A Bound on GA Convergence Using Lyapunov-like Functions

KURT BURNETTE
School of Engineering
University of Portland
Portland, Or 97203
U.S.A.

BART RYLANDER
School of Engineering
University of Portland
Portland, Or 97203
U.S.A.

Abstract: - We extend the work done in modeling the Genetic Algorithm (GA) from fundamental principles [8] by calculating a bound for convergence time for two selection schemes. When a Lyapunov function is used to show convergence, we demonstrate a method whereby the same Lyapunov function can give convergence time. We then calculate that bound for proportional selection and ranking selection. Our result is important because to date there has only been a described expected convergence for the GA given that the assumptions for the Schema Theorem [3] hold [1].

Key-Words: - Complexity, Convergence Bound, Genetic Algorithm, Evolutionary Computation

1 Background

Prior work that described the expected convergence rate of the GA was based on the many assumptions underlying the Schema Theorem [3], [1]. These assumptions are not universally agreed upon. Consequently, attempts at modeling the GA from fundamental principles have been undertaken. Using techniques from dynamical systems Vose showed that certain genetic algorithms are focused [8]. We use the same techniques to give actual bounds on convergence times. Vose showed that if a Lyapunov function is monotone, then an algorithm is focused. We show that if the *rate of change* of the Lyapunov function is monotone then convergence time can be bounded. And, we show that the rate of change of the Lyapunov function is indeed monotone for two specific selection schemes.

2 Convergence Time Bounded

We assume a simple genetic algorithm with no crossover and no mutation. The genetic algorithm is modeled by specifying a search space Ω and a simplex Λ of population vectors, the vertices of Λ being the points of Ω. The algorithm is a map G from Λ to Λ. G is focused if it is continuously differentiable and for every p contained in Λ and for every p contained in Λ the following sequence converges: p, G(p), $G^2(p)$, … G will be broken down into a composition of two functions F and M, F being the selection scheme and M being a combination of mutation and crossover. For our results, we assume M is the identity function and then G is the same as F.

If G is continuously differentiable and has finitely many fixed points, and if there is a continuous function ϕ satisfying $x \neq G(x) \Rightarrow \phi(x) > \phi(G(x))$ then G is focused [8]. In this setting, ϕ is called a Lyapunov function.

Theorem 1: Bounded Convergence Time

If F is a continuously differentiable function from Λ to Λ, and F has finitely many fixed points, and ϕ is a Lyapunov function for F, then if $\phi(F^m(p))/\phi(F^{m-1}(p)) \geq 1 + k$ (except when $F^m(p) = F^{m-1}(p)$) where $k > 0$ does not depend on m, then the convergence time of F is bounded.

Proof:

By hypothesis,

$$\phi(F(p)) \geq (1+k)\,\phi(p)$$
$$\phi(F^2(p)) \geq (1+k)\,\phi(F(p))$$
$$\phi(F^3(p)) \geq (1+k)\,\phi(F^2(p))$$

By back substituting,

$$\phi(F^3(p)) \geq (1+k)^3\,\phi(p)$$

Continuing we get,

$$\phi(F^n(p)) \geq (1+k)^n \phi(p)$$

Since Λ is compact, let M be the maximum of ϕ on Λ. Then,

$$M \geq \phi(F^n(p)) \geq (1+k)^n \phi(p)$$

So $M/\phi(p) \geq (1+k)^n$.

Therefore,

$$\log(M/\phi(p)) \geq n\log(1+k),$$

and

$$\log(M/\phi(p))/\log(1+k) \geq n.$$

QED

3 Bound For Proportional Selection

Proportional Selection is defined as $F(x) = f \cdot x/f^T x$, where $f \cdot x$ means (diag f) (x). For proportional selection, all the components of f are assumed to be positive. If p is the initial population vector, define j so that $f(j) = \max\{f(i):p(i)\neq 0\}$. If ϕ is defined to be $\phi(x) = x(j)$, then ϕ is a Lyapunov function for proportional selection [8].

Theorem 2: Convergence Bound for Proportional Selection

If F is a proportional selection scheme, then

$$\phi(F^m(p))/\phi(F^{m-1}(p)) \geq 1 + k$$

(except when $F^m(p) = F^{m-1}(p)$) where $k > 0$ does not depend on m.

Proof:

It suffices to prove

$$\phi(F(p))/\phi(p) \geq \phi(F^2(p))/\phi(F(p))$$

By definition

$$\phi(F(p))/\phi(p) \geq (f_j p_j /\Sigma f_i p_i)/p_j = f_j/\Sigma f_i p_i$$

For convenience $F(p) = q$

By definition

$$\phi(F^2(p))/\phi(F(p)) \geq (f_j q_j /\Sigma f_i q_i)/q_j = f_j/\Sigma f_i q_i$$

Comparing the last two equations, it suffices to prove

$$\Sigma f_i q_i \geq /\Sigma f_i p_i$$

By the definition of F

$$q_k = f_k p_k/\Sigma f_i p_i$$
$$f_k q_k = f_k^2 p_k/\Sigma f_i p_i$$
$$\Sigma f_i q_i = \Sigma f_i^2 p_i/\Sigma f_i p_i$$

Then it suffices to prove

$$\Sigma f_i^2 p_i/\Sigma f_i p_i \geq \Sigma f_i p_i$$

which is the same as

$$\Sigma f_i^2 p_i \geq (\Sigma f_i p_i)^2.$$

To anyone experienced with inequalities, the last inequality looks very unlikely to be true, but it turns out to be true in this setting because $\Sigma p_i = 1$.

To prove $\Sigma f_i^2 p_i \geq (\Sigma f_i p_i)^2$ we proceed by induction. If N is the size of Ω, assume that N = 2.

First observe that

$$(f_1 - f_2)^2 \geq 0$$
$$f_1^2 - 2f_1 f_2 + f_2^2 \geq 0$$

and so

$$f_1^2 + f_2^2 \geq 2 f_1 f_2$$

Now $\Sigma f_i^2 p_i = f_1^2 p_1 + f_2^2 p_2$

And because $\Sigma p_i = 1$,

$$\Sigma f_i^2 p_i = f_1^2 p_1 + f_2^2 p_2 = (f_1^2 p_1 + f_2^2 p_2)(p_1 + p_2)$$
$$= f_1^2 p_1^2 + f_2^2 p_1 p_2 + f_1^2 p_1 p_2 + f_2^2 p_2^2$$

On the other side,

$$(\Sigma f_i p_i)^2 = (f_1 p_1 + f_2 p_2)2$$
$$= f_1^2 p_1^2 + 2 f_1 f_2 p_1 p_2 + f_2^2 p_2^2$$

Then comparing these last two equations, and using the observation, we conclude that $\Sigma f_i^2 p_i \geq (\Sigma f_i p_i)^2$ for the special case N = 2.

For N > 2, the same inequality can be proved by induction. There is one complication, namely that $\Sigma p_i = 1$is only true for the given N, however, the induction proof can be used only for the formal manipulation of symbols, and the actual inequality can be proved in the special case where N is the actual size of Ω for our instance.

Now look at $\phi(F^n(p))/\phi(F^{n-1}(p))$, where n is the last n where $F^n(p) \neq F^{n-1}(p)$.

Choose k so that
$\phi(F^n(p))/\phi(F^{n-1}(p)) > 1 + k$.

QED

This latter theorem then places the proportional selection scheme within the framework of Theorem 1, which gives a bound on the convergence time.

4 Bound For Ranking Selection

Ranking selection is defined as the selection function corresponding to the selection scheme given by

$F(x)_i = \int_{\sum[f_j<f_i]x_j}^{\sum[f_j\le f_i]x_j} \rho(y)dy$ where □ is any continuous increasing probability density over [0,1]. If f is defined as for proportional selection, then f is a Lyapunov function for ranking selection [8].

Theorem 3: Convergence Bound for Ranking Selection

If F is a ranking selection scheme, then
$\phi(F^m(p))/\phi(F^{m-1}(p)) \geq 1 + k$

(except when $F^m(p) = F^{m-1}(p)$) where $k > 0$ does not depend on m.

Proof:

$$F(p)_i = \int_{\eta_i}^{\eta_i+p_i} \rho(y)dy$$

where ψ is a permutation of {0, 1, ..., n-1} such that $i<j \Rightarrow f(\psi_i) < f(\psi_j)$ and η is defined recursively by $\eta_{\psi 0}= 0$, $\eta_{\psi i+1} = \eta_{\psi i} + p_{\psi i}$.

By the mean value theorem,
$F(p)_i = \rho(\xi_i)p_i$, where ξ_i is a point in the interval
(η_i , η_i+p_i).

As in the previous proof, we would like to show that
$\phi(F(p))/\phi(p) \geq \phi(F^2(p))/\phi(F(p))$,

and we use q for the next generation, q = F(p).

The ξ_i will be different in each generation because the interval is divided up by the p_i in the first generation and q_i in the next generation, so for notational convenience, we will use ξ_i for the p generation and ζ_i for the next generation q.

In order to show
$\phi(F(p))/\phi(p) \geq \phi(F^2(p))/\phi(F(p))$

it suffices to show
$\Sigma\, \rho(\zeta_i)q_i \geq \Sigma\, \rho(\xi_i)p_i$.

By definition,
$\phi(F(p))/\phi(p) \geq \phi(F^2(p))/\phi(F(p))$

is the same as
$\rho(\xi_i)p_j / p_j \geq \rho(\eta_j)\rho(\xi_j)p_j / \rho(\xi_j)p_j$

This latter inequality is true if
$\rho(\xi_j) \geq \rho(\eta_j)$.

Because $\xi_j \geq \eta_j$ from geometry, and because ρ is increasing,
$\rho(\xi_j) \geq \rho(\eta_j)$.

QED

Once again, this theorem places ranking selection within the framework of Theorem 1, and provides a bound for convergence time.

5 Conclusions

We have extended the work at modeling the GA from fundamental principles. In particular, using the techniques from dynamical systems that Vose used to model the GA [8], we have demonstrated a method using the same Lyapunov function to derive a convergence time. We showed that the *rate of change* of the Lyapunov function that indicates convergence is monotone. We then used this fact to prove convergence time bounds for GAs that use proportional selection and ranking selection schemes.

Real world GAs are typically more complex than the GA we use in our assumptions. This does not necessarily detract from the significance of the result. Since there is no universally agreed upon GA, any assumptions for mathematical modeling would necessarily be incomplete. Since modeling the GA has proved mathematically rigorous, simplifications were required. However, a more complete understanding of one form of GA will provide insight into how all GAs behave if only as a baseline.

We have only proven the convergence time bound for two typical selection schemes. It seems likely that proving convergence time bounds for other selection schemes such as Tournament selection is possible and worthwhile. The calculation for Tournament selection would be an excellent problem for a graduate student. If such an effort is undertaken, we recommend using the results in Theorem 1 which is a very powerful and general tool. It also seems possible that convergence time bounds can be developed for GAs that use Mutation and/or Crossover, but this seems to be a truly complex task.

References

[1] Ankenbrandt, C., An Extension To the Theory of Convergence and A Proof of the Time Complexity of Genetic Algorithms, *Foundations of Genetic Algorithms*, pp. 53-68, Morgan Kaufman, 1991.

[2] Goldberg, D. E., Deb, K., A Comparative Analysis of Selection Schemes used in Genetic Algorithms. *Foundations of Genetic Algorithms*, pp. 69-93, Morgan Kaufman, 1991.

[3] Goldberg, D. E., Genetic Algorithms in Search, Optimization, and Machine Learning. Addison- Wesley, 1989.

[4] Hancock, P., An Empirical Comparison of Selection Methods in Evolutionary Algorithms. *Evolutionary Computing: AISB Workshop* Leeds, U.K. April 1994, Selected Papers, Springer-Verlag, 1994.

[5] Holland, J., Adaptation in Natural and Artificial Systems, Ann Arbor, MI, University of Michigan Press, 1975.

[6] Liepins, G., Vose, M., Representation Issues in Genetic Optimization, *Journal of Experimental and Theoretical Artificial Intelligence 2*, pp. 101-115, Morgan Kaufman, 1990.

[7] Liepins, G., Vose, M., Deceptiveness and Genetic Algorithm Dynamics, *Foundations of Genetic Algorithms*, pp. 36-52, Morgan Kaufman, 1991.

[8] Vose, M., The Simple Genetic Algorithm, Morgan Kaufman, 1999.