Synthetic Ethology: A New Tool for Investigating Animal Cognition (Extended Version)*

Technical Report UT-CS-01-462

Bruce J. MacLennan

Computer Science Department
University of Tennessee, Knoxville
maclennan@cs.utk.edu

May 31, 2001

Abstract

This report presents *synthetic ethology* as a new tool for the investigation of animal cognition. In synthetic ethology a population of simulated organisms is created inside a computer and allowed to evolve within a specified environment. Since we create the organisms and the world they inhabit, we are free to make them as simple or as complex as required for the investigation. The mechanisms underlying the organisms' behavior is fully explicit and accessible to the investigator; there can be no "ghost in the machine." Synthetic ethology permits investigations spanning a wide range of time and space scales, from simulated nervous system activity, to individual behavior, to group behavior and communication, up to populations. Thus cognitive capacities can be investigated in their full ethological context.

1 Goals

Synthetic ethology is based on several methodological commitments. First, it is based on the conviction that research into cognition should investigate behavior and the

*This report is an extended version of a chapter invited for *The cognitive animal: Empirical and theoretical perspectives on animal cognition*, ed. by Colin Allen, Marc Bekoff & Gordon Burghardt, MIT Press, in press.

mechanisms underlying that behavior in the agents' environment of evolutionary adaptiveness. Second, this investigation should extend over structural scales from the neurological mechanisms underlying behavior, through individual agents, to the behavior of populations, and over time scales from neurological processes, through agents' actions, to the evolutionary time scale. Obviously, such a wide range of scales is difficult to encompass in investigations of natural systems. Third is the observation that the discovery of deep scientific laws (especially quantitative ones) requires the sort of control of variables that can be achieved only in an artificial experimental setup.

Therefore we are faced with conflicting demands. On the one hand, we need precise experimental control. On the other, ecological validity dictates that agents be studied in their environment of evolutionary adaptiveness, where are there are innumerable variables, which are not amenable to independent control. *Synthetic ethology* intends to reconcile these conflicting requirements by constructing a synthetic world in which the phenomena of interest may be investigated. Because the world is synthetic, it can be much simpler than the natural world and thereby permit more careful experimental control. However, although the world is synthetic and simple, it is nevertheless complete in that the agents exist, live, and evolve in it.

The original motivation for synthetic ethology came from one of the central problems in cognitive science: the nature of *intentionality*, the property that makes mental states *about* something. We felt that an understanding of intentionality would have to encompass both the underlying mechanisms of intentional states and the social-evolutionary structures that lead to the creation of shared meaning. Our analysis of intentionality concluded that something is *intrinsically meaningful* to an agent when it is potentially relevant to an agent or to its group in its environment of evolutionary adaptedness (MacLennan 1992). Therefore intentionality must be studied in an evolutionary context.

We began our investigation with communication, since it involves both intentionality and shared meaning. We will show in this chapter how synthetic ethology permits the investigation of signals that are inherently meaningful to the signalers, as opposed to those to which we, as observers, attribute meaning.

2 Methods

The agents that populate our synthetic worlds can be modeled in many different ways; in particular there are a variety of ways of governing their behavior, including simulated neural networks and rule-based representations. In the experiments described here, an agent's behavior was controlled by a set of stimulus-response rules (64 rules, in these experiments). These rules were determined by an agent's (simulated) genetic string, but they could be modified by a simple learning mechanism (described below).

Since our goal is to investigate the synthetic agents in their environment of evolutionary adaptedness, they must evolve. Therefore our world includes a simplified form of simulated evolution, which proceeds as follows. Periodically two agents are chosen to breed, the probability of which is proportional to their "fitness" (as described later). The genetic strings of the two parents are mixed so that each of the offspring's genes comes form one or the other of the parents. In addition, there is a small probability of a gene being mutated. The resulting genetic string is used to create the stimulus-response rules for the single offspring, which is added to the population. In order to maintain a constant population size (100, in these experiments), one agent was chosen to "die" (i.e. to be removed from the population), the probability of dying being inversely related to "fitness."

We take this opportunity to illustrate the sort of experimental control permitted by synthetic ethology. Because we have complete control over the experimental setup and the course of evolution, we may begin with genetically identical populations and observe their evolution under different experimental conditions. If something interesting is observed in the course of an experiment, we may rerun the exact course of the evolution of the population to that point, and then make additional observations or experimental interventions to investigate the phenomena. Finally, whenever any interesting phenomena are observed, there can be no fundamental mystery, for all the mechanisms are transparent. If some agent exhibits interesting behavior, its entire mechanism is available for investigation. There can be no "ghost in the machine."

In synthetic ethology there is no requirement to model the natural world, so long as the synthetic world retains the essential characteristics of the natural world. That is, although determinate laws govern the evolution of our experimental populations, we are able to decide our world's "physical laws," which determine whether an agent "lives" or "dies," and which select agents to reproduce. The goal, of course, is to create synthetic worlds that are like the natural world in relevant ways, but are much simpler to study. The following experiment will illustrate what can be accomplished.

3 Demonstrating the Evolution of Communication

3.1 Methods

Our first series of experiments investigated whether it was even possible for genuine, meaningful communication to evolve in an artificial system. Therefore we decided to construct the simplest possible system that could be expected to lead to real communication.

Although there are many purposes for which an agent might be expected to communicate, we decided to focus on cooperation. Our reasoning was that communication could be expected to evolve in a context in which some agents have information that other agents could use to facilitate cooperation. Therefore we gave each agent a *local environment*, which could be sensed by that agent but by no other. It can be thought of as the situation in an animal's immediate vicinity, but to keep the model as simple as possible, we limited the local environments to be in a small number of discrete states (eight, in these experiments).

To make the state of one agent B's local environment relevant to another agent A, we arranged that they could cooperate only if A performed an action suitable for B's environment. To make this cooperation as simple as possible, we made our agents capable of producing an action from the same set as the local-environment states. Thus A could cooperate with B only by producing the same item as was in B's local environment, which A could not sense directly.

To select for cooperation we simply measured the number of times, over a specified period, that each agent was involved in successful cooperations. The probability of an agent reproducing was made proportional to this rate of cooperation, and its probability of dying was inversely related to the rate in a simple way. Thus we placed selective pressure on cooperation but not directly on communication; indeed, limited cooperation can be achieved by random action (which has a $1/8$ chance of succeeding).

Our experiments implemented only micro-evolution, so our agents were unable to evolve new sensor or effector organs. Therefore, we gave our agents organs that might be used for communication, but we did not construct the agents to use them in this or any other way.

Again, simplicity was our principle aim. Therefore we equipped our synthetic world with a simple *global environment*, shared by all the agents, which could be in one of a few discrete states (eight, in these experiments). The agents had the physical capability of sensing and modifying this global environment. Specifically, the state of the global environment is part of the stimulus to which an agent reacts, and the response can be either a new state for the global environment or an attempt to cooperate.

To test the potential effects of communication on cooperative behavior, we implemented a mechanism for making communication impossible. Specifically, when communication is being suppressed we periodically randomize the state of the global environment. This allowed us to measure the effect of apparent communication on the fitness (rate of cooperation) of the population, since genuine communication is defined in terms of its effect on the fitness of the communicators (Burghardt 1970).

We also investigated a very simple form of single-case learning, which could be enabled or disabled. When enabled, learning took place when an agent attempted to cooperate, but failed. Specifically, if agent A attempted cooperation L, but the last signaler's local-environment state was $L' \neq L$, then the rule of A that led to this action was changed to try L' instead. In other words, the rule used was changed to what would have been correct in this situation (although there is no guarantee that it would be the correct response in the future). This simple, single-case learning rule is potentially destabilizing, since it allows occasional errors to corrupt effective communicators, but it is a start towards investigating learning.

We initialized our population with 100 individuals with random genetic strings. Therefore, the stimulus-response rules governing their behavior, which were determined by their genomes, were also initially random.

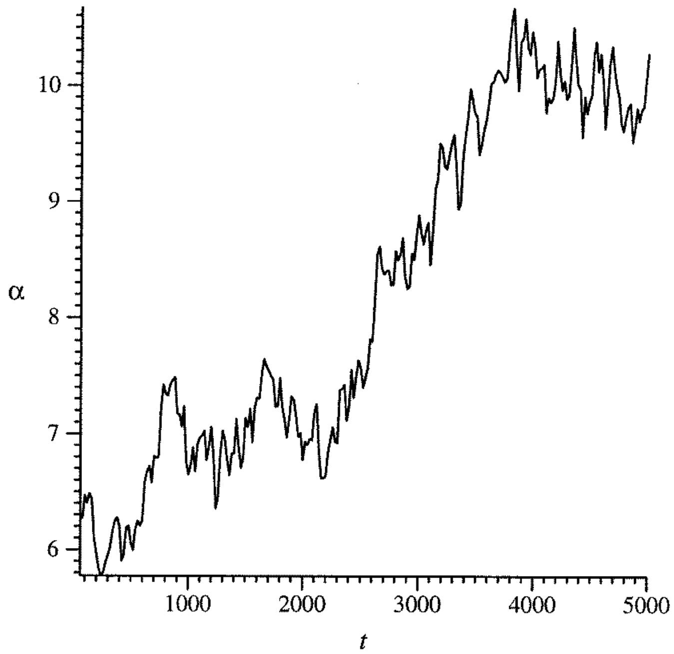

Figure 1: Degree of Coordination α: Communication Permitted with Learning Disabled

3.2 Results

To be able to measure the effect of communication on the fitness of a population, we quantified the fitness by the number of successful cooperations per unit time, which we call the *degree of coordination* of the population. (The unit of time is a "breeding cycle," in which one individual dies and one is born.) Because there is considerable random variation in the degree of coordination, the time series was smoothed by a moving average. Linear regression was used to establish the rate at which the degree of coordination (fitness) was increasing or decreasing. Details can be found in MacLennan (1990, 1992) and MacLennan & Burghardt (1993).

The baseline for comparison is determined by suppressing all possible communication, as previously described. In this case the degree of coordination stays near to 6.25 cooperations per unit time, the level predicted by analysis to occur in the absence of communication. Linear regression shows a slight upward trend in the degree of cooperation, which can be expressed as a rate of fitness increase: 3.67×10^{-5} cooperations per unit time per unit time. (The reason for this upward trend is discussed in the papers cited).

On the other hand, when communication was not suppressed, we found that the degree of coordination increased at a rate of 9.72×10^{-4} cooperations / unit time / unit time, which is 26 times faster than when communication was suppressed. Over an

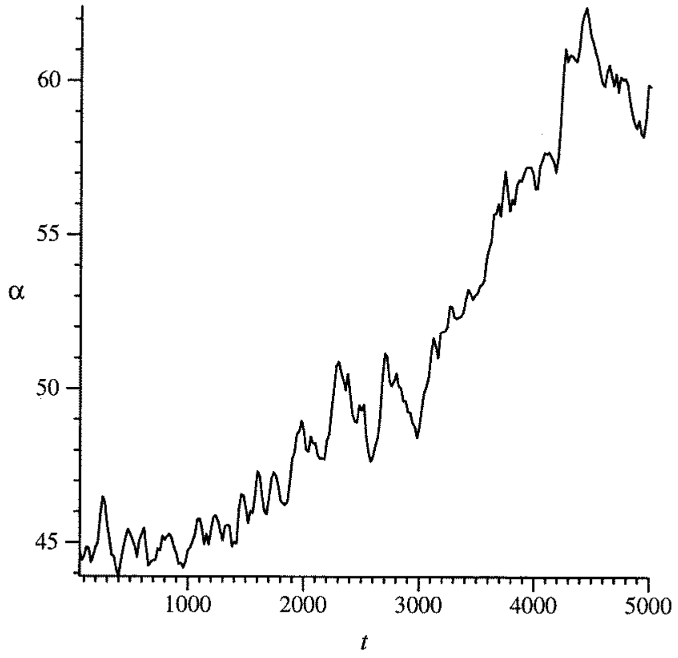

Figure 2: Degree of Coordination α: Communication Permitted and Learning Enabled

interval of 5000 breeding cycles, the degree of coordination reached 10.28 cooperations / unit time, which is 60% higher than the 6.25 achieved when communication was suppressed (Fig. 1).

When the agents were permitted to learn from their mistakes, fitness increased even faster: 3.71×10^{-3} cooperations / unit time / unit time, which is 3.82 times the rate when learning was disabled, and approximately 100 times the rate when communication was suppressed. Furthermore, the degree of cooperation begins higher than in the other cases, because the population is given several opportunities to respond to a particular situation before the local environments are re-randomized. Therefore, an agent has the opportunity to learn from its mistake and to respond correctly several more times before the local environments are changed. Thus we observe the degree of coordination to begin at approximately 45 cooperations / unit time (as opposed to the 6.25 without learning), and to climb rapidly to 59.84 cooperations / unit time, which is 857% above the level achieved without communication (Fig. 2).

As would be expected for experiments of this kind, there is considerable experimental variation from run to run. Nevertheless, the results we have described are typical over more than one hundred experiments. Therefore, we can conclude that genuine, meaningful communication is taking place, for it is enhancing significantly the fitness of the population. Furthermore, since communication evolves in our population when it is not suppressed, we may investigate genuine communication in its environment of evolutionary adaptedness.

Since it is genuine communication, the signals passed among the agents are meaningful *to them*, but not necessarily to us as observers. That is, we have the opposite situation from artificial intelligence, in which the computer manipulates symbols that are meaningful to us but meaningless to it (or, more precisely, have only derived meaning dependent on the meaning we attribute). Here we are in the same situation as in natural ethology: we are faced with apparently meaningful communication and must discover its meaning *for the communicators*.

Even in these simple experiments, signals and their interpretation are complex functions of the total situation. The signal emitted by an agent may depend on both its local environment and the shared global environment. Further, an agent's interpretation (use) of a signal may (and typically is) influenced by its own local environment.

Nevertheless, we would expect that over time a simple meaning would emerge for the signals; that is, that there would be a one-to-one correspondence between signals and local-environment states. To determine if this was occurring, we compiled a *co-occurrence table*, which recorded the number of times particular pairings of signal (global-environment state) with meaning (local-environment state) occurred in successful cooperations. So that we could track changes over time, the table at any given point of time reflected only recent activity by the agents.

If no communication were taking place, one would expect all signal/meaning combinations to be about equally likely, and that is what we found when communication was suppressed, and at the beginning of the simulations when it was not. However, when communication was not suppressed, the co-occurrence tables became more structured as the "language" self-organized.

We quantified the organization of the co-occurrence tables in a number of different ways, including entropy, a measure of disorder (so lower numbers represent greater organization). With our experimental design, the maximum entropy, when all signal/meaning pairs are equally likely, is 6 bits, but when there is a one-to-one symbol/meaning correspondence, the entropy is 3 bits. When communication was suppressed we observed an entropy of 4.95 bits, which shows that it is not completely disordered, but when communication was not suppressed, the entropy decreased (after 5000 breeding cycles) to 3.87, representing a much higher degree of organization.

Visual inspection of the evolved co-occurrence tables showed a number of cases in which, almost always, a particular signal corresponds to a particular meaning and vice versa. However, we also observe cases of ambiguity, in which a signal is more or less equally likely to correspond to two or more meanings, and cases of synonymy, in which two or more signals are about equally likely to correspond to a particular meaning. These cases could result from individual agents using ambiguous or synonymous symbols, or from two or more competing "dialects" in the population, but Noble & Cliff (1996) have evidence supporting the former hypothesis.

The observations heretofore described can be called behavioral and are analogous to those made in natural ethology. However, synthetic ethology affords additional possibilities, for the structure of the agents is completely transparent. At any time we may "dissect" the agents and analyze their behavioral programs (see MacLennan 1990 for examples). Thus we may relate the mechanisms of behavior to their manifestations in the population.

3.3 Brief Overview of Other Experiments

We have been interested in whether a population would evolve to use sequences of symbols for communication, were there selective advantage in doing so. Therefore my students and I have conducted a number of experiments, similar to those already described; details may be found in MacLennan (in press) and in the references cited therein. In these experiments the agents evolved the ability to communicate with pairs of symbols displaying a very rudimentary "syntax," but the results have been less than we expected. One explanation may be that the very simple behavioral model we used (finite-state machines) is too weak for the sequential perception and control required for more complex communication. Animals, in contrast, have rich, highly structured perceptual-motor systems, which evolution can recruit for communication. Therefore, future experiments might need to use more complex agent models, as well as a more structured environment about which they might communicate.

We have conducted some experiments using neural networks as behavioral models, but they have not produced substantially different results. Most likely this is because the nets that we have used are about as unstructured as the finite-state machines.

4 Discussion

Of necessity, our discussion of related and future work and of implications must be brief. Noble & Cliff (1996) have replicated our earliest studies and extended them in a number of informative ways. A somewhat different approach can be found in Werner & Dyer (1992), which demonstrated the evolution of communication by making it necessary for effective reproduction. Steels (1997a, 1997b) has conducted fundamental studies of the emergence of meaningful symbols.

In discussing related work, it may be worthwhile to make a few remarks about the connection between synthetic ethology and two related disciplines, *artificial life* and *artificial language*. First, it must be stressed that there is substantial overlap between the three, so that the difference is at most one of emphasis.

Artificial life studies artificial systems that are significantly "lively" in some sense. Some investigators are attempting to create systems that are literally alive, while others are content with systems that faithfully imitate life. In ether case, the artificial systems may exist as robots or as patterns of electrical activity in a computer's memory. Synthetic ethology differs in that the agents need not be alive in either of these senses, although they may be. Certainly, we make no claim that the agents described in this report are alive in any literal sense.

Artificial language research, the newest discipline of the three, uses computers to study the formal processes governing the formation and evolution of languages (e.g. Kirby 2000a, 2000b). It tends to concentrate on the sorts of complex syntactic and semantic structures found in human languages, and tends to treat the languages as autonomous systems independent of the behavior and evolution of the agents that use them. Although one of the original goals of synthetic ethology was to study the evolution of human-like languages, to date it has been limited to very simple communication systems. Another difference is that synthetic ethology may be applied to other kinds of behavior besides communication. Future work, however, might combine synthetic ethology and artificial language techniques.

Current experiments in synthetic ethology are too simple to exhibit psychological states, but future ones may be; we do claim that even the current experiments do exhibit genuine intentionality. Nevertheless, synthetic ethology indicates how psychological states may be made accessible to scientific investigation.

We have claimed that our agents (although they are not conscious, nor even alive) exhibit genuine intentionality. The point is certainly arguable and depends on our analysis of intentionality. Nevertheless, all subtleties aside, we claim that the signals are inherently meaningful to the agents because the agents' continued persistence as organized systems depends on their use of the signals.

Are these synthetic worlds and agents so alien that results will not be seen as relevant to nature? In particular, we have argued that we can use abstract, ad hoc selection rules (since the "laws of physics" are under our control), but it can be objected that selection should be more naturalistic (e.g. Werner & Dyer 1992). Certainly, this is an important issue, and in the long run we want to explore ever richer synthetic worlds, but to introduce gratuitous complexity would defeat the goals of synthetic ethology.

One of the advantages of synthetic ethology is that we can make our worlds as simple as possible, so long as they include the phenomena of interest. On the other hand, we must construct these worlds from scratch; they are not given to us. This becomes a challenge as we begin to investigate phenomena that require larger populations of more complex agents acting in more complex environments. Simulating such worlds requires ever more powerful computers. Therefore synthetic ethologists must strike a delicate balance between the sophistication of the synthetic world and the resources required to implement it. Indeed, as we move in the direction of greater complexity, synthetic ethology will face some of the same problems as natural ethology. Nevertheless, by affording greater control and an alternative to natural life, it will remain a worthwhile approach.

5 Bibliography

1. Burghardt, G. M. (1970). Defining ‘communication’. In J. W. Johnston Jr., D. G. Moulton & A. Turk (Eds.), *Communication by chemical symbols*, pp. 5–18. New York (NY): Appleton-Century-Crofts.

2. Kirby, S. (2000a). Learning, bottlenecks and the evolution of recursive syntax. In E. J. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models.* Cambridge (UK): Cambridge University Press.

3. Kirby, S. (2000b). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert-Kennedy & J. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic forms.* Cambridge (UK): Cambridge University Press.

4. MacLennan, B. J. (1990). *Evolution of communication in a population of simple machines* (Tech. Rep. No. CS-90-99). Knoxville, TN: Computer Science Department, University of Tennessee.

5. MacLennan, B. J. (1992). Synthetic ethology: An approach to the study of communication. In C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen (Eds.), *Artificial life II: The second workshop on the synthesis and simulation of living systems*, pp. 631–658. Redwood City (CA): Addison-Wesley.

6. MacLennan, B. J. (in press). The emergence of communication through synthetic evolution. In V. Honavar, M. Patel & K. Balakrishnan (Eds.), *Advances in evolutionary synthesis of neural systems.* Cambridge (MA): MIT Press.

7. MacLennan, B. J., & Burghardt, G. M. (1993). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, **2**, 161–188.

8. Noble, J., & Cliff, D. (1996). On simulating the evolution of communication. In P. Maes, M. Mataric, J.-A. Meyer, J. Pollack & S. W. Wilson (Eds.), *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pp. 608–617. Cambridge (MA): MIT Press.

9. Steels, L. (1997a). Constructing and sharing perceptual distinctions. In M. van Someran & G. Widmer (Eds.), *Proceedings of the European Conference on Machine Learning.* Berlin: Springer-Verlag.

10. Steels, L. (1997b). The synthetic modeling of language origins. *Evolution of Communication*, **1**, 1–34.

11. Werner, G. M., & Dyer, M. G. (1992). Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen (Eds.), *Artificial life II: The second workshop on the synthesis and simulation of living systems*, pp. 659–687. Redwood City (CA): Addison-Wesley.