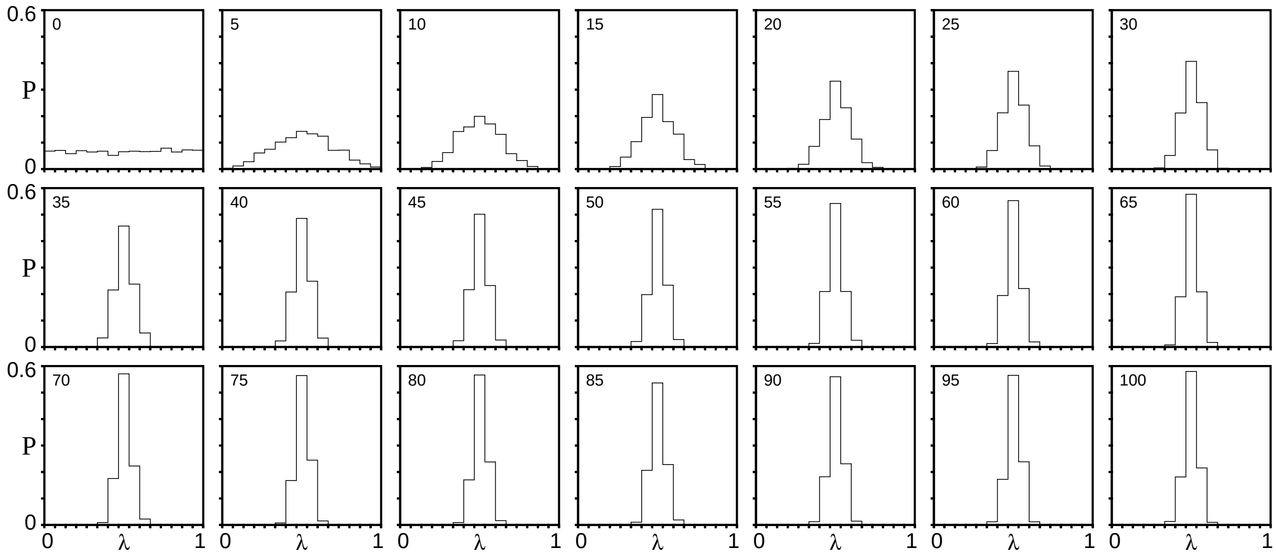

Figure 8: *No-selection mosaic:* Rule-frequency-versus-λ histograms given every five generations for populations evolved under the genetic algorithm with no selection; that is, the fitness function was not calculated. As before, each histogram is merged from 30 runs; each run had a population of 100 rules. The generation number is given in the upper left corner of each histogram.

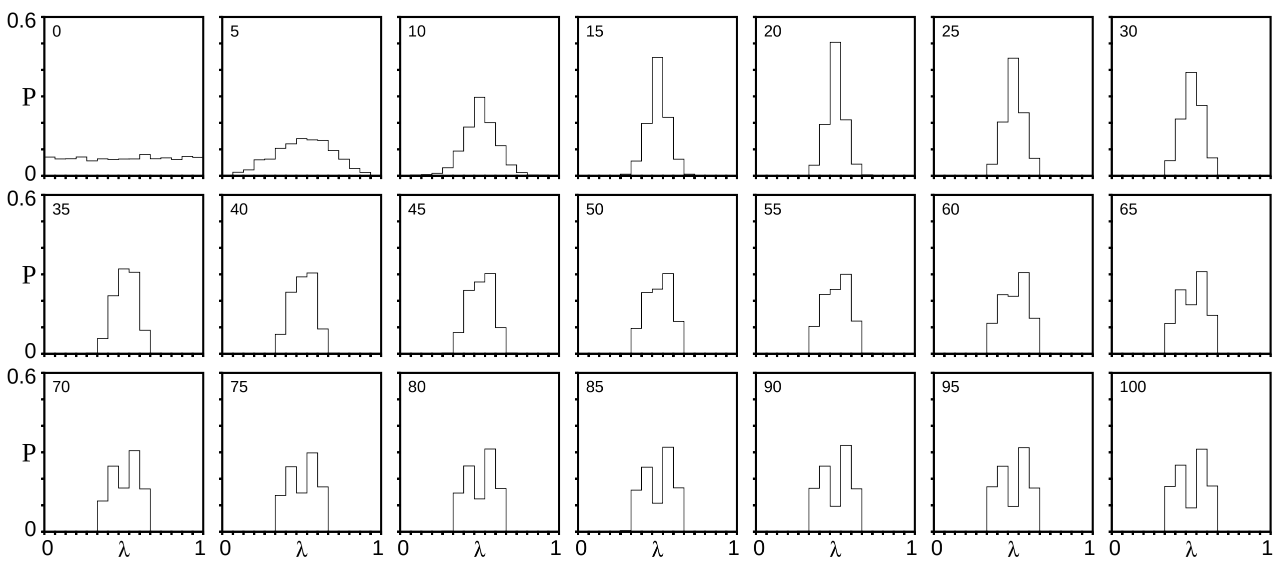

Figure 9: *Proportional-fitness mosaic:* Rule-frequency-versus-λ histograms given every five generations, merged from the 30 GA runs with proportional fitness. Each run had a population of 100 rules.

bins are beginning to rival it in magnitude. By generation 40 the right-of-center bin has exceeded the central bin, and by generation 65 the histogram has developed two peaks surrounding a dip in the center. The dip becomes increasingly pronounced as the run continues, but stabilizes by generation 85 or so.

The differences between Figure 9 and Figure 8 over all 100 generations shows that the population's structure in each generation is not entirely due to drift. Indeed, after generation 35 the distinctive features of the population indicates new, qualitatively different, and unique properties due to the selection mechanism. The two peaks represent a symmetry breaking in the evolutionary process—the rules in each individual run initially are clustered around $\lambda = 1/2$ but move to one side or the other of the central bin by around generation 35. The causes of this symmetry breaking will be discussed in the next subsection.

7.4 Evolutionary Mechanisms: Symmetry Breaking and the Dip at $\lambda = 1/2$

At this point we move away from questions related to the original experiment and instead concentrate on the mechanisms involved in producing our results. Two major questions need to be answered: Why in the final generation are there significantly fewer rules in the central bin than in the two surrounding bins? And what causes the symmetry breaking that begins near generation 35 seen in Figure 9?

In the briefest terms, the answer, obtained by detailed analysis of the 30 GA runs, is the following. The course of CA evolution under our GA roughly falls into four "strategy" epochs. Each epoch is associated with an innovation discovered by the GA for solving the problem. Though the absolute time at which these innovations appear in each run varies somewhat, each run basically passes through each of these four epochs in succession. The epochs are shown in Figure 10, which plots the best fitness, the mean fitness of the elite strings, and the mean fitness of the population versus generation for one typical run of the GA. The beginnings of epochs 2 through 4 are pointed out on the best-fitness plot. Epoch 1 begins at generation 0.

Epoch 1: Randomly generated rules

The first epoch starts at generation 0, when the best fitness in the initial generation is approximately 0.5, and the λ values are uniformly distributed between 0.0 and 1.0. No rule is much fitter than any other rule, though as was seen in Figure 5(a), rules with very low and very high λ tend to have slightly higher fitness. The strategy here—if it can be called this at all—derives from only the most elementary aspect of the task. Rules either specialize for $\rho > \rho_c$ configurations by mapping high-density neighborhoods in the CA rule table to 1 or specialize for $\rho < \rho_c$ configurations by mapping low-density neighborhoods to 0.

Epoch 2: Discovery of two halves of the rule table

The second epoch begins when a rule is discovered in which most neighborhood patterns in the rule table that have $\rho < \rho_c$ map to 0 and most neighborhood patterns in the rule table that have $\rho > \rho_c$ map to 1. Under the coding scheme we have used, this is roughly

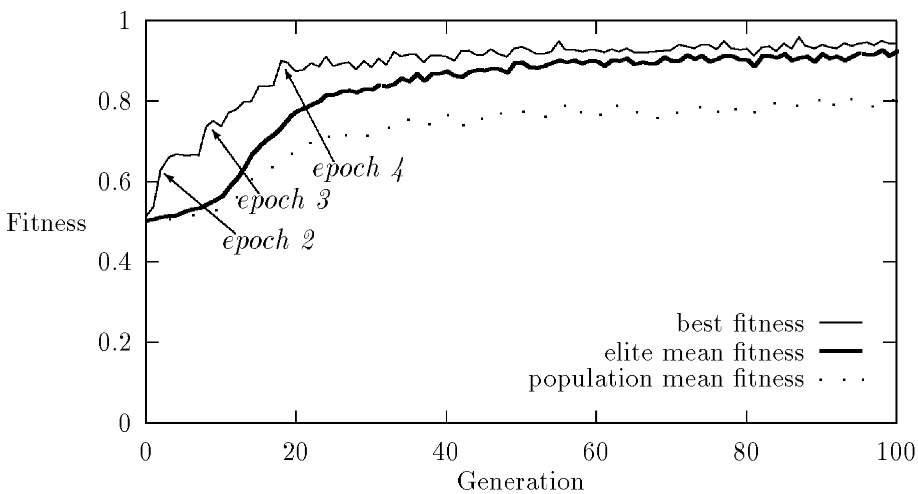

Figure 10: Best fitness, elite mean fitness, and population mean fitness versus generation for one typical run. The beginnings of epochs 2–4 are pointed out on the best-fitness plot. Epoch 1 begins at generation 0.

correlated with the left and right halves of the rule table: namely, neighborhoods 0000000 to 0111111 and 1000000 to 1111111, respectively. Given our encoding, such a strategy is presumably easy for the GA to discover due to single-point crossover's tendency to preserve contiguous sections of the rule table. It differs from the accidental strategy of epoch 0 in that there is now an organization to the rule table: output bits are roughly associated with densities of neighborhood patterns. It is the first significant attempt at distinguishing initial configurations with more 1's than 0's and vice versa. Under our fitness function, the fitness of such rules is approximately between 0.6 and 0.7, which is significantly higher than the fitness of the initial random rules. This innovation typically occurs between generations 1 and 10; in the run displayed in Figure 10 it occurred in generation 2, and can be seen as the steep rise in the best-fitness plot at that generation. All such rules tend to have λ close to 1/2. There are many possible variations on these rules with similar fitness, so such rules—all close to $\lambda = 1/2$—begin to dominate in the population. This, along with the natural tendency for the population to drift to $\lambda = 1/2$, is the cause of the clustering around $\lambda = 1/2$ seen by generation 10 in Figure 9. For the next several generations the population tends to explore small variations on this broad strategy. This can be seen in Figure 10 as the leveling off in the best-fitness plot between generations 2 and 10.

Epoch 3: Growing blocks of 1's or 0's

The next epoch begins when the GA discovers one of two new strategies. The first strategy is to increase the size of a sufficiently large block of adjacent or nearly adjacent 1's; the second strategy is to increase the size of a sufficiently large block of adjacent or nearly adjacent 0's.

Examples of these two strategies are illustrated in Figures 11 and 12. These figures give

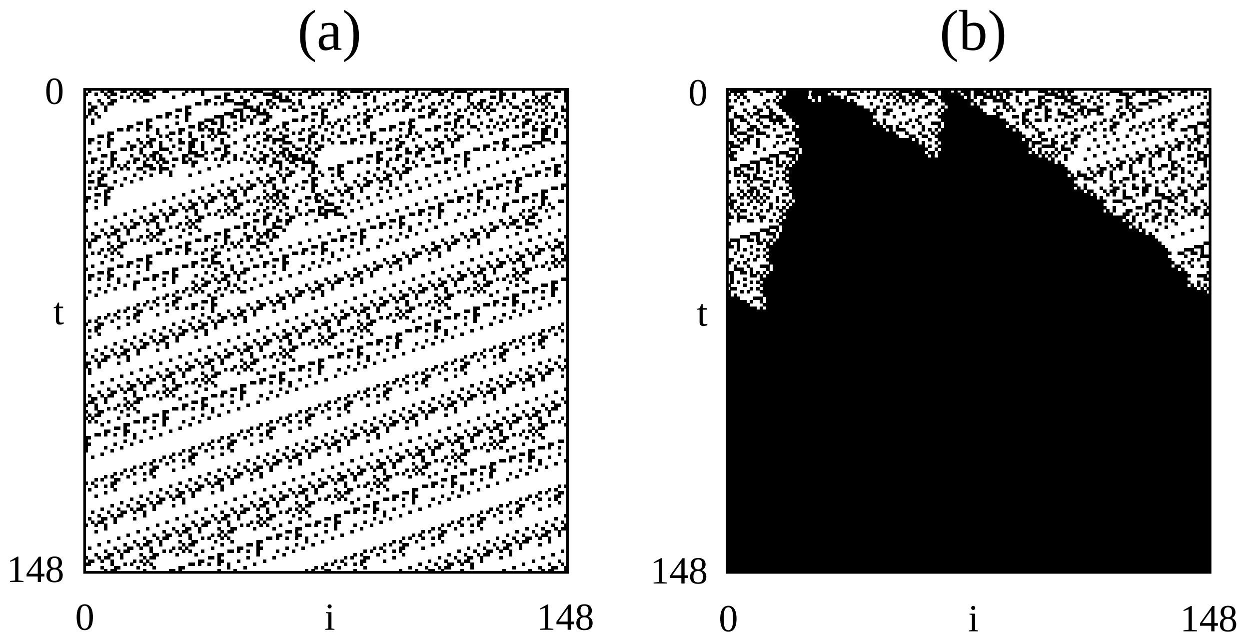

Figure 11: Space-time diagrams of one epoch-3 rule with $\lambda \approx 0.41$ that increases sufficiently large blocks of adjacent or nearly adjacent 1's. Both diagrams have $N = 149$ and are iterated for 149 time steps (the time displayed here is shorter than the actual time allotted under the GA). In (a) $\rho(0) \approx 0.40$ and $\rho(148) \approx 0.17$. In (b) $\rho(0) \approx 0.54$ and $\rho(148) = 1.0$. Thus, in (a) the classification is incorrect, but partial credit is given; in (b) it is correct.

space-time diagrams from two rules that marked the beginning of this epoch in two different runs of the GA. Figure 11 illustrates the action of a rule discovered at generation 9 of one run. This rule has $\lambda \approx 0.41$, which means that the rule maps most neighborhoods to 0. Its strategy is to map initial configurations to mostly 0's—the configurations it produces have $\rho < \rho_c$, unless the initial configuration contains a sufficiently large block of 1's, in which case it increases the size of that block. The left space-time diagram Figure 11(a) shows how the rule evolves an initial configuration with $\rho < \rho_c$ to a final lattice with mostly 0's. This produces a fairly good score. The right space-time diagram Figure 11(b) shows how the rule evolves an initial configuration with $\rho > \rho_c$. The initial configuration contains a few sufficiently large blocks of adjacent or nearly adjacent 1's, and the size of these blocks is quickly increased to yield a final lattice with all 1's for a perfect score. The fitness of this rule at generation 9 was ≈ 0.80.

Figure 12 illustrates the action of a second rule, discovered at generation 20 in another run. This rule has $\lambda \approx 0.58$, which means that the rule maps most neighborhoods to 1. Its strategy is the inverse of the previous rule. It maps initial configurations to mostly 1's unless the initial configuration contains a sufficiently large block of 0's, in which case it increases the size of that block. The left space-time diagram (a) illustrates this for an initial configuration with $\rho < \rho_c$; here a sufficiently large block of 0's appears in the initial configuration and is increased in size, yielding a perfect score. The right space-time diagram (b) shows the action of the same rule on an initial configuration with $\rho > \rho_c$. Most neighborhoods are mapped to 1 so the final configuration contains mostly 1's, yielding a fairly high score. The fitness of this rule at generation 20 was ≈ 0.87.