

Sexual and Asexual Paradigms in Evolution: The Implications for Genetic Algorithms

No Author Given

No Institute Given

Abstract. In this paper, we consider the role of sexual and asexual reproduction in evolution and relate this to the issue of optimization with genetic algorithms. In particular, we examine the results and analyses provided by [1] that imply that sexual reproduction has a general advantage over asexual reproduction. We generalize the class of models beyond those considered in [1] while still maintaining their analytical tractability. With this, we are then able to provide a more qualified description of when and why sexual-recombination can exceed mutation-based methods, and vice versa.

1 Introduction

In [1], MacKay has reported an advantage for evolutionary strategies that are based on sexual, rather than asexual, reproduction methods. MacKay compared systems that evolve by genetic recombination to ones that evolve by random mutations alone. He reports consistently faster rates of convergence to optimal fitness levels in systems based on recombination. In addition, MacKay reports that the rate of information acquired through natural selection is always greater in the case of genetic recombination. In other words, if N bits of information are needed to specify an optimal genome configuration, then the maximum rate that these bits can be acquired is faster in systems that evolve by recombination. It was proposed in [2] that the upper limit on the speed of information acquisition in natural selection is 1 bit per generation. MacKay argues that only in cases where reproduction is based solely on random mutations is the speed limit of evolution 1 bit per generation. In cases based on recombination, the upper limit on information acquisition is \sqrt{G} , where G is the size of genome.

While MacKay does not directly address the field of genetic algorithms, his results are clearly relevant to the issue of the relative power of genetic algorithms over search methods based upon stochastic hill-climbing, e.g. [3], [4]. To the extent that MacKay's results hold, they both justify the use of genetic algorithms in optimization, and contribute to the theoretical understanding of why they work. In this paper, we consider the generality of MacKay's findings and consider where and why genetic algorithms based on crossover excel over stochastic hill-climbing methods. We begin with an exposition of MacKay's study, and we illustrate his main theoretical and experimental results. We go on to generalize the models that MacKay uses, and extend his analysis to this wider class.

2 MacKay’s Model

Introducing some notation, we say that a species or organism consists of a set of individuals, each with a binary genome. The genome of individual i is denoted by $\mathbf{x}^{[i]}$, and each genome is of length G , i.e. $|\mathbf{x}^{[i]}| = G$. The set of all possible individuals in the species is denoted by \mathcal{X} and $|\mathcal{X}| = 2^G$, while a sample population of N individuals will be denoted by $\{\mathbf{x}^{[i]}\}_{i=1}^N \subseteq \mathcal{X}$. We refer to the j^{th} bit on the genome of $\mathbf{x}^{[i]}$ by x_j^i , such that $\mathbf{x}^{[i]} \equiv [x_1^i, x_2^i, \dots, x_G^i]$.

Fitness Function: The fitness of each genome $\mathbf{x}^{[i]}$ is assessed with a fitness function $\phi: \mathcal{X} \mapsto \mathbb{R}$. This function is given by

$$\phi(\mathbf{x}^{[i]}) = \sum_{j=1}^G x_j^{[i]}. \quad (1)$$

Reproduction: Reproduction occurs either through mutation, or through mutation and crossover combined. Mutation occurs by flipping each bit $x_j^{[i]}$ independently with a fixed probability m . In the case of sexual reproduction, an offspring $\mathbf{x}^{[k]}$ has parents $\mathbf{x}^{[i]}$ and $\mathbf{x}^{[j]}$. Each bit $x_i^{[k]}$ is chosen to be either $x_i^{[i]}$ or $x_i^{[j]}$, with equal probability.

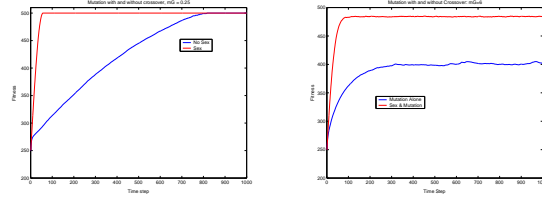
Selection: At each generation, all individuals scoring above the 50th percentile survive and reproduce to the extent that the population size stays constant at N .

2.1 Results

In Figure 1, we plot some simulations of MacKay’s model. The population N was 500. In both figures, we compare the effects of mutation alone and mutation with crossover. In the upper figure, a low mutation rate was used, i.e. $mG = .25$. While both populations converge to optimal fitness values, convergence is remarkably faster when there is sexual reproduction. In the lower figure, we use a higher mutation rate, i.e. $mG = 6$. In this case, we see considerably poorer performance by mutation alone. Sexual reproduction quickly converges to a near optimal value.

2.2 Analysis

Asexual Reproduction In a population of N individuals, each with a binary genome of length G , a mutation rate of m will flip-bits in the genome with a mean frequency of $\mu = mG$, and variance $\sigma^2 = mG(1 - m)$. If each individual in the population has exactly F positive bits, or a normalized fitness value of $f = F/G$, then the mean normalized fitness of the population after reproduction and mutation will be $f(1 - 2m) + m$, and its variance will be $m(1 - m)$. The distribution of the fitness will be roughly Gaussian, with a mean $f(1 - 2m) + m$, and variance $m(1 - m)$. After selection, the upper half of the Gaussian will be



(a) Mutation, with
and without sex:
 $mG = .25$.

(b) Mutation, with
and without sex:
 $mG = 6$.

Fig. 1. In MacKay's study, it was shown how sexual reproduction outperforms asexual reproduction. MacKay argues that, in the case of asexual reproduction, the speed limit of evolution is 1 bit per generation. However, with sexual reproduction, there is an upper limit of \sqrt{G} , where G is the size of genome.

retained. The mean will now be shifted to $\mu + \sigma\sqrt{2/\pi}$, where σ is the standard deviation, or $\sqrt{m(1-m)}$. From this, we can see that the average normalized fitness f of the population will evolve (roughly) as

$$f^{t+1} \approx f^t(1-2m) + m + \frac{\sqrt{mG(1-m)}}{G} \sqrt{2/\pi} \quad (2)$$

$$= f^t(1-2m) + m + \sqrt{\frac{2m(1-m)}{G\pi}}. \quad (3)$$

Sexual Reproduction Under the model of genetic recombination used by MacKay, the gene at position j in an offspring is inherited from the gene at the same position in one parent or the other. If both parents have exactly fG good bits on a length G genome, then there is f^2 chance of two good genes occurring at the same position on both parents. There is a $2f(1-f)$ chance that another position will have at least one good gene on either parent. The chance of inheriting a good gene will be the result of an unbiased coin-flip (so to speak). The number of good genes inherited will therefore be f^2 plus the result of $2f(1-f)G$ fair coin-tosses. This will have a mean of $f(1-f)G$, and variance of $\frac{1}{2}f(1-f)G$. The distribution of the fitness after reproduction will be roughly Gaussian, with a mean f and variance $\frac{1}{2}f(1-f)G$. Just as in the asexual reproduction case, after selection, the upper half of the Gaussian will be

retained. The mean will now be shifted to $\mu + \sigma\sqrt{2/\pi}$, where σ is the standard deviation, or $\sqrt{\frac{1}{2}f(1-f)G}$. The fitness will therefore evolve as as

$$f^{t+1} = f + \sqrt{\frac{1}{2}f(1-f)G}\sqrt{2/\pi} \quad (4)$$

$$= f + \sqrt{\frac{f(1-f)G}{\pi}}. \quad (5)$$

Why sex is good? It is immediately apparent from this analysis where the advantage of sexual reproduction lies. After asexual reproduction, the average fitness of the population tends towards a value of $\frac{F}{G} = .5$, where bits are randomly distributed, i.e. half are good and half are bad. It is only with the selection of the upper-half of the distribution, that the average fitness increases. In comparison, with sexual reproduction the average fitness does not decline. With selection, the fitness is then increased further by an amount proportional to \sqrt{G} .

3 An Extension to MacKay’s Study

In these studies, the fitness function, Equation 1, is a simple linear function of the genome. The bits in the genome do not interact:

$$P(x_i^{[k]}|x_j^{[k]}) = P(x_i^{[k]}), \quad \forall i \neq j.$$

It is questionable whether the observed results will generalize to more complex fitness models. MacKay proposes that these results will generalize: “...we believe these qualitative results would still hold if more complex models of fitness and crossover were used: *The relative benefit of sex would still scale as \sqrt{G} .*” (p. 279).

He suggests the use of a boolean function such as exclusive or, as an example of nonlinear function. As exclusive-or is the equivalent of an odd-parity function of a pair of binary variables, such a nonlinear function can be generalized to arbitrarily large blocks of variables, i.e. > 2 . These functions take the form

$$\phi(\mathbf{x}^{[i]}) = \sum_{j=1}^{G/b} \Omega_j(\mathbf{x}^{[i]}), \quad (6)$$

where b is the number of variables per block, and Ω_i is the parity of block i in the genome.

In our studies, we preserve MacKay’s original model with the exception of the fitness function.

Fitness Function: We replace the linear fitness function, Equation 1, with the nonlinear function, Equation 6.

Reproduction: Each of N individuals has a binary genome \mathbf{x} of length G . As before, mutation occurs by flipping each gene $x_j^{[i]}$ independently with a fixed

probability m . In the case of sexual reproduction, we follow MacKay: An offspring $\mathbf{x}^{[k]}$ has parents $\mathbf{x}^{[i]}$ and $\mathbf{x}^{[j]}$. Each gene $x_i^{[k]}$ is chosen to be either $x_i^{[i]}$ or $x_i^{[j]}$, with equal probability. In sexual reproduction, there may be either a zero or nonzero likelihood of mutation.

Selection: In both sexual and asexual reproduction, reproduction exactly doubles the population and selection occurs by selecting only and all the individuals that are above the median fitness level. This doubling and halving process, keeps the population size N constant in time.

3.1 Results

In the following, we report simulations using a range of mutation rates, $m \in \{.006, .00025, .0001, .00001\}$, and a range of block sizes, $b \in \{2, 5, 10, 15, 25\}$. The genome was always a fixed constant (i.e. 100) times the block length b , i.e. $G = 100b$. The population N was fixed at 500. All simulations lasted a maximum of 2500 generations.

3.2 Asexual reproduction

All results, for both sexual and asexual reproduction are shown in the Figure 2. In subfigures 2a, 2b, 2c, and 2d, we illustrate the mean fitness of the population at each generation until convergence. In each subfigure, color represents the block-size of the parity function. Mutation rates are shown in successive subfigures, i.e. a) $m = 10^{-5}$, b) $m = 10^{-4}$, c) $m = 2.5 * 10^{-4}$, d) $m = 6 * 10^{-3}$.

From these graphs, it is evident that convergence to optimal fitness occurs for $m \in \{.00025, .0001, .00001\}$, while sub-optimal performance results from $m = .006$. Interestingly, as can be seen in 2a, 2b, and 2c, in the cases of mutation rates that ultimately lead to optimal performance, large block lengths result in faster convergence. For the case of $m = .006$, longer block lengths lead to slower learning rates, and increasingly deteriorating performance, see Figure 2d. Finally, we notice that when convergence to optimal fitness occurs, it does so in approximately 400 generations.

The reason why larger block size leads to faster convergence may be explained by the fact that the mutation rate m is constant as the block length b is varied. A larger block size with constant m means a higher likelihood of flipping a bit in the block, and hence flipping the fitness of that block. It is proposed that if instead of varying m , we were to vary mb , this effect is likely to vanish. If mb is varied, then the likelihood of mutating a bit in a block is constant through different block sizes.

3.3 Sexual reproduction

In cases of sexual reproduction, convergence to optimal performance is slower than asexual reproduction. Rates decrease with increasing block size, and with increasingly mutation rates. The case of $m = .00001$, is shown in Figure 2e,

$m = .0001$ is shown in Figure 2f, $m = .00025$, is shown in Figure 2g, and $m = .006$, is shown in Figure 2h. In these sexual reproduction cases, we notice a long and increasing plateau-period prior to increasing fitness rates. These plateaus also increase with increasing block length, and increasing mutation rates. For the case of $m = .00025$ and $b = 25$, we notice that the plateau lasts for 2000 generations, while for $m = .00025$ and $b \in \{5, 10, 15, 25\}$ these plateaus exceed the maximum of 2500 generations.

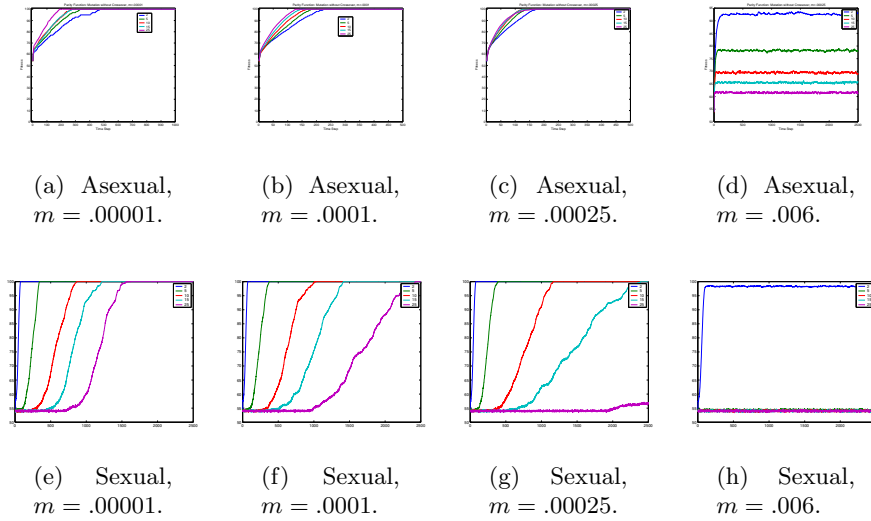


Fig. 2. Each subfigure shows fitness curves for $b \in \{2, 5, 10, 15, 25\}$ block-sizes. Different subfigures illustrate either sexual or asexual reproduction at different mutation rates.

4 Analysis

Using the parity fitness function described, the genome divides into G/b blocks, where b is the length of block to which the parity function will apply. There are 2^b possible block configurations that can occur, $\mathcal{B}_b = \{B_0, B_1, B_2, \dots, B_{2^b}\}$. For simplicity, consider the case where $b = 2$. In this case, we have $2^2 = 4$ possible

block configurations. These are

$$B_0 = \{0, 0\}, \quad (7)$$

$$B_1 = \{0, 1\}, \quad (8)$$

$$B_2 = \{1, 0\}, \quad (9)$$

$$B_3 = \{1, 1\}. \quad (10)$$

These configurations have different fitness values: $\Omega(B_0) = \Omega(B_3) = 1$, and $\Omega(B_1) = \Omega(B_2) = 0$. Let us denote $P(B_0)$ by α and $P(B_3)$ by β . We can assume that $P(B_1) = P(B_2)$. In other words, the configurations B_1 and B_2 have a symmetry, and there is no reason for the genetic algorithm to prefer an abundance of B_1 over B_2 , or vice versa. We can denote $P(B_1)$ and $P(B_2)$ by γ .

Our genetic matrix has N rows and G columns. In order to understand the evolution of the genome, we need only need consider a single column-block, which is an adjoint pair of columns. For the case of $b = 2$, this will be N by 2 matrix (Each genetic matrix can be seen as consisting of $G/2$ column-blocks.). In this matrix, the rows will be comprised of the configurations $\{B_0, B_1, B_2, B_3\}$. These configurations will occur with probabilities $\{\alpha, \gamma, \gamma, \beta\}$ respectively. Sexual reproduction, via genetic recombination, occurs by pairing two rows randomly, and creating an offspring by the method described earlier.

Sexual Reproduction Given the initial probabilities of $\{\alpha, \gamma, \gamma, \beta\}$, and using this method of sexual reproduction, we can calculate the probabilities of the configurations at the next epoch of evolution. The probability of creating a B_0 by sexual reproduction is determined by the probability of choosing a 0 as the first bit and as the second bit in the configuration. Choosing 0 as the first bit in a configuration is $\alpha + \gamma$: Both B_0 and B_1 have 0's as their first bits, and they have the probabilities of α and γ respectively. Choosing 0 as the second bit is also $\alpha + \gamma$: Both B_0 and B_2 have 0's as their second bits, and they also have the probabilities of α and γ respectively. Thus, creating a B_0 configuration through recombination is $(\alpha + \gamma)^2$. Following this reasoning, we can calculate the probabilities of the four possible configurations $\{B_0, B_1, B_2, B_3\}$ after sexual reproduction. This can be denoted by $\{\alpha', \gamma', \gamma', \beta'\}$ respectively, i.e.

$$\alpha' = (\alpha + \gamma)^2 \quad (11)$$

$$\beta' = (\beta + \gamma)^2 \quad (12)$$

$$\gamma' = (\alpha + \gamma)(\beta + \gamma). \quad (13)$$

After reproduction, the probability of a good block is $\alpha' + \beta'$. We can denote this by p , i.e. $p = (\alpha' + \beta')$. The probability of a bad block is $1 - p$. In a genome of length G and with a block length of 2, we have $G/2$ locations, and so the fitness will have a binomial distribution. The probability of k good blocks in the genome will be given by

$$\binom{\frac{G}{2}}{k} p^k (1 - p)^{\frac{G}{2} - k}.$$

The mean and variance of this distribution are given by $\mu = \frac{G}{2}p$ and $\sigma^2 = \frac{G}{2}p(1-p)$, respectively.

Selection Selection chooses the upper-half of the binomial distribution. The mean of the distribution after selection will now be positioned at the 75th percentile of the distribution prior to selection. A binomial distribution is approximately Gaussian, and 75th percentile of a Gaussian is approximately $\mu + \sqrt{\frac{2}{\pi}}\sigma$, where σ is $\sqrt{\frac{G}{2}p(1-p)}$.

Both α' and β' will be increased proportionally to the overall increase in fitness, i.e. by $\frac{1}{\mu} \left(\mu + \sqrt{\frac{2}{\pi}}\sigma \right)$. The proportions of B_0 and B_3 after selection, given by α'' and β'' respectively, will be as follows:

$$\alpha'' = \alpha' \left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu} \right) = (\alpha + \gamma)^2 \left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu} \right) \quad (14)$$

$$\beta'' = \beta' \left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu} \right) = (\beta + \gamma)^2 \left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu} \right). \quad (15)$$

As $\{\alpha'', \gamma'', \gamma'', \beta''\}$ must sum to 1,

$$\gamma'' = \frac{1 - (\alpha'' + \beta'')}{2}. \quad (16)$$

4.1 Summary of Results

The results obtained in this paper differ from those obtained for the case of the linear fitness function, i.e those reported by [1]. Moreover, they fail to support the claim of generality made in [1]. For the nonlinear fitness functions used here, asexual reproduction outperforms sexual reproduction. In cases where sexual reproduction does lead to optimal fitness, rates are slower than their asexual counterparts. In addition, we see considerable differences in the behaviors of sexual and asexual reproduction. Sexual reproduction is hindered by increasing block size, and by increasing mutation rates. In contrast, asexual reproduction shows U-shaped dependence on mutation rates, i.e. as mutation rates increase from zero, learning rates increase, and continue until a threshold where learning rates are fastest, after which point performance is sub-optimal. In addition, below the critical threshold, performance improves with increasing block size. Above the threshold, performance is hindered by increasing block size.

5 Conclusion

It is hoped that this work can contribute to the debate on the relative abilities of genetic algorithms and local search methods, e.g.. [3], [4]. It is clear that the

fitness function $\phi(\mathbf{x}^{[i]}): \mathcal{X} \mapsto \mathbb{R}$ determines largely the success of the search method. In cases where this fitness function is a linear (or presumably, quasi-linear) function of \mathcal{X} , we may see recombination-based methods exceeding the performance of mutation-based methods. In these cases, if two parents with high fitness values breed, then their offspring will also have a high fitness value. On the other hand, in cases where the fitness function ceases to have this favorable property, crossover may lead to poorer performance. Two parents with high fitness may, when they breed, lead to distinctly unfit individuals. This appears to be the case in the results that we obtain. In Figures 2e-2h, a plateau is evident at the start of evolution. Presumably in this period of time, the population is homogenizing. This homogenizing is necessary so that, when crossovers between fit individuals do occur, they will lead only to fit offspring. This homogenizing period extends with increasing mutation rates, resulting in increasingly slower rates to convergence. In these situations, the mutation-based method may excel.

We believe that although the cases studied were simple, the analytical methods used may be more generally applied and thus provide insight into the speed and performance of other evolutionary paradigms. We believe that the approach followed here may facilitate an understanding of the role of other evolutionary paradigms in search algorithms.

References

1. MacKay, D.J.C.: Information Theory, Inference and Learning Algorithms. Cambridge University Press (2003)
2. Maynard Smith, J.: The Evolution of Sex. Cambridge University Press (1978)
3. Mitchell, M., Holland, J.H., Forrest, S.: When will a genetic algorithm outperform hill climbing. In Cowan, J.D., Tesauro, G., Alspector, J., eds.: Advances in Neural Information Processing Systems. Volume 6., Morgan Kaufmann Publishers, Inc. (1994) 51–58
4. Baum, E.B., Boneh, D., Garrett, C.: Where genetic algorithms excel. In: Proceedings COLT 1995, Santa Cruz, California (1995)