

Acquisition of General Adaptive Features by Evolution.

Dan Ashlock¹ and John E. Mayfield²

¹ Iowa State University, Department of Mathematics
Ames, Iowa, 50010
danwell@iastate.edu

² Iowa State University, Department of Zoology and Genetics
Ames, Iowa, 50010
jemayf@iastate.edu

Abstract. We investigate the following question. Do populations of evolving agents adapt only to their recent environment or do general adaptive features appear over time? We find statistically significant appearance of general adaptive features in a spatially distributed population of prisoner’s dilemma playing agents in a noisy environment. Multiple populations are evolved in an evolutionary algorithm structured as a cellular automaton with states drawn from a rich set of prisoner’s dilemma strategies. Populations are sampled early and at the end of a ten-thousand generation simulation. Modern and archaic populations are then placed in competition. We test the hypothesis that competition between an archaic and modern population yields probability $p = 0.5$ of modern populations out-competing archaic ones. The hypothesis is rejected at a confidence level of 99.5% using a binomial probability model in each of seven variations of our basic experiment.

1 A Question about Evolution

The modern evolutionary synthesis successfully combines the principles of genetics with Darwin’s concept of natural selection acting on variation in a population. The current synthesis very successfully explains the action of natural selection generation by generation; but is not a complete theory. For example, the modern synthesis does not provide insight into the long standing question of whether or not “evolutionary progress” is a legitimate concept. In this paper we report seven variations on an experiment that clearly documents one kind of evolutionary progress, the production of a generalized ability to compete at the iterated prisoner’s dilemma.

We simulate evolution on a spatial grid, called the *world*. Natural selection is simulated by allowing a player to colonize adjacent grid cells if it obtains a higher total score in play of noisy iterated prisoner’s dilemma than its neighbor did. A source of variation is provided by occasional random mutations of players on the grid. The state of the world is sampled early in the course of evolution and after substantial time has passed. Modern and archaic world states, with and without a historical relationship, are then brought together and allowed to compete in the

absence of mutation for world domination. This yields a Bernoulli variable with success defined to be world domination by a modern population. A probability for this Bernoulli variable close to (or below) 0.5 would indicate no acquisition of general adaptive features. A probability significantly above 0.5 indicates that general adaptive features are appearing in our simulation of evolution.

2 Model Specifications

2.1 Overview

We adopt a modification of a model of Lindgren and Nordhall [6] [8], [7], a *fashion-based* cellular automata (terminology due to J. Lutz), with states drawn from a set of finite state machines (FSMs) specialized to play the iterated prisoner’s dilemma. This section presents only a quick, global, picture; details are given subsequently. A cellular automaton (C, S, \mathcal{N}, μ) is a discrete dynamical system, comprised of four parts, a set C of cells, a set of S states that cells may take on, an assignment \mathcal{N} of neighborhoods to cells, and an updating rule μ mapping the current states of cells in the neighborhood of a cell into a new state for the cell. Possibly the most famous example of a cellular automaton is Conway’s game of life. A good popular introduction to cellular automata appears in Stephen Wolfram’s article in *Scientific American* [12]. An early collection of papers on cellular automata, is provided by Wolfram et al.[3].

The cellular automata we use in this research have a 100-by-100 rectilinear grid as its set, C , of cells. The grid wraps at its edges creating a toroidal spatial structure without edges. The neighborhood of a cell comprises the cell itself together with those above, below, right, and left of it in the grid. The set, S , of states is the set of FSMs with at most 255 states using the Mealy machine architecture [11]. The updating rule, μ , has two parts, *fashion* and *mutation*, as follows. The FSM in each cell plays noisy iterated prisoners dilemma against its four neighbors and possibly against a fixed FSM called the *environment*. This generates a score for each cell. Each cell then adopts the state of the neighbor that has the highest score (following the “fashion” of the most successful neighbor). Recall that each cell is its own neighbor, so the state of a cell will not change if the current occupant out-scores its neighbors. After the fashion updating has taken place, a small number of FSMs in randomly selected locations are mutated, i.e., small random changes are made to their data structures.

2.2 The Prisoner’s Dilemma

Prisoner’s dilemma is a two player simultaneous game. The game is explored in detail with many excellent examples in *The Evolution of Cooperation* by Robert Axelrod [2]. Briefly, two players simultaneously decide to defect or cooperate. If both defect they receive a score of D . If both cooperate both receive a score of C . If one defects and the other cooperates then the defector receives a score of

H and the cooperator receives a score of L . Prisoner’s dilemma requires that the payoffs obey the following pair of constraints.

$$L \leq D \leq C \leq H \tag{1}$$

$$L + H \leq 2C \tag{2}$$

Equation 1 simply requires that the payoffs come in the intuitive order. Equation 2 forces sustained cooperation to average higher than alternate cooperation and defection. We use the common values $L = 0$, $D = 1$, $C = 3$, and $H = 5$.

In the [single shot] prisoner’s dilemma, strategy is simple. No matter what the other player does, the best score comes from defecting. If two players play repeatedly, the strategy becomes more complex. Repeated play is termed the *iterated prisoner’s dilemma* (IPD).

In this research we use a noisy version of IPD. There is a small chance $\alpha = 0.01$ that any action during the course of play will be replaced with its opposite. Adding noise substantially changes the balance of power among strategies, see [6], and serves to add some biological character to the simulation as we shall see later. In particular, the strategy tit-for-tat, which figures prominently in Axelrod’s analysis, is transformed by noise from an excellent strategy to one that is not asymptotically better than flipping coins to pick an action.

2.3 The Finite State Machines

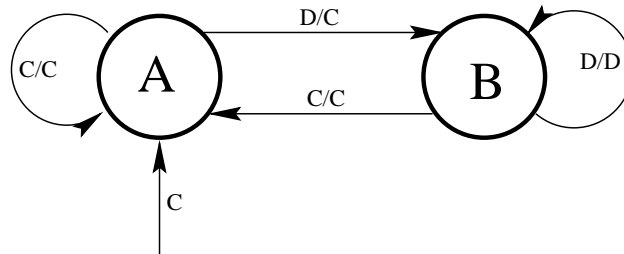


Fig. 1. A two state FSM, using the Mealy machine architecture that implements the Prisoner’s Dilemma strategy tit-for-two-tats. The arrow entering from outside the diagram gives the initial state of the machine, state A, and is labeled with the machines initial action, (C)operation. Subsequent actions and states are found by following the arrow out of the current state that is labeled with (opponent’s last action)/(the machines current action).

The states of the cellular automata used in this research are FSMs specialized to play the iterated prisoner’s dilemma. A seminal paper in this area is [9], other papers that explain the use of FSMs of the type used in this research are [11], [1], [4]. Caution should be exercised in regard to the word “state”; the FSMs

that are the states of the cellular automata themselves have internal states. An example of one of these FSMs, implementing the strategy tit-for-two-tats, is shown in Figure 1. The FSMs used in this research start with eight states and with uniform, randomly assigned initial states, transitions, and responses. The number of states in an FSM can be changed, by mutation, to any number in the range 1 to 255.

2.4 Updating the Cellular Automata

In order to update the states of our cellular automata, as per the fashion rule, we need a score for each cell. There are two sorts of play that yield values added into the score: neighbor play and environment play. In all but one of the experiments described in this paper there is a single, fixed strategy called the *environment* that each player plays against for n_e rounds of IPD. This environmental strategy does not change or adapt in any way, but rather gives the FSMs in the simulation a fixed environmental feature to adapt to. As a control, one set of simulations is performed against the “null environment” in which no environmental play takes place. In addition to the environment, each finite state machine plays each adjoining cells for n_n rounds of the IPD. The environment and neighbor scores are then added to get the total score used in fashion updating. We take $n_e = n_n = 60$.

The number of states (strategies for playing IPD) present in the cellular automata during any time step of the simulation is much smaller than the total number of cells. Initially we generate sixteen random FSMs and use these FSMs as the initial state values available to the cells of the cellular automata. In the course of a simulation, mutation produces an average of less than one new potential state (type of FSM) per time step. It is also the case that many types of FSMs die out. This allows us to save computer time. Rather than play all 1,200,000 pairwise games of IPD needed to get the scores for the fashion updating, we instead do the following. We compute a score matrix indexed by the FSMs (cellular automata states) actually appearing in cells in the current time step. The entry $a_{i,j}$ of the matrix is the average score obtained by FSM i in five sessions of play, each consisting of n_n rounds of IPD, against FSM j . We also compute the average score in the same manner for each type of machine present against the environment. Since an early choice during play of the IPD can place an FSM irrevocably into one or another part of its internal state space, this average is more representative of the behavior of the strategy coded than a single set of plays. Representative behavior is required as the score obtained is used for all members of a potentially large population of FSMs of a given type.

The score one FSM obtains is then computed by looking up its scores against its neighbor *types* in the score matrix and then adding these to the environmental score to get a total score for a given cell of the world. The score matrix is recomputed at the beginning of each time step. With the score of each cell computed, the state of the cells is updated by the fashion mechanism.

2.5 Mutation

After the states of the cellular automata have been updated in the fashion phase we introduce mutants at a rate of four per hundred thousand to the population. This gives an expectation of 0.4 mutants per time step. We use three types of mutation. The first type of mutation, termed *basic mutation*, randomly changes the initial action, a transition, or a response of an FSM to a new value, selected uniformly at random. The second type of mutation, termed *growth*, adds a state to an FSM (unless it has 255 states already). An added state is created by duplicating an existing state. Transitions from the duplicated state, and their associated responses, are copied exactly. Transitions to the duplicated state are assigned randomly to the original state or its duplicate. This has the effect of adding a state to the FSM without changing its behavior during play of the IPD but the effects of subsequent mutations are modified. The growth mutation is inspired by the gene doubling mutation in [6]. The third type of mutation, termed *contraction*, deletes a state from the FSM. Any transitions to the state deleted are randomly reassigned to remaining states of the automata. During updating of the cellular automata, those mutations performed are 50% basic, 25% growth, and 25% contraction.

3 Experiments Performed

A single simulation consists of generating 16 initial random FSMs, generating an initial configuration for the cellular automaton by choosing the state of each cell uniformly from the sixteen available FSMs, and then updating the automata as described in Section 2.4 10,000 times. The entire contents of the world are saved at time-steps 100 and 10,000. We term the worlds saved at time step 100 *archaic* and those saved at time step 10,000 *modern*. We performed 210 simulations, 30 for each of seven environmental strategies. The environmental strategies used included the null strategy, all four strategies that can be implemented with a single state FSM, an extremely cooperative strategy, and an error correcting strategy. A brief description of these environmental strategies follows.

- **Null.** The null strategy is the absence of a strategy. Environmental play does not take place.
- **Always Cooperate.** This strategy cooperates no matter what. This gives the FSMs in the simulation a reason to defect a lot, but they will still also benefit from recognizing and cooperating with one another.
- **Always Defect.** This strategy defects no matter what. It should force the FSMs in the simulation to defend themselves against defection.
- **Tit-for-Tat.** This is the classic strategy from Axelrod’s analysis of the iterated prisoner’s dilemma. It cooperates initially and then repeats its opponents move thereafter.
- **Anti-Tit-for-Tat.** This strategy is the reverse of tit-for-tat. It defects initially and returns the opposite of its opponents last move thereafter. This

is an insane strategy for a member of an evolving population but is a challenging environment as it is both exploitable and impossible to cooperate with.

- **Tit-for-Two-Tats.** Tit-for-two-tats is a nicer version of tit-for-tat. An FSM that plays tit-for-two-tats is shown in Figure 1. It cooperates initially and in any turn in which its opponent has not defected twice in a row previously. This behavior encourages an opponent to defect and cooperate against it alternately, reaping an average score of $(H + C)/2 = 4$.
- **Pavlov.** Pavlov is an error correcting strategy. It was reported (but not named) in [6]. Pavlov cooperates initially and cooperates in subsequent turns when it and its opponent did the same thing in the last play. This has the effect of correcting errors caused by noise when two copies of Pavlov are playing one another. Before a noise event, two copies of Pavlov engage in mutual cooperation. If an action is inverted by noise then one copy cooperates and the other defects. As both copies have made a different move, they both defect on the next time step. This paired defection sets the stage for cooperation thereafter, at least until the next noise event.

For each simulation, in addition to saving the state of the automata in time-step 100 and 10,000 a number of summary statistics are taken every 10th time-step. The mean and standard deviation of the average prisoner’s dilemma score made by the FSMs in the population are saved. The mean and standard deviation of the number of states in an FSM and the mean number of states that can be reached from the initial state of an FSM are also saved. This latter number, given noise, is a good approximation of the number of states the FSM uses.

Once we have run the 30 simulations for a given environment we then compare archaic and modern populations. To make the comparison we load the left half of an archaic world and the right half of a modern world (this choice, while arbitrary, is unbiased). We turn off mutation entirely and proceed via fashion updating for 100 time-steps. Whichever population constitutes a majority of the cells in the 100th time step is declared to have out-competed the other. We designate success as a modern population out-competing an archaic one to obtain a Bernoulli trial. All 30 modern populations are compared to all thirty archaic ones, generating 900 such trials. The number of time-steps used in the comparison, 100, is twice the time needed for a superior strategy to overwhelm and inferior one. This follows from the fact that a superior strategy can spread one cell per time-step and the region occupied by an archaic or modern population is 50 cells across in its smallest dimension. A graphical representation of four of the cellular automata produced by simulation is shown in Figure 2.

4 Experimental Results

For all seven environments tested we reject at the 99.5% confidence level the hypothesis that the probability of a modern population out-competing an archaic population is 0.5. The confidence intervals uniformly estimate this probability

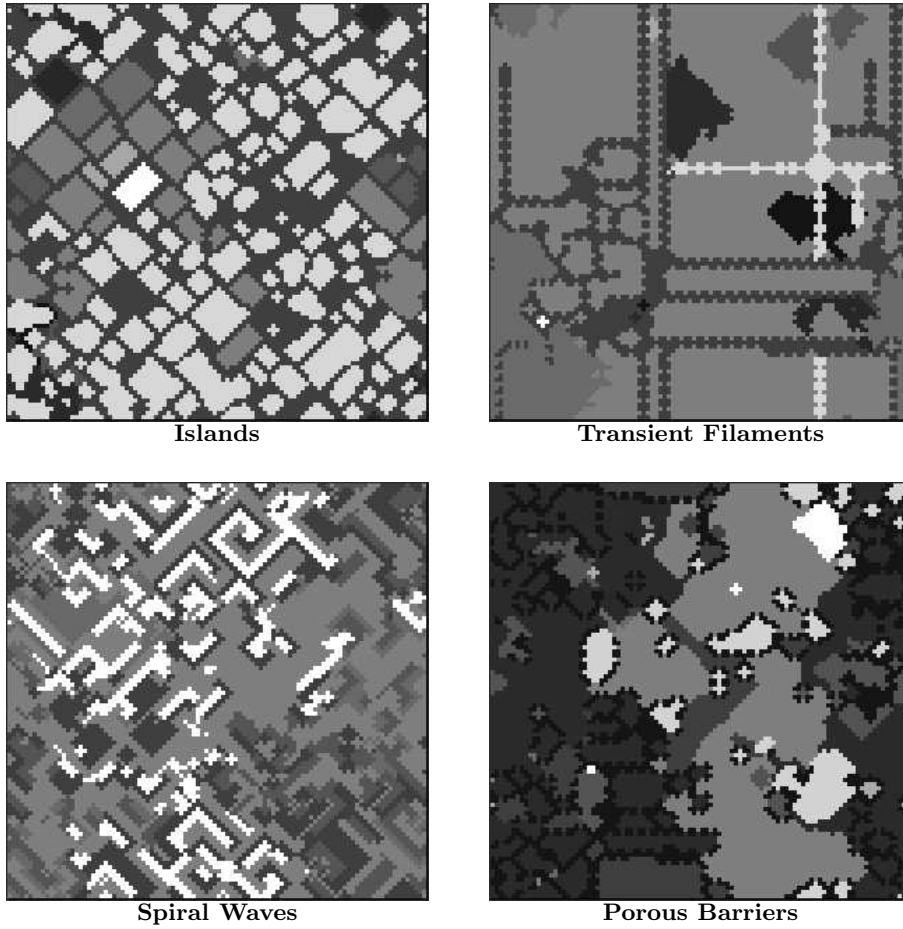


Fig. 2. Shown here are the cellular automata from update 10,000 of four different simulations. The upper-left picture, from null run 24, shows a filamentous network that isolates islands. The upper-right picture, from run null 16, shows transient filament builders. The lower-left picture, from always defect run 12, exhibits a spiral waves composed of three distinct FSM types. The lower-right picture, from tit-for-two-tats run 26, shows curious porous barriers, something observed only three times in the worlds saves at update 10,000.

to be significantly above 0.5. A summary of the number of successes, empirical probabilities, and resulting confidence intervals is given in Figure 3. The statistical test used a normal approximation to the binomial distribution [5]. It would appear that some sort of general ability to play the iterated prisoners dilemma is conferred over time by the digital evolution performed in these experiments. In addition to being quite distinct from 0.5 the probabilities approximated by our experiments in the different environments are sometimes significantly different from one another, having disjoint confidence intervals.

Environment	Count	p	99.5% confidence for p
Null	697	0.7744	$0.7353 \leq p \leq 0.8135$
Always Defect	692	0.7689	$0.7294 \leq p \leq 0.8084$
Always Cooperate	809	0.8989	$0.8707 \leq p \leq 0.9271$
Tit-for-tat	809	0.8989	$0.8707 \leq p \leq 0.9271$
Anti-tit-for-tat	756	0.8400	$0.8057 \leq p \leq 0.8743$
Tit-for-two-tats	628	0.6978	$0.6548 \leq p \leq 0.7408$
Pavlov	763	0.8477	$0.8141 \leq p \leq 0.8813$

Fig. 3. Results of Bernoulli trials comparing modern and archaic populations.

There is a marked lack of pattern in the probabilities relative to the type of environment that produced them. The “nice” strategies, tit-for-two-tats, Pavlov, always cooperate, and tit-for-tat comprise both ends of the range of probabilities. The error tolerant strategies, Pavlov and to a lesser extent tit-for-two-tats are at opposite ends of the distribution. The easily exploited strategies, anti-tit-for-tat, always cooperate, and tit-for-two tats are firmly at both ends of the spectrum. The difficult to exploit strategies, always-defect and tit-for-tat, are widely separated. It is not clear that the attributes niceness, error correction, and exploitability, are relevant for predicting the effects of an environmental strategy - though it is intuitive that exploitability, at least, should be. At present we simply conjecture that we have not run experiments in enough strategies to see any pattern in the effect of the environment. Interestingly, the lack of an environment is most like pure defection in its effect on the evolving populations. Given only their fellow automata as an “environment” the population acts as though the environment is uniformly hostile.

5 Discussion

The major stumbling block in any attempt to detect evolutionary progress is identifying an objective computable metric of such progress. In this paper we have studied the ability of a large number of sample populations to compete with one another. Relative competitive ability, measured in terms of differential reproduction under the effects of natural selection, is an objective measure of evolutionary progress. Such experiments can only be done with living organisms if an uncontaminated survival from an previous epoch is found or when the time scale of adaption is quite short, for example as in [10].

The experiments reported here consisted of varying the amount of time populations have had to evolve while controlling all other factors. The assessment of relative competitive ability of modern and archaic populations yielded a strong positive response to the question “was there evolutionary progress?” Collateral data saved on the FSMs over the course of evolution indicates that the mechanism of progress may well have been in learning how to deal with noise. Given

that there are often large numbers of identical FSMs in our simulations, being able to cooperate with a copy of yourself is a selective advantage. Given the need to defend yourself against either harsh FSMs or the environment you must also have the capacity to respond to defection. In a noisy environment this can lead to retaliation against noise events. This in turn places a premium on the ability to recover from error. An example of such an FSM is Pavlov, described in the list of environmental strategies in Section 3. Any error recovery path can, however, be exploited by an opponent that adapts to it. Notice that always defect, for example, can exploit Pavlov's (very short) error recovery path. This situation creates the potential for an arms race of ever more subtle ways to recover from errors made while playing nominally cooperative opponents while shutting doors on rapacious opponents that are exploiting existing error recovery paths.

A remark is needed on the choice of the times 100 and 10,000 for population sampling. The number 10,000 is simply "large." Typically many displacements of the entire set of FSMs occur in 10,000 updatings. The choice of 100 timesteps requires more justification. In earlier experiments [9], [1], [11] it was found that a population had lost its initial random character and diversity roughly by timestep 35-60. This was the reason for initially choosing 100 updatings as the time of emergence of a primitive population from its random beginnings. In the cellular automata used, a populations loose much of their initial diversity by timestep 50, settling on one or a few FSMs. By update 100 several mutants typically appear and either die or spread. This indicates that the population at update 100 of the automaton is not untested by selection. Finally, another set of 30 experiments were run in the null environment and saved at update 2000 as well as 100 and 10,000. Progress, of the sort reported in Section 4 was seen in comparing generation 100 to 2,000 and 2,000 to 10,000. These results will be reported more fully in a subsequent paper.

6 Future Work

This paper represents an initial foray into the use of evolutionary algorithms to test the notion of evolutionary progress. A large number of *ad hoc* choices of parameters were made implying the need for sensitivity studies. Of more interest is the repetition of the experiment for other open-ended tasks besides the iterated prisoner's dilemma. It may be possible, by taking samples more often, to get a notion of the rate of progress. A predictive characterization of the effect of the environmental strategy on the rate of evolutionary progress would be a desirable result but it is not clear that such a characterization is an obtainable result. Finally, we have put forth one intuitive measure of evolutionary progress. Other possible measures of evolutionary progress should be explored.

References

1. Dan Ashlock, Mark D. Smucker, E. Ann Stanley, and Leigh Tesfatsion. Preferential partner selection in an evolutionary study of prisoner's dilemma. *Biosystems*, 37:99-125, 1996.

2. Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.
3. Stephen Wolfram Doyne Farmer, Tommaso Toffoli. *Cellular Automata*. North Holland, New York, 1984.
4. D.B. Fogel. Evolving behaviors in the iterated prisoners dilemma. *Evolutionary Computation*, 1(1):77–97, 1993.
5. Richard J. Larsen and Morris L. Marx, editors. *An introduction to mathematical statistics and its applications*. Prentice-Hall Inc., 1981.
6. Kristian Lindgren. Evolutionary phenomena in simple dynamics. In D. Farmer, C. Langton, S. Rasmussen, and C. Taylor, editors, *Artificial Life II*, pages 1–18. Addison-Wesley, 1991.
7. Kristian Lindgren and Mats G. Nordahl. Artificial food webs. In Christopher G. Langton, editor, *Artificial Life III*, pages 73–103. Addison-Wesley, 1994. SFI Studies in the Sciences of Complexity, Proc. Vol. XVII.
8. Kristian Lindgren and Mats G. Nordahl. Evolutionary dynamics of spatial games. *Physica D*, 75:292–309, 1994. To appear.
9. John H. Miller. The coevolution of automata in the repeated prisoner’s dilemma. A Working Paper from the SFI Economics Research Program 89-003, Santa Fe Institute and Carnegie-Mellon University, Santa Fe, NM, July 1989.
10. Virginia Morell. Flies unmask evolutionary warfare between sexes. *Science*, 272:953–954, 1996.
11. E. Ann Stanley, Dan Ashlock, and Leigh Tesfatsion. Iterated prisoner’s dilemma with choice and refusal. In Christopher Langton, editor, *Artificial Life III*, volume 17 of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 131–176, Reading, 1994. Addison-Wesley.
12. Stephen Wolfram. Computer software in science and mathematics. *Scientific American*, 251:188–203, 1984.