Exploring Adaptive Agency II:
Simulating the Evolution of Associative Learning

Peter M. Todd
Department of Psychology
Jordan Hall Building 420
Stanford University
Stanford, CA 94305
todd@psych.stanford.edu

Geoffrey F. Miller

Abstract
We consider psychology as the study of adaptive agency, investigating the processes and mechanisms resulting in fitness-increasing behavior in the world. A central issue in psychology so construed becomes: what are the relations between the primary adaptive process of evolution by natural selection, and the adaptive processes psychologists call ‘learning’? In particular, under what conditions would learning evolve? To explore this issue, we use genetic algorithms to simulate the evolution by natural selection of neural networks, which in turn control the behavior of simple creatures in virtual environments. We have developed what we consider the simplest possible environmental challenge in which unsupervised associative learning could prove adaptive: ‘bootstrapping’ the learned use of one highly accurate, but individually varying, sensory modality by another less accurate, but evolutionarily stable, modality. We have found a possibly quite general U-shaped curve relating the time (in number of generations) to evolve the use of unsupervised learning on the varying ‘bootstrapped’ modality, to the accuracy of perception in the stable modality which guides this learning. This U-shaped curve appears to represent a trade-off between the adaptive pressure to evolve learning (which peaks when perception accuracy in the stable guiding modality is at chance) and the ease of learning during a given lifespan (which peaks when this accuracy is perfect.)

1 Introduction
Natural selection has constructed animals’ minds and behavioral repertoires for adaptive fit to the environmental problems they must face. As the study of such minds and behaviors, psychology should focus on the notion of adaptive agency -- the generation of action in the world in response to challenges to individual fitness. This framework encompasses many approaches, including (1) the elucidation of complex species-typical adaptations (as in human and animal experimental psychology and cognitive ethology), (2) the comparison of psychological adaptations across species and consideration of their phylogenetic origins (as in comparative psychology), and (3) the general exploration of the adaptive processes themselves that yield adaptive agency (e.g. by simulation methods, including those in the field of artificial life -- see Langton, 1989). In the current paper, we consider the phenomenon of ‘learning’ as an aspect of adaptive agency, by investigating via evolutionary simulations some conditions under which the ability to learn may prove adaptive and so spread through a population. The theoretical, historical, and methodological background for this work is presented more extensively in Miller and Todd (1990), and further extensions of this method applied to habituation and sensitization as adaptations to short-term environmental dynamics appear in Todd and Miller (in press).

Evolution as an adaptive process has itself undergone changes: "survival of the stable" probably preceded "survival of the fittest" (Dawkins, 1976). Evolution in the earth’s early environment is likely to have selected for physical systems with relative stability in the shifting primordial soup. After stability came replication and metabolism: the ability to turn external material into copies and extensions of oneself. With the evolution of larger, more complex phenotypes came the opportunity to evolve behavior-generating systems that could produce innately programmed sequences of activity and movement. Sensory systems could then evolve to guide these behavior-generators more adaptively, based on sensitivity to particular environmental cues. Thus, blind activity based on preprogrammed behaviors may have preceded sensory-guided reactivity, that is, the ability to adaptively adjust to the current changing environment on a moment-by-moment basis. Only after these first two stages had evolved could a further adaptive process evolve -- ‘learning,’ defined as the ability to make long(ish)-term adaptive changes in behavior-generators in response to the environment. In this hierarchy, learning emerges not as the primary adaptive force it is sometimes assumed to be, but rather as a tertiary one, following genotypic evolution and short-term environmental reactivity (see also Shepard, 1987, 1989). Once we re-conceptualize ‘learning’ as merely one process among several that generate adaptive agency, the questions we might ask about this

1 By this definition, learning includes such processes as experience-guided development not commonly included in this category. For examples of such processes, see Krudsen, 1988, and Singer, 1984, 1988.
process begin to change as well.

Exploring adaptive agency by simulation thus requires simulating genotypic evolution, the generation of behavior, and the ability to change behavior generators by learning. Specifically, we use a genetic algorithm to evolve successive generations of a population of neural network architectures, which in turn control the behavior of simple creatures which can learn as they live in a simulated environment. This is an extension of earlier work exploring the use of genetic algorithms to design network architectures capable of learning specific input/output mapping tasks (Todd, 1988; Miller, Todd, & Hegde, 1989); here, the algorithm's measure of fitness depends not on learning an arbitrary task, but on behaving adaptively in the simulated environment.

Traditionally, researchers using simulation in this area have started by assuming learning as a primary adaptive process, and then asked how might evolution interact with, shape, and be shaped by learning? Examples include the work of Hinton and Nowlan (1987) and Belew (1989), who explicate the Baldwin effect in which learning "guides" evolution; and Stork and Keasing (1990) and Belew, McNerney, and Schraudolph (1990), who investigate the ways evolution and learning can combine to affect the initial structure of neural networks. But considering the hierarchy of adaptive processes spelled out above, the question we want to ask is, rather, given the already-powerful adaptive processes of genotypic evolution and environmental reactivity, under what conditions would the tertiary adaptive process of learning ever prove useful in terms of increasing individual fitness? Assuming no learning, when should learning evolve? Why should this extra level of adaptiveness exist?

(For analysis of learning mechanisms as evolutionary adaptations, see Pulliam & Dunford, 1980; Dawkins, 1983; and Cosmides & Tooby, 1987.)

To answer such questions, 'learning' must be viewed not as a single monolithic process, but as the diverse set of distinct mechanisms, abilities, and dynamics that ethologists and psychologists have shown it to be (see Tinbergen, 1951; Lorenz, 1966; Hinde & Stevenson-Hinde, 1973; Bower & Hilgard, 1981; and Staddon, 1983). The adaptive (fitness-increasing) functions of each type of learning must be considered separately, with attention to the type of organism, ecologische, and adaptive problem in question; this is the approach followed in the field of ecological learning theory (Davey, 1989), from which we take inspiration. In simulations, we must first specify some environment and what defines fitness in that environment, and then observe which adaptive processes are most important for maximizing individual fitness in "evolving" the relevant environmental problems. Ackley and Littman's (1990) simulations go far to address these issues, showing the utility of evolving motivational systems to drive learning, but still his simulations either operate with learning or without, rather than setting up an environment whose adaptive problems allow the evolution of learning itself to be studied.

2 Adaptive Functions of 'Learning' Processes

Why should an organism allow environmental conditions to change how its behavior-generating mechanisms work (by 'learning?'), rather than allowing natural selection to optimize those mechanisms (by 'hardwiring') just as it's has optimized so many other physical adaptations (see Mayr, 1974; Staddon, 1983; Menzel, 1984)? We see two main adaptive functions for learning. First, and perhaps most importantly, 'learning' may serve to increase an organism's 'developmental leverage', allowing it to build a larger, more complex, more finely organized phenotype given a certain size genotype. To the extent that an organism can sense predictable environmental regularities during neural development, for example, it could use the resulting sensory activation patterns to guide the self-organization of its behavior-generating mechanisms (e.g. Singer, 1984). Learning may allow the genotype to 'store information in the environment' and let environmental regularities do much of the hard work of wiring up adaptive behavior-generators. The environmental regularities used in this way may take any abstract form. For example, parental 'imprinting' in birds (Lorenz, 1935/1970; Immelman, 1984) can be viewed as a way of building a behavior-generator sensitive to the appearance of one's parent, based on the following environmental regularity: the first large moving thing one sees after hatching is very likely one's parent. Of course, the particular behavior-generator constructed by different birds for recognizing their parents will be different (the birds will 'learn' different parental images), but the species is a whole relies on the same environmental regularity to do the construction.

Second, 'learning' as sensitivity to dynamic changes in environmental conditions throughout life may allow organisms to adjust their behavior-generators faster than natural selection can track (see Belew, 1989). In this case, learning adapts phenotypes to ongoing changes and particularities of the environment, rather than building on environmental regularities during phenotype-construction. This is probably the adaptive function of learning most familiar to adult humans (e.g. learning a new restaurant location, or a new person's name), and therefore more commonly studied by psychologists, but it is likely to be less common than the phenotype-construction function throughout the animal kingdom (see Miller & Todd, 1990).

For our simulation studies of the evolution of learning, we attempted to devise the simplest, cleanest scenario in which learning could prove adaptive, focusing on the phenotype-construction function. If we make the genotype unable to know ahead of time which one of two alternate econiches it will be born into, natural selection will be unable to choose the proper behavioral mechanisms to guide the phenotype's actions. Instead, it must select for the evolution of more general mechanisms that can flip into one of two states depending on which econiche the genotype finds itself in during phenotypic development. Thus, we settled on a kind of "imprinting" or parameter-setting based on the early environment as the simplest possible case in which learning, construed as adaptation to specific environmental regularities, could evolve.
3 Inclusive Fitness as the only Natural 'Supervisor' for Learning

Several previous efforts using genetic algorithms to evolve neural network architectures have evaluated architecture fitness by training the networks with a supervised learning procedure (i.e. one with an externally provided "target value" the network is to produce), typically backpropagation (Miller, Todd, & Hegde, 1989; Belew, McAdams, & Schraudolph, 1990; for a review, see Weiss, 1990). While supervised learning paradigms may be appropriate when engineering connectionist systems for particular industrial applications, they are problematic and perhaps misleading in scientific studies of adaptive agency. In particular, to be biologically plausible, the source of the "targets" must be justified. Organisms as whole functioning agents in real environments rarely receive patterns of information analogous to training signals in back-propagation (see Miller & Todd, 1990). Although the distinction between supervised and unsupervised learning procedures can be blurred, we have chosen to focus on the more defensible latter end of the spectrum, including self-organizing, associative, and simple feedback-based mechanisms. But even if we sidestep the issue of target-based training, the concept of feedback still raises problems.

Years of learning by "being taught" instill in us intuitions as to the usefulness of corrective feedback to promote learning. But such intuitions make it easy to overlook the fact that it is at least as difficult for natural selection to evolve the ability to perceive feedback signals from the environment to guide an organism's learning, as it is to evolve the perception of any other external stimuli. Consider for example the complexities involved in registering the information that one has just been rebuffed in a social exchange. Feedback signals cannot be assumed to be just somehow "provided" to an organism for it to use in adjusting its behavior. Instead, feedback systems must be understood as special sensory systems evolved to provide information to special learning mechanisms that in turn adaptively change the functioning of certain behavior generators. These feedback systems are likely to be complex articulated systems themselves separate from the behavior generators; it is neurologically implausible that feed-forward signals are sent in one direction along axons while error (feedback) information is sent in the other. (Other sensory systems bias those behavior generators more directly and temporarily in response to ongoing environmental events -- reactivity, as we described earlier.) Feedback systems, whether motivational, emotional, volitional, or proprioceptive, evolve just like other aspects of adaptive agency -- by cumulative selection of incrementally better-adapted designs.

What an organism would like to have to guide its learning is direct information about how its inclusive fitness changes as a result of its behavior. But there is no such thing as an inclusive fitness transducer that can be used to supervise learning. Organisms must instead evolve to sense inclusive fitness indirectly, through whatever proximal sensory cues are reliably associated with increased fitness in their environment. Thus, natural selection itself is ultimately the only source of supervision for learning systems.

Natural selection's indirect supervision of learning leads to the complexity inherent in evolving learning systems. The learning scenario we present in this paper illustrates this point. To perform even the simplest kind of unsupervised associative (or correlational) learning we address, with real adaptive consequences through behavior rather than just forming perceptual categories, the following building blocks are required: (A) one input modality that has evolved to sense an environmental cue that has been reliably associated with a change in inclusive fitness; (B) another input modality that has evolved to sense a different cue which has not been reliably associated with a change in fitness over evolutionary time, but which can have such an association for some time during the organism's lifespan; (C) an output modality (effector) capable of generating some behavior that has fitness consequences related to the environmental situation sensed by A; (D) a connection between input A and output C with an evolved behavioral valence indicating whether or not that environmental cue should activate or inhibit that behavior (e.g. should there be a positive or negative weight from A to C?); and (E) a connection between input B and output C, whose valence can be built up by association with the activation of C caused by A. In the simulations to follow, we have limited the maximum processing capabilities of the creatures as closely as possible to these building blocks, as we show in section 5.

4 A Simple Scenario for the Evolution of Unsupervised Learning

To investigate conditions under which learning might evolve, we wanted to start with the simplest simulated world we could devise where learning could prove adaptive, and then see if it would spread through a population of simulated creatures behaving in that world. We explain the scenario by outlining a biological metaphor that identifies the structure of the econiche and the nature of the adaptive problem. The actual simulation methods used to instantiate this metaphor are presented in the next section.

Our scenario can be imagined as an underwater realm, in which parent creatures produce gametes that float around randomly, possibly undergoing sexual recombination with other gametes, until they fall to a particular feeding patch, where they hatch into offspring. The offspring then attach to the substrate, grow as sessile (stationary) feeders, and attempt to ingest the appropriate kinds of substances that float past. The offspring live for a fixed lifespan, and just before death they produce some number of gametes of their own, based on the amount of food they've eaten (which thus determines their fitness). These creatures thus never interact with each other (only their gametes may interact genetically), and at each moment in their life they merely decide whether or not to eat the particular substance currently in their vicinity.

The floating substances come in two forms, food and poison, that appear equally often (same base rate) and at random. Eating food increases the creature's internal energy store (by an arbitrary 10 points) and, since the average number of gametes released is proportional to lifetime energy accumulation, it likewise increases potential fitness. Eating poison decreases the creature's energy by the same
amount (10 points), representing the metabolic activity necessary to counteract its effects, and thus reduces lifetime fitness. If a creature decides not to eat at a particular time, its energy (and hence its fitness) remains unchanged. Once a creature’s lifespan ends, the average number of gametes released (and thus average fitness) is proportional to the amount of food eaten minus the amount of poison eaten. Thus, those creatures with "brains" (simulated by neural networks) that allow them to eat food and avoid poison will save up more energy and have more offspring, so that their beneficial brain-designs will spread through the population.

Food and poison can be distinguished by olfactory and visual features. Food is sweet and poison is sour across the entire world, forever. However, turbulence in the watery environment may carry odors in such a way that a creature receives more sour odors than sweet odors even though the nearest thing is really food. Turbulence thus decreases the accuracy of the smell sense, so food might not always smell sweet, at a given instant for a given creature, and poison might not always smell sour. We define a smell-sense accuracy for a given world (based on the degree of turbulence in the water) which determines the percentage of time each substance has its proper smell for the creatures. For example, 75% smell-accuracy means food smells sweet to each creature a random 3/4 of the time, and sour the other 1/4 of the time, and vice-versa for poison.

Food and poison also have associated colors: red or green. But this world is divided up into different food patches, and in half of these patches, food is always red and poison always green; in the other half, food is green and poison is red. Thus, during each sessile creature’s lifetime, food will always be one color and poison the other. But the mapping from color to object will vary between individual creatures, depending on which kind of food patch they happen to land in at birth. The creatures’ visual sense, though, unlike their olfactory sense, is completely accurate -- unaffected by the turbulence -- so that food and poison always appear their true color for that patch.2

To summarize, then, each creature in this world lives a fixed lifespan, eating or ignoring food and poison, and amassing energy which helps determine its eventual number of offspring in the next generation. Food smells sweet and poison sour across all creatures, but with some error rate determined by the smell-sense accuracy. Food and poison each have characteristic fixed colors within one creature’s life, but the meaning of each color varies between creatures, food being red for some and green for others. Thus natural selection can predict the association between smell and substance, but not between color and substance -- this will be the task for learning.

5 The Genetic Algorithm for Evolving Neural Networks

To simulate the evolution of learning in the scenario just described, we use a relatively standard form of Holland’s (1975) genetic algorithm, combined with a simple “developmental” scheme which translates genotypes into neural network architecture phenotypes. Each network, representing an individual creature, is evaluated in the simulated world over several time-steps (e.g. 200) representing the creature’s lifespan. During each time-step, the network receives sensory input based on the current environmental cues available to the creature, processes that input according to its architecture and current weights, generates behavior based on the activation of its output units, and changes its connection weights using an unsupervised correlational learning rule (e.g. Hebbian association). The effects of the creature’s network-generated behavior on the world and its own fitness are then registered, and the next time-step begins.

The genotypes of all initial populations are randomly-generated bit-strings. A strong genetic specification scheme (as defined by Miller, Todd, & Hegde, 1989) interprets each genotype as a connectivity constraint matrix that directly specifies the nature of each unit and connection in the network architecture. (Given the extremely simple design and small size of the networks investigated, more powerful development schemes seemed inappropriate.)

For the current scenario, we use the simplest network architecture we can: three units connected in a feedforward manner (unit 1 to units 2 and 3, and unit 2 to 3 alone). This compact genotype severely constrains the amount of neural “machinery” evolution has to play with in any given creature. However, this size is sufficient to incorporate the five building blocks necessary for simple associative learning as described in section 3, since all we need are two input units (A and B), an output unit (C), and connections between them (D and E).

To fill in this architectural structure, genotypes can specify six possible types of units: four possible types of input unit, a hidden unit, or an output unit. Different input units can sense sweet smell, sour smell, red color, or green color, being activated if their corresponding stimulus is present; hidden units compute the standard logistic function of their summed inputs; and output (motor) units generate eating behavior if turned on over some threshold level, or indicate abstinence if not. Genotypes also specify feedforward connections between and biases on the units (no recurrence is allowed in the simulations discussed here). Connections and biases can have positive or negative weights of four possible magnitudes (0, 1, 2, 3); and each is either genetically fixed or learnable via a simple Hebbian rule: \( \Delta w_{ab} = lr \ast a_a \ast a_b \), where the change in connection strength \( \Delta w_{ab} \) from unit a to unit b is equal to some learning rate \( lr \) times the activation \( a_a \) of unit a and the activation \( a_b \) of unit b. Hidden and output unit activations may range from -1 to +1, while input activations may range from 0 to 1. With unit types specified by 4 bits (the extra bits helping to protect unit-type genes against mutation and crossover), initial biases and weights by 4 bits, and learnability by one bit, the genotype length for a 3-unit network is \((4 \ast 3) + (1 \ast 3 + 3 \ast 3) \ast (4 + 1) = 72 \).
Grefenstette's (1987) standard genetic operators are used in the genetic algorithm: fitness-proportionate reproduction (with range-based fitness scaling), two-point crossover (p = 0.6 per offspring), and point mutation (p = 0.005 per genotype bit). Genes are ordered within genotypes by having all weights to a particular unit placed consecutively. Thus crossover tends to preserve associated groups of incoming connections to a given unit, but crossover points can occur anywhere within the genotype. Each generation contained 100 individuals for all runs reported here.

6 Some Poor Designs in Initial Populations

The problem facing our creatures is to eat food and avoid eating poison, when one sensory modality is noisy but evolutionarily stable, and the other sensory modality is accurate but evolutionarily variable. In this section we consider how well various poorly-adapted designs that might be randomly generated in our initial populations could do in this scenario. We expect these maladapted creatures to be quickly weeded out, but they are worth discussing because their defects illuminate important aspects of our watery world.

The simplest failure occurs in those creatures who do not even have any genetically specified motor unit (building block C) -- they would be unable to eat, so they'd always have zero fitness. Even worse, though, would be a creature with a smell unit (building block A), a motor unit (C), and a fixed connection between them (D) that has the wrong 'behavioral valence', i.e. that has the wrong fixed weight. For instance, if a creature's nervous system contains a smell unit connected by a fixed negative weight to the motor (eating) unit, it will tend to eat poison and avoid food -- even worse for its fitness than not eating at all.

Another design that has a somewhat more subtle flaw would be a network with a smell unit (A), a color unit (B), and a motor unit (C), but with only learnable connections (each with zero initial strength) between these units. If such a creature found itself in a green-food patch, for example, it could gradually build up a Hebbian association between green color and sweet smell, that is, between units A and B, because the activity on those two units would be highly correlated. But it would be unable to build up the appropriate positive or negative weights from either of these input units to the motor unit (i.e., the necessary connections D and E) would never be properly strengthened, because there would be nothing turning on and off the motor (eating) unit appropriately in response to food, and thus no association of sensation to food would be achievable. In other words, the creature would never be able to attach an adaptive behavioral valence to either color or smell. Such a creature would have found the statistical regularities in its world without being able to map them onto actual behavior. Thus, its average fitness would be zero.3

7 Evolving Better Networks

As the genetic operators process genotypes again and again through many generations, recombination and mutation will sometimes produce better designs. One reasonable design we can expect consists of a sweet smell unit (A) connected to a motor unit (C) by a fixed positive weight (D), with the other unit having only zero-strength incoming and outgoing connections and so being disregarded.4 A creature equipped with this nervous system will depend purely on smell to decide when to eat, with its behavioral accuracy constrained by the inherent noisiness of the smell sense in its world (i.e. by amount of turbulence). Although this design will sometimes make mistakes (ignoring food that smells sour because of the turbulence -- a 'miss' -- and eating poison that smells sweet -- a 'false alarm'), still on average it will eat more food than poison. Thus these creatures' energy, and number of offspring, will be higher than if they were just eating randomly. We call this design the color-blind eater, or, more fancifully, Chromignarus simplex, and it is shown in the top portion of Figure 1.

The best design though is an elaboration on the color-blind eater chassis. Creatures of this type gain an additional adaptive edge by including a red or green color unit (B), along with a learnable connection (E) to the eating unit. This creature design we name Chromodiscipulus simplex, color-learning eaters, as shown at the bottom of Figure 1.

Figure 1. The network designs of Chromignarus (without color learning) and Chromodiscipulus (with color learning), showing the corresponding building blocks (A-E).

With this design, an appropriately excitatory or inhibitory link (E) between the color unit (B) and the eating unit (C) will be built up over successive time-steps in the creature's life by the Hebbian associative weight-change

3 This raises a general problem for unsupervised learning algorithms that allow only the development of self-organizing perceptual maps or the extraction of statistical regularities in input vectors (e.g. Linsker, 1988, 1990). Although the resulting systems will in a sense 'know' the structure of the world they inhabit, they will have no way of translating that knowledge into appropriate behavior and thus into increased fitness.

4 The sweet-smell sensing unit can always be replaced by a sour-smell sensing unit, with the appropriate reversal of connection strength to the eating unit. Since mutation is unrestricted in our simulations, though, we do not tend to see the evolution of populations containing both subspecies simultaneously, because their interbreeding would tend to disrupt the linkage between smell-sense type and connection-valence.
mechanism. This will happen as follows: The sweet-smell unit (A) will, via its positive connection (D), preferentially turn on the eating unit (C) in the presence of food; the eating unit will have a bias so that it stays off otherwise (i.e. for poison). Now suppose the creature also has a red color-sensing unit (B), and food for this creature’s patch is red. Then the red unit and the eating unit typically will be firing in unison in the presence of food, causing the weight between them (E) to increase. Eventually, this weight will be large enough that the red unit alone can cause the eating unit to come on, regardless of what the sweet unit says to do — the creature has now learned that red means food. And since red, for this creature, always means food, the creature can rely solely on this visual cue, rather than the inaccurate smell cue, and always choose to eat properly, thereby increasing its fitness further. On the other hand, if food was green for this creature, the red unit and the eating unit would mostly come on separately, rather than in unison, and the weight between them would be decreased and eventually become inhibitory; the red color unit would thus serve to tell the creature when not to eat, and once again increase its fitness. So in either case, learning has given the creature a fitness bonus, and thereby proved adaptive.

8 Initial Simulation Results

We initially ran this simulation with a world where smell accuracy was 75%. The average population fitnesses of eight runs over 750 generations in this world are plotted in Figure 2. The populations tend to start with a collection of the sorts of poor designs discussed in section 5, with an average fitness of about zero. But very rapidly, the Chromignarus design evolves and spreads through the population, causing a leap in average fitness (to a value of about 2). All of the runs exhibit the rapid spread of this design’s strong fixed positive connection between a sweet-smell unit and an eating unit within about the first 100 generations. Chromignarus usually rules the turbulent seascape for many generations, resulting in the first population plateau (at average fitness value 2) seen along each curve in Figure 2. After this plateau is reached, evolution typically takes a while to add a color unit with a learnable connection to the eating unit onto the Chromignarus design.⁵ But once this addition occurs, and Chromodiscipulus appears, it rapidly takes over the population, increasing the average population fitness to a new plateau close to ceiling level. The switch to Chromodiscipulus happens quite quickly (over only a few tens of generations), as the step-like rise between the two plateau-levels on the curves in Figure 2 shows; but for each population, this switch occurs at different times.

![Figure 2. Average population fitness for several evolutionary runs with 75% smell-sense accuracy. Chromignarus predominates at the first fitness plateau, then Chromodiscipulus takes over at the second, with the evolution of color learning.](image)

The fact that the creatures evolved to use learning indicates that our world was properly constructed with a behavioral challenge for which learning would prove adaptive. As such, witnessing the evolution of learning in our simulations is a confirmation of our approach, but is not in itself surprising. We continued our investigation of the conditions under which learning could evolve with the more interesting question of how quickly learning would evolve, given various smell accuracies in different worlds. By tracking population average fitness values, it is possible to tell when the use of learning has spread through the population, as we showed in Figure 2. Initially, the average fitness quickly rises to a plateau where the fixed sweet-smell to eating unit connection is present in most of the creatures (the Chromignarus species). After remaining fairly level at this average fitness value for possibly many more generations, the average population fitness again jumps, indicating that the learnable connection from a color unit has penetrated the population (the Chromodiscipulus species). After this jump, the average fitness again levels out around its final highest value. Thus recording when fitness jumps occur can tell us when the different creature designs predominate in the population.

9 Analysis of the Adaptive Problem at Different Smell Accuracies

No matter what network design has evolved, adaptiveness in this scenario depends critically on smell accuracy. If the smell-sense is 50% accurate (chance level), so that food smells sweet half the time and sour the other half, then no useful information can be gained from any smell-sense.

---

⁵ Note that the particular color (red or green) sensed by this unit is not universal throughout the population, since as described in the previous section either color can be used equally well to guide the learned behavior.
units. Can the creatures still evolve to behave adaptively? No, because there would be no evolutionarily stable information for the system to build upon. The color-blind eater design with a fixed connection from smell to eating will have an average fitness of zero in such a world, since there’s no systematically helpful information for it to access. Any color-sensing creature designs with only a fixed connection from a color unit to the motor unit will also have zero average fitness across the two types of color patches in the world: in one patch, a given such design would do perfectly (always eating food and avoiding poison), but in the other kind of patch, it would be wired up exactly the wrong way and do terribly, always eating poison and avoiding food, so that the genotype’s average performance would be nil. Nor would a learnable connection from a color unit to an eating unit be of any help: there would be no way of bootstrapping the learning by using a statistically dependable association between smell and eating.

If smell accuracy is 100%, the color-blind eater design, Chromignarus, will take over the population and no design could possibly do better. With the association between smell and substance type both evolutionarily stable and completely accurate, nothing more is required for maximally adaptive behavior. Food will always smell sweet and poison always sour, so the creatures should evolve with just two utilized units, a sweet-sensing unit and an eating unit, connected by a strong positive weight (and both with zero or negative biases) -- exactly the Chromignarus design. This setup is sufficient to always make the right decision: eat when there’s a sweet smell, since it’s guaranteed to be food, and don’t eat when there’s a sour smell, from poison. No additional color information is needed.

But we now wondered what would happen for smell accuracies between 50% and 100% -- would learning evolve for all of these accuracies, and if so, how quickly? We proceeded to run many populations for accuracies in this range to explore these questions. To determine whether or not in fact learning evolved in each population, we used the average fitness plateau-monitoring method mentioned in the previous section.

The easiest way to discover when a population has evolved networks that learn color associations adaptively is to compare population average fitness at each generation with the average fitness that would be expected if the creatures did not learn to use color information. That is, if the creatures are on average doing better than creatures without color learning typically could do, then a large number of the creatures have most likely (barring large random effects) evolved to learn to use color. Computing expected fitness without color learning for different levels of smell accuracy is relatively easy.

In general, for a creature with a sweet-smell unit linked to an eating unit by a fixed positive connection, expected fitness without color learning is just

\[
E(F) = P(f)\cdot F_f \cdot P(s|f) + P(p)\cdot F_p \cdot P(s|p),
\]

where \(P(f)\) is the probability of food, \(F_f\) is the fitness benefit of eating food, \(P(s|f)\) is the probability of food smelling sweet (a sweet smell given the presence of food), that is, the smell accuracy, \(P(p)\) is the probability of poison, \(F_p\) is the fitness decrement from eating poison, and \(P(s|p)\) is the probability of poison smelling sweet. Since \(P(f) = P(p) = 0.5\) (food and poison appear equally often), \(F_f = +10\) points, \(F_p = -10\) points, and \(P(s|f) = (1 - P(s|f))\), expected fitness varies simply with smell accuracy. For example, if smell accuracy is 100%, then \(E(F) = (0.5)(+10)(1.0) + (0.5)(-10)(0.0) = 5\). For 90% accuracy, we get \(E(F) = (0.5)(+10)(0.9) + (0.5)(-10)(0.1) = 4\). Likewise, smell accuracies of 80%, 70%, 60%, and 50% result, respectively, in expected fitnesses of 3, 2, 1, and 0, without color learning.

Calculating expected fitness given color learning is more complex, depending not only on smell accuracy (which determines ease of learning), but on learning rate and number of life-steps. Given any inaccuracy in the smell-sense, creatures will start out performing relatively poorly, until the connection-strength from their color-sensing unit to the eating unit has been built up (to override the imperfect smell-sense). Thus their expected fitness even with color learning will generally be lower than the maximum achievable with perfect behavior, here 5 (the expected fitness computed above for 100% smell accuracy). Longer lifespans tend to raise the expected fitness ceiling by making the period of improper behavior before complete learning relatively shorter. Various learning rates tend to affect the expected fitness in a non-monotonic way, with a trade-off between noisy weight oscillation given high learning rates and slow weight build-up given low ones. (Adding ‘momentum’ to the Hebbian learning rule would tend to smooth the biased random walk build-up of weights, speed up learning, and result in higher expected fitness.) Thus, many factors combine in still-uncharacterized ways to determine the expected average population fitness once color learning has evolved; however, these considerations are unnecessary for our present goal of simply determining when learning has evolved in the population.

10 Parametric Exploration of the Learning Scenario

For each of 17 smell accuracies between 50% and 100%, we ran 20 populations of 100 individuals for 1500 generations each. Figure 3 shows how many generations it took each population to make each of the two jumps to new fitness plateaus. These two jumps correspond to the widespread appearance of Chromignarus (without color learning) -- indicated by asterisks -- and of Chromodiscipulis (with color learning) -- indicated by bullets. The bottom curve shows the average number of generations taken to evolve Chromignarus across the 20 runs at each accuracy level, and the top curve indicates the analogous average generations to evolve Chromodiscipulis, and thus learning itself.\(^6\)

\(^6\) 15 approximately evenly spaced accuracy levels from 52% to 92% inclusive.

\(^7\) If Chromodiscipulis had not evolved by the 1500-generation simulation limit for a particular run, we simply took that limit as the evolution time for color learning. Thus, for accuracies that tend to produce many ‘failed evolutions’, the average we plotted underestimates the actual evolution time.
75%, and took longer and longer for accuracies diverging on either side of that middle range (as shown by the upper, solid, curve in Figure 3). For example, it takes about three times longer for color learning to evolve with 60% or 90% smell accuracy, than with 75% accuracy. For both 50% accuracy (chance) and 100% accuracy, time to evolve color learning appears to asymptote to infinity (as suggested by the analysis in section 9). In further simulations, we have found that this U-shape is quite robust over a variety of population sizes and creature lifespans.

11 Theoretical Interpretation of the U-Shaped Learning Evolution Function

This U-shaped learning-evolution function may be a quite general pattern across certain adaptive learning scenarios. Such a relationship was not anticipated when we devised the learning scenario, so we have since endeavored to explore its significance both theoretically and empirically, by considering the factors thought to underlie it, and by manipulating those factors and observing the resulting evolutionary patterns. For expository convenience, the U-shaped relation between smell accuracy and time taken to evolve color learning will hereafter be referred to as the learning evolution curve, or LEC.

We view the U shape as emerging indirectly from a trade-off between the phylogenetic adaptive pressure to evolve learning (during species-wide evolution), and the ontogenetic ease of learning (during each individual creature lifetime). These two forces have roughly mirror-image effects on the time it will take to evolve learning as we move away from middle levels of smell-sense accuracy (75%), each being high (infinite in fact) at one end of the 50%-100% accuracy scale, and low at the other. They combine to produce the symmetrical U shape we see in Figure 3.

Consider first the phylogenetic adaptive pressures for learning. As outlined previously, Chromignarus performs poorly at low smell accuracies. Evolving color learning in such cases would add significantly to this creature’s fitness, by overcoming its error-prone, color-blind, smell-guided behavior. But the higher the smell-sense accuracy, the less color learning will be able to add to the average fitness, because the smell-sense alone will yield increasingly adaptive behavior. Thus, the phylogenetic adaptive pressure to evolve color learning peaks when smell accuracy is closest to chance (50%), and decreases as accuracy improves, ending up at zero when smell accuracy is perfect (100%).

Now consider the ontogenetic problems associated with actually doing color learning. Although Chromignarus in a low smell accuracy world would benefit greatly from color learning, that very inaccuracy makes learning difficult. The changes in the weight from the color unit to the motor unit are dependent on the correlation between the motor unit’s activity, controlled by the sweet-smell unit, and the activity of the particular color-sensing unit; the lower the smell accuracy, the smaller this correlation will be, and the smaller the correlation, the longer it takes for the weight changes to build up in the appropriate direction. In fact, at low accuracies, Chromodiscipulus’ finite lifespan may elapse before it has time to build up a weight from its color unit to its motor unit that is strong enough to begin to allow
color information to override the inaccurate smell information. But conversely, at high smell accuracies, the correlation between the eating activity and the color will be large, since the motor unit will come on (via its sweet-smell unit connection) almost exclusively in the presence of food, which is always one particular color. As a result, the weight changes on the learnable color connection will add up quickly in the proper direction. Thus, the ease of learning color-guided behavior peaks when smell accuracy is highest (100%), and becomes increasingly more difficult as accuracy reaches chance (50%).

All else being equal, when the adaptive pressure to evolve learning is high, the number of generations required to evolve learning should be low. Similarly, all else being equal, when the ease of learning is high, the number of generations required to evolve learning should be low. The mapping from each of these two factors in isolation to the time required for evolution thus could be a reciprocal function. If this is the case, then since the adaptive pressure to evolve color learning is zero when smell accuracy is perfect (100%), the time taken to evolve color learning (based on this first factor alone) should be infinite. At the other end of the accuracy scale, the adaptive pressure peaks, and so the amount this factor contributes to the time to evolve learning will be small. Conversely, since ease of learning is zero (i.e., learning is impossible) when smell accuracy is at chance (50%), the time taken to evolve color learning (based on this second factor alone) will also be infinite. And at 100% accuracy, where ease of learning peaks, the amount this factor contributes to learning evolution time will also be small. When the two factors combine, whether additively or multiplicatively, the factor indicating infinite time to evolve learning will win out. This creates the two arms of the U shape in the LEC which asymptote to infinity, at 50% and 100% accuracies. In the middle ranges of the smell accuracy, though, the two factors contribute relatively small times to the evolution of learning, and their combined effect is also relatively small. This gives the bottom of the U, at around 75% smell accuracy.8

To spell this out a bit further, at low smell accuracies, the difficulty of gaining benefit from learning offsets the (large potential) benefit itself, leading to increased time for learning to evolve. At high smell accuracies, in contrast, there is little adaptive pressure to evolve color learning, although color learning would be easy to perform. Natural selection cannot distinguish Chromignarius from Chromodiscipillus if they are both doing almost perfectly; this 'ceiling effect' will keep Chromodiscipillus from proliferating, so again color learning will take a long time to evolve.

But for middle smell accuracies, color learning is relatively adaptive and relatively easy. Color learning gives a significant fitness increase over using smell alone, and learning can occur fairly quickly, since the eating unit comes on rather more often to food than to poison. Mid-

8 Given our learning rate and lifespan, the two evolution-time factors happened to combine in such a way that they yielded a roughly symmetric LEC about 75% accuracy. The LEC could undoubtedly be skewed in either direction by changing the learning rate and/or lifespan, which would affect ease of learning, and hence the relative contributions of the two factors.

12 Planned Extensions and Future Research

The implications of the U-shaped learning evolution curve have yet to be fleshed out. Just how general it is across parameter variations, and how it changes with manipulation of learning rates, lifespans, genetic operators and representations, and informational structure in the environment all need to be explored. More importantly, though, we seek evidence of this type of relation in actual evolutionary processes in terrestrial environments. Such empirical verification of our results will be the true test of their significance.

In further work, our theoretical motivation will continue to be the exploration of adaptive agency and interactions among adaptive processes; our methodological strategy will continue to focus on the search for simple, elegant scenarios that reveal unanticipated, potentially general patterns and dynamics underlying adaptive agency. We intend to focus on the evolution of learning for a variety of adaptive purposes, beginning with issues of foraging and energy recruitment. This work will serve to provide further support for the emerging models of ecological learning theory. But we also hope that, through careful and thorough simulation of the processes embodied in adaptive agency, a general theory concerning the interaction of adaptive processes might emerge -- a concrete understanding of the interactions among adaptive pressures, cue structures in different environments, genetic representations and operators, developmental mechanisms, learning, information-processing, and behavior-generation.

Acknowledgements

This research was partially supported by National Science Foundation Graduate Fellowships to both authors. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We wish to thank Dave Rumelhart, Roger Shepard, the Evolutionary Psychology Group at the Center for Advanced Study in the Behavioral Sciences, and a population of other colleagues who have influenced and guided this research along the way.

References


