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

Mark W. Andrews[1] and Christopher Salzberg[2]

[1] Gatsby Computational Neuroscience Unit,
Alexandra House, 17 Queen's Square,
London WC1N 3AR, United Kingdom,
mark@gatsby.ucl.ac.uk,
[2] Sayama Lab, Department of Human Communication,
University of Electro-Communications,
Tokyo, Japan,
chris@cx.hc.uec.ac.jp

Abstract. In this paper, we generalize the models used by MacKay [1] in his analysis of evolutionary strategies that are based on sexual, rather than asexual, reproduction methods. This analysis can contribute to the understanding of the relative power of genetic algorithms over search methods based upon stochastic hill-climbing, e.g. [2], [3].

1 Introduction

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 $\|\mathcal{X}\| = 2^G$. We refer to the jth 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 $\phi \colon \mathcal{X} \mapsto \mathbb{R}$,

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

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

Reproduction: Reproduction occurs by mutation alone, or by recombination and mutation. Mutation involves 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   Analysis

Using the parity fitness function described, the genome divides into G/b blocks. In the case where $b = 2$, we have $2^2 = 4$ possible block configurations: $B_0 = \{0,0\}, B_1 = \{0,1\}, B_2 = \{1,0\}, B_3 = \{1,1\}$. 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 $P(B_1) = P(B_2)$, and denote both by γ. Given the initial probabilities of $\{\alpha, \gamma, \gamma, \beta\}$, the probabilities of the four possible configurations $\{B_0, B_1, B_2, B_3\}$ after sexual reproduction are given $\{\alpha', \gamma', \gamma', \beta'\}$ respectively, where

$$\alpha' = (\alpha + \gamma)^2, \; \beta' = (\beta + \gamma)^2, \; \gamma' = (\alpha + \gamma)(\beta + \gamma).$$

After reproduction, the probability of a good block is $p = (\alpha' + \beta')$. 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 chooses the upper-half of this binomial distribution, which will be positioned at approximately $\mu + \sqrt{\frac{2}{\pi}}\sigma$. 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:

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

$$\beta'' = \beta'\left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu}\right) = (\beta + \gamma)^2\left(\frac{\mu + \sqrt{\frac{2}{\pi}}\sigma}{\mu}\right). \tag{3}$$

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

$$\gamma'' = \frac{1 - (\alpha'' + \beta'')}{2}. \tag{4}$$

References

1. MacKay, D.J.C.: Information Theory, Inference and Learning Algorithms. Cambridge University Press (2003)
2. 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
3. Baum, E.B., Boneh, D., Garrett, C.: Where genetic algorithms excel. In: Proceedings COLT 1995, Santa Cruz, California (1995)