output units, since both are functions of the current sensory input (as described in sections 2.2.2 and 2.2.3). Evolution will determine the number and types of sensory units in each creature's network, and the connectivity pattern and weights between the world and the sensors, and between the sensors and the internal state units and output units. (The final recurrent structure of the full network here resembles that of an Elman- or Jordan-style recurrent network — see Elman, 1988, and Jordan, 1986. Here we have drawn the

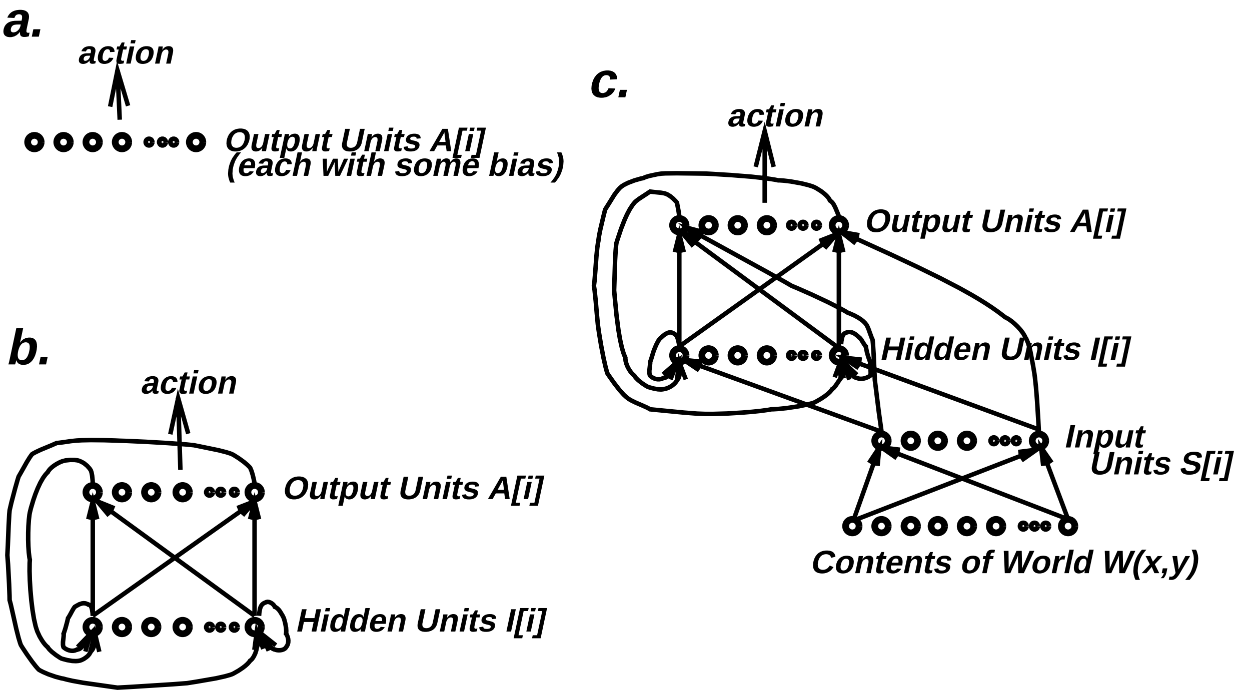

Figure 2. Network implementations of behavioral components. a. A simple network for producing actions stochastically. b. A network for generating actions based on internal states. c. A network that uses both internal states and sensory inputs from the contents of the world to produce its actions.

input units offset, to emphasize their secondary role in guiding and adjusting internally-generated behavior.)

For the genetic and evolutionary representation and modification of these neural networks, we use a scheme like that developed by Miller, Todd, and Hegde (1989). Units can use various activation functions, allowing a wide range of nonlinear functions to be implemented between different variables in the three layers; but obviously this framework does not allow the unlimited flexibility that evolving entire Lisp routines does. Still, we believe it will be flexible (and manipulatable) enough to yield very interesting results in this system (as it has in many earlier studies), and so we are eager to compare the results of this representation scheme to the Lisp-based one. (If further computational power is needed in the network, we can add additional hidden layers of computing units between the existing layers.)

5 Characterizing the Environment, and Other Research Directions

At this point, we do not have space (nor results) enough left to do much other than describe some of the types of environments, and environmental classifications, that we are beginning to explore with the framework we have just laid out. We choose to consider two-dimensional grid world environments in general because they are easily visualizable and concrete (rather than described by unintuitive parameter settings). They should also be more readily translatable into real-world settings so that we can tie in with empirical, ecological data. However, classifying these environments will probably require returning to abstract parameters and categories; but we will maintain the link to the grid world description and the valuable visualization it allows.

Traditional machine-learning conceptions of environments, as “noisy,” “regular,” “unknown,” etc., are crude and unhelpful in our current endeavor. Slightly more sophisticated conceptions of environments as finite state automata (FSA’s), which return sensory inputs to the creature in response to its motor outputs, can describe quite complex environment/behavior interactions (see, e.g., Mozer & Bachrach, 1991; Rivest & Schapire, 1987). We are interested in developing mappings between such models of environments, and our 2-d grid world representation. However, there are at least three drawbacks to the FSA approach to environmental characterization: first, for environments of reasonable size (like our grid worlds), the FSA will have to be enormous to capture all the possible states and transitions between them (e.g., for a small 10x10 grid which either contains food or nothing at each location, there are 2^100, or approximately 10^30, possible states); second, the actions and sensations are all assumed to be defined, rather than allowed to emerge through the course of evolutionary adaptation; and third, in focussing on the sensory consequences of actions, the FSA model overlooks the interesting possibilities where no sensors are used.

Wilson (1991), Smith (1991), and Littman (these proceedings) have proceeded from the FSA environment models to more abstract categories based on the amount of memory or state a creature needs to exploit an environment, and how long the creature must wait for rewards following its behavior. Littman carries this further, characterizing environments in terms of the creatures that can behave optimally in them. These approaches also seem promising, and certainly emphasize the interrelationship between environmental structure and the creatures that behave in those environments. But it is not easy to see how to move back and forth between these conceptions of environments and more concrete visualizable forms like the grid world representation. Also, the notion of reward, while important for learning systems, again obscures the important classes of non-learning, and even non-sensing, creatures and behaviors, as well as their evolution. And finally, we believe that it should be possible to describe some aspects of the environment in terms that are not entirely creature-centric, that is, in ways that do not rely solely on the capabilities of creatures themselves; but we may be forced to recant this position.

An example classification scheme closer to what we are hoping for was indicated in the third scenario in section 2.2, where we spoke of food being distributed in the world in separate clumps or patches. We are working on a way of parameterizing the clumpiness of food in a 2-d world (which we have already done in one way for the 1-d case — see Todd, 1992, chapter 5). With this characterization in hand, we will be able to construct worlds of different clumpinesses, and see what sorts of behaviors evolve to exploit them, and whether there are interesting correlations between the two. Patch finding and foraging is a well-studied area in the animal behavior literature, so we will have much to draw on and compare to there.

As we have mentioned earlier, we are also interested in how well creatures without various of the behavioral components might fare in different environments. Thus we might compare creatures without internal state to those with, or ones with limited sensory abilities to others with sophisticated senses. Another interesting variation would be to explore the tradeoffs that evolution makes when it must choose *between*, say, senses and internal states: by keeping the sum $n_S + n_I$ fixed during evolution, we could see how evolution balances the two when it has limited resources to work with.

We are clearly embarked on an exploratory, open-ended research program, but one which we believe will yield useful insights into the nature of environments and the behaviors that are adaptive in them. By starting at the ground-level, creating a system in which all aspects of a creature's behavior-generating mechanisms can evolve as needed, and working our way up to consideration of different types of environments where this evolution will take place, we hope to have a framework rich enough to lead us to understanding of phenomena beyond those we have already imagined.

References

Ackley, D., and Littman, M. (1992). Interactions between evolution and learning. In C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (Eds.), *Artificial life II* (pp. 487-509). Redwood City, CA: Addison-Wesley.

Cariani, P. (1990). Adaptive connection to the world through self-organizing sensors and effectors. In A. Meystel, J. Herath, and S. Gray (Eds.), *Proceedings of the 5th IEEE International Symposium on Intelligent Control*, vol. 1 (pp. 73-78). Los Alamitos, CA: IEEE Computer Society Press.

Elman, J.L. (1988). *Finding structure in time.* Technical Report 8801. La Jolla, CA: Center for Research in Language, University of California at San Diego.

Gallagher, J.C., and Beer, R.D. (these proceedings). A qualitative dynamical analysis of evolved locomotion controllers.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston: Houghton-Mifflin.

Holland, J.H. (1992). *Adaptation in natural and artificial systems.* Cambridge, MA: MIT Press.

Jordan, M.I. (1986). *Serial order: A parallel, distributed processing approach.* Technical Report ICS-8604. La Jolla, CA: Institute for Cognitive Science, University of California at San Diego.

Koza, J. (1992). *Genetic programming.* Cambridge, MA: MIT Press/Bradford Books.

Langton, C.G. (1992). Life at the edge of chaos. In C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (Eds.), *Artificial life II* (pp. 41-91). Redwood City, CA: Addison-Wesley.

Littman, M.L. (these proceedings). An optimization-based categorization of reinforcement learning environments.

Maes, P. (these proceedings). Behavior-based artificial intelligence.

Miller, G.F. (1992). *The evolution of protean behavior through psychological selection: Adaptive unpredictability, sexual selection, and human encephalization.* Unpublished doctoral thesis. Stanford, CA: Psychology Department, Stanford University.

Miller, G.F., and Todd, P.M. (1990). Exploring adaptive agency I: Theory and methods for simulating the evolution of learning. In D.S. Touretzky, J.L. Elman, T.J. Sejnowski, and G.E. Hinton (Eds.), *Proceedings of the 1990 Connectionist Models Summer School* (pp. 65-80). San Mateo, CA: Morgan Kaufmann.

Miller, G.F., and Todd, P.M. (in preparation). *A theory of perception based on fitness affordances.* Unpublished manuscript. Stanford, CA: Psychology Department, Stanford University.

Miller, G.F., Todd, P.M., and Hegde, S.U. (1989). Designing neural networks using genetic algorithms. In J.D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 379-384). San Mateo, CA: Morgan Kaufmann.

Mozer, M.C., and Bachrach, J. (1991). SLUG: A connectionist architecture for inferring the structure of finite-state environments. *Machine Learning, 7,* 139-160.

Pattee, H.H. (1989). Simulations, realizations, and theories of life. In C.G. Langton (Ed.), *Artificial life* (pp. 63-77). Redwood City, CA: Addison-Wesley.

Rivest, R.L., and Schapire, R.E. (1987). A new approach to unsupervised learning in deterministic environments. In P. Langley (Ed.), *Proceedings of the Fourth International Workshop on Machine Learning* (pp. 364-375). San Mateo, CA: Morgan Kaufmann.

Sims, K. (1991). Artificial evolution for computer graphics. *Computer Graphics, 25,* 319-328.

Smith, R.E. (1991). *Default hierarchy formation and memory exploitation in learning classifier systems.* Unpublished doctoral thesis. Tuscaloosa, AL: Department of Engineering Mechanics, University of Alabama.

Todd, P.M. (1992). *The evolution of learning: Simulating the interaction of adaptive processes.* Unpublished doctoral thesis. Stanford, CA: Psychology Department, Stanford University. (Available from UMI.)

Todd, P.M., and Miller, G.F. (1991). On the sympatric origin of species: Mercurial mating in the Quicksilver model. In R.K. Belew and L.B. Booker (Eds.), *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 547-554). San Mateo, CA: Morgan Kaufmann.

Todd, P.M., and Miller, G.F. (submitted). *Evolution "without" natural selection: Sexual selection for speciation and learning.* Submitted to the Proceedings of the Third Artificial Life Conference.

Werner, G.M., and Dyer, M.G. (these proceedings). Evolution of herding behavior in artificial animals.

Wilson, S.W. (1991). 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 Simulation of Adaptive Behavior* (pp. 15-21). Cambridge, MA: MIT Press/Bradford Books.