Linear Selection

Mario Graff
mgraff@essex.ac.uk

Riccardo Poli
rpoli@essex.ac.uk

Alberto Moraglio
amoragn@essex.ac.uk

Department of Computer Science
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom

ABSTRACT

We investigate a form of selection, linear selection, where parents are not selected independently. One form of dependent selection, semi-linear selection, where the parents are jointly selected with a probability proportional to the average of their selection probabilities, leads the GA to behave half-way between an algorithm driven by crossover and one driven by mutation.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Automatic Programming

General Terms

Theory

Keywords

Genetic Algorithms, Selection, Theory

1. LINEAR SELECTION

Different selection methods have been analysed mathematically in depth in the last 15 years [2, 1, 4, 5]. These theoretical studies appeared to have completely characterised selection, fundamentally making it a largely understood process. However, something important has been neglected: all studies have considered forms of selection where the parent individuals are selected independently. The more general case of dependent selection has, therefore, remained a totally unchartered terrain. In this work we start filling this theoretical gap.

When crossover is used, two parents need to be selected. These are typically drawn independently, so the probability of a pair of parents (x, y), is given by the product of their selection probabilities, i.e., $p(x, y) = p(x)p(y)$. However, in principle, any assignment of $p(x, y)$ such that $p(x, y) \geq 0$ and $\sum_x \sum_y p(x, y) = 1$ would be an acceptable form of joint parent selection. In this paper we study the following forms:

- The simplest of such combinations is where a pair of parents is selected based on the straight average of the selection probabilities of the parents. That is

$$p(x, y) = \frac{p(x) + p(y)}{2} \cdot \alpha^{-1} \tag{1}$$

where α is a normalization factor such that $\sum_{x,y} p(x, y) = 1$. We will term this form of selection *pure linear selection.*

- We also consider a second form of linear selection, *semi-linear selection*, which has the following form:

$$p(x, y) = \frac{p(x) + p(y)}{2\alpha} \delta_x \delta_y \tag{2}$$

where α is a normalization factor and δ_x is 1 if x is in the population and 0 otherwise.

- We also consider

$$p(x, y) = \frac{F(x) + F(y)}{2} \phi(x)\phi(y) \cdot \alpha^{-1} \tag{3}$$

where, $\phi(x)$ is the proportion of individuals of type x in the population and $F(x)$ is a function of the fitness of individual x (but does not necessarily coincide with it). We call this form of selection *Holland's selection*.

In the presence of crossover, the infinite-population dynamics of a GA with these selections schemes is described by the following equation

$$\phi(z, t + 1) = \sum_{x\in\Omega} p(x, t) \sum_{y\in\Omega} \tilde{p}(y, t)p(x, y \to z) \tag{4}$$

where Ω is the search space. For semi-linear selection $\tilde{p}(y, t) = \frac{\delta_y(t)}{\sum_{w\in\Omega} \delta_w(t)}$, for linear selection $\tilde{p}(y, t) = \frac{1}{|\Omega|}$ and for Holland's selection $\tilde{p}(y, t) = \phi(y, t)$.

By analysing Eq. (4) and comparing it with the dynamic equations of more traditional types of GA, we have found the following: (a) linear selection behaves like normal selection with headless chicken crossover, (b) Holland's selection behaves exactly like the selection method proposed in [3], however, (c) one form of dependent selection, semi-linear selection, showed no exact connection with any pre-existing form of selection. What is interesting about it is that, when used in conjunction with crossover, in provides the GA with novel features that are somehow in between those of a crossover-based and a mutation-based GA with ordinary (independent) selection.

Our theoretical analysis was complemented by extensive experimental results which fully confirmed the predictions of the theory, including the fact that semi-linear selection leads the GA to behave half-way between an algorithm driven by crossover and one driven by mutation.

GECCO'07, July 7–11, 2007, London, England, United Kingdom.
ACM 978-1-59593-697-4/07/0007.

2. REFERENCES

[1] T. Blickle and L. Thiele. A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation*, 4(4):361–394, 1997.

[2] D. E. Goldberg and K. Deb. A comparative analysis of selection schemes used in genetic algorithms. In G. J. E. Rawlins, editor, *Foundations of Genetic Algorithms*. Morgan Kaufmann Publishers, 1991.

[3] J. H. Holland. *Adpatation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.

[4] T. Motoki. Calculating the expected loss of diversity of selection schemes. *Evolutionary Computation*, 10(4):397–422, 2002.

[5] R. Poli and W. B. Langdon. Backward-chaining evolutionary algorithms. *Artificial Intelligence*, 170(11):953–982, Aug. 2006.