ELSEVIER

European Journal of Operational Research 96 (1996) 195–201

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

Theory and Methodology

A Markov chain analysis of genetic algorithms with power of 2 cardinality alphabets

Haldun Aytug [a], Siddartha Bhattacharrya [b], Gary J. Koehler [c,*]

[a] *School of Business and Engineering Administration, Michigan Technological University, Houghton, MI 49931, USA*
[b] *Department of Management, Southern Illinois University at Carbondale, REHN Hall, Carbondale, IL 62901, USA*
[c] *Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907, USA*

Received 1 March 1995; accepted 1 April 1996

Abstract

In this paper we model the run time behavior of GAs using higher cardinality representations as Markov Chains, define the states of the Markov Chain and derive the transition probabilities of the corresponding transition matrix. We analyze the behavior of this chain and obtain bounds on its convergence rate and bounds on the runtime complexity of the GA. We further investigate the effects of using binary versus higher cardinality representation of a search space.

Keywords: Genetic algorithm; Stopping criteria; Higher cardinality

1. Introduction

Genetic algorithms (GAs) are general purpose stochastic search methods modeled on natural genetics and the survival of the fittest. GAs are typically used on complex problems that are not easily solvable by conventional algorithms in reasonable time. A typical GA will start the search on a set of randomly selected strings where each string represents a solution in the search space. By applying genetic operators like crossover, mutation and selection, a new generation of solutions are probabilistically generated. The elements of each string can be a binary number (binary representation) or may come from a set of numbers with cardinality greater than two (high cardinality representation).

The Schema Theorem (see Goldberg, 1989) has traditionally formed the basis for theoretical analyses of GAs. Though providing insights into the nature of the evolutionary process, the schema theorem does not provide an adequate characterization of the genetic search (Vose, 1993) and has come under increasing criticism (Grefenstette, 1991; Grefenstette, 1992; Forrest and Mitchell, 1993; Muhlenbein, 1991).

Recent work provides a more exact analysis of the behavior of GAs. Based upon a model obtained by Vose and Liepins (1991), (Davis, 1991b also derived a similar model), a series of papers investigated the run-time behavior of GAs (Nix and Vose, 1992; Suzuki, 1993; Aytug and Koehler, 1994). As with most theoretical studies, these too have, however, considered only binary representations. Goldberg (1989) has argued for binary representations over higher cardinality representations for maximizing implicit parallelism in genetic processing. Notwithstanding this argument, practitioners report better performance with non-binary representations in many applications (Davis, 1991a) and higher level repre-

* Corresponding author. Email: koehler@mgmt.purdue.edu.

PII S0377-2217(96)00121-X

sentations also provide greater intuitive appeal. With relatively little theoretical work done on non-binary GAs (Goldberg, 1990; Wright, 1991; Eschelman and Schaffer, 1992; Bhattacharrya and Koehler, 1994), this binary vs. higher cardinality debate remains largely unresolved.

In this paper we generalize the works of Nix and Vose (1992), Aytug and Koehler (1994), and Bhattacharrya and Koehler (1994). We model the run-time behavior of GAs using higher cardinality representations as Markov Chains, define the states of the Markov Chain and derive the transition probabilities of the corresponding transition matrix. We analyze the behavior of this chain and obtain bounds on its convergence rate and thereby the bounds on the run-time complexity of the GA. We further investigate the effects of using binary versus higher cardinality representation of a search space. We consider GAs with cardinality 2^x, where x is a positive integer. Though restrictive in its generalizability to arbitrary alphabets, it nonetheless allows comparison of binary and non-binary GAs. Similar restrictions have been adopted in Eschelman and Schaffer (1992).

In Section 2 we introduce notation, review some relevant work and derive our model. Section 3 is devoted to a complexity analysis and some discussion of the behavior of GAs with 2^x cardinality alphabets compared to binary GAs. A summary of results is given in Section 4.

2. Notation

Let

- $K = 2^x$ be the cardinality of the alphabet,
- γ be the string length (or the number of digits),
- N be the number of possible different strings of length γ, ($N = K^\gamma$)
- χ be the crossover rate,
- μ be the mutation rate,
- n be the population size,
- s_i be the i-th digit of string s,
- $\delta(s_i)$ be 1 if i-th digit of string s is non-zero, 0 otherwise,
- $|s|$ be the number of non-zero digits in s,
- $r_{i,j}(k)$ be the probability of generating child k from parents i and j by using single point crossover, mutation and selection, where $i,j,k = 0 \ldots N-1$,
- F be an $N \times N$ diagonal matrix where $f_{i,i} = f(i) \equiv f_i$ is the function of interest,
- $\rho(A)$ be the spectral radius of a square matrix A,
- $\lceil x \rceil$ be the smallest integer greater than x
- $\|v\|$ be the norm of vector v defined as $\sum_l v_l$, and
- $\oplus$ be the bitwise exclusive or operator.

2.1. Preliminaries

In the following sections we analyze GAs that seek to maximize a function f, on a search space where a solution in the search space can be represented by a string of length γ. We will use the term "binary GA" to mean "a GA operating on binary coded strings".

Bhattacharrya and Koehler (1994) have shown that the equivalence described in Vose and Liepins (1991),

$$r_{i,j}(k \oplus l) = r_{i\oplus k, j\oplus k}(l),$$

still holds for GAs with 2^x cardinality alphabets. They further show that the probability of generating zero as a child from parents i and j is given by

$$\begin{aligned} r_{i,j}(0) = \frac{1}{2}(1-\mu)^\gamma \Bigg[& (1-\chi)\left(\frac{\eta}{K-1}\right)^{|i|} \\ & + \frac{\chi}{\gamma-1}\sum_{h=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|i|-\Delta_{i,j,h}} \\ & + (1-\chi)\left(\frac{\eta}{K-1}\right)^{|j|} \\ & + \frac{\chi}{\gamma-1}\sum_{h=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|j|+\Delta_{i,j,h}}, \end{aligned}$$

where

$$\eta = \frac{\mu}{1-\mu}$$

and

$$\Delta_{i,j,h} = \sum_{m=1}^{h}\delta(i_m) - \sum_{m=1}^{h}\delta(j_m),$$

where $h \le \gamma - 1$ is the crossover location. Note that substituting $K = 2$ gives us the formula derived by Vose and Liepins (1991).

Nix and Vose (1992) have shown that binary GAs can be modeled by Markov Chains. The state space of the chain is equivalent to all possible populations that can be constructed by a GA during search. They further show that the resulting chain is ergodic for $0 < \mu < 1$ and thus a steady state behavior exists. We will use their results on binary GAs to build a Markov Chain corresponding to an alphabet of $K = 2^x$, $x > 1$.

Let Z be an incidence matrix where column j corresponds to a state of the Markov Chain (i.e., a population) and the i-th entry in column j correspond to the number of copies of string i in population j. Let $z_{i,j}$ denote the i,j-th entry of matrix Z.

It is easy to show that for arbitrary $K > 1$ this matrix will have a size of $N \times T$ where T is given by

$$\binom{K^\gamma - 1 + n}{K^\gamma - 1}.$$

(See, Nix and Vose, 1992 for a proof for $K = 2$).

Nix and Vose (1992) have shown that the probability that a GA will move to population j given that the current population is i (i.e., i,j-th entry of the $T \times T$ Markov state transition matrix Q) is given by the following multinomial distribution.

$$Q_{i,j} = n! \prod_{y=0}^{N-1} \frac{\left[\overline{M}(Fz_i/\|Fz_i\|)\right]^{z_{y,j}}}{z_{y,j}!}, \quad i,j = 1 \dots T,$$

where $\overline{\mathrm{M}}$ is defined as follows. Let σ_j be a permutation defined on R^N by

$$\sigma_j s = \sigma_j \langle s_0, \dots, s_{N-1} \rangle^T = \langle s_{0 \oplus j}, \dots, s_{(N-1) \oplus j} \rangle^T.$$

Define $\overline{\mathrm{M}}: \mathfrak{R}^N \to \mathfrak{R}^N$ as

$$\overline{M}(s) = \langle (\sigma_0 s)^T M(\sigma_0 s), \dots, (\sigma_{N-1} s)^T \times M(\sigma_{N-1} s) \rangle^T$$

where M is a $N \times N$ matrix with $m_{i,j} = r_{i,j}(0)$.

Since our state definitions are exactly the same as those of Nix and Vose (1992), the state transition probabilities can be calculated in exactly the same way.

Aytug and Koehler (1994) have derived a lower bound and an easily computed upper bound on the run-time performance of binary GAs. In the following section we will use their ideas to derive an upper and a lower bound on the run-time performance of 2^x cardinality problems. They have shown that the number of iterations t required to achieve a confidence level $0 < \delta < 1$ such that we have seen an optimal solution after t iterations is bounded from below by

$$t^* = \left\lceil \max_j \left\{ \frac{\ln(1-\delta)}{\ln \rho\left(Q - Qe_j e_j^T\right)} \right\} \right\rceil \tag{1}$$

and from above by

$$\bar{t} = \left\lceil \max_j \left\{ \frac{\ln(1-\delta)}{\ln\left(1 - \min_i Q_{i,j}\right)} \right\} \right\rceil. \tag{2}$$

The upper bound requires finding the minimum element of matrix Q. Aytug and Koehler (1994) have derived such a minimum for the binary case.

The lower bound is not simplified further since we do not yet have an easily computed expression for $\rho(Q - Qe_j e_j^T)$.

3. Derivation of the bounds for K = 2^x

Since Eq. (2) requires the minimum element of Q we will first derive such an expression as summarized in the next theorem (Theorem 3.1). We show that the minimum element of Q is a function of the minimum element of M and prove that the minimum of M is bounded below by $\min(\mu^\gamma, (1-\mu)^\gamma)$ by Lemma 3.2. Demonstrating that the lower bounds in Lemma 3.2 are tight is shown in Theorem 3.1. Finally, Theorem 3.2 combines Theorem 3.1 and Eq. 2 yielding the desired bound.

Theorem 3.1. *The minimum element of Q is*

$$\left(\frac{\mu}{K-1}\right)^{n\gamma}, \quad \text{for } \mu < \frac{K-1}{K}$$

and

$$(1-\mu)^{n\gamma}, \quad \text{for } \mu \ge \frac{K-1}{K}.$$

Proof. Before we proceed we will present a result about the minimum element of multinomial distributions obtained by Aytug and Koehler (1994).

Lemma 3.1. (Aytug and Koehler, 1994). *The minimum element of a multinomial distribution*

$$\Pr\{n_1,\dots,n_k\} = n! \prod_{i=1}^{k} \frac{P_i^{n_i}}{n_i!},$$

taken over values of n_i ($\sum_{i=1}^{k} n_i = n$) *and where* p_i *is the probability of event* A_i *happening from a set of mutually exclusive events* $A_1,\dots,A_k$ *is given by*

$$(\min\{p_i\})^n.$$

Expanding the $Q_{i,j}$ *expression defined earlier gives,*

$$Q_{i,j} = n! \prod_{y=0}^{N-1} \left[\left(\left(\sum_{m=0}^{N-1} f_m z_{m,i} \right)^{-2} \sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h \times z_{h,i} r_{m,h}(y) \right)^{z_{y,j}} (z_{y,j}!)^{-1} \right].$$

Consider

$$\sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h z_{h,i} r_{m,h}(y).$$

Clearly

$$\sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h z_{h,i} \min_{m,h} \{ r_{m,h}(y) \} \le \sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h z_{h,i} r_{m,h}(y).$$

Note that

$$\min_{m,h} \{ r_{m,h}(y) \} = \min_{m \oplus y, h \oplus y} \{ r_{m \oplus y, h \oplus y}(0) \} = \min_{k,l} \{ r_{k,l}(0) \}.$$

Hence,

$$\min_{k,l} \{ r_{k,l}(0) \} \sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h z_{h,i} \le \sum_{m=0}^{N-1} \sum_{h=0}^{N-1} f_m z_{m,i} f_h z_{h,i} r_{m,h}(y).$$

Thus

$$Q_{i,j} \ge n! \prod_{y=0}^{N-1} \frac{\left(\min_{k,l} \{ r_{k,l}(0) \} \right)^{z_{y,j}}}{z_{y,j}!}.$$

We first seek to determine the minimum of the $r_{k,l}(0)$ values. This gives a bound on the minimum element of Q. We then show this bound is attained. This provides a minimum element of Q. This notion is pursued with the next result.

Lemma 3.2. *The minimum element of matrix M is bounded below by*

$$\left(\frac{\mu}{K-1} \right)^{\gamma} \quad \text{for } \mu \le \frac{K-1}{K}$$

and

$$(1-\mu)^{\gamma} \quad \text{for } \mu \ge \frac{K-1}{K}.$$

Proof. Consider the expression for $r_{i,j}(0)$ representing the probability of producing 0 from parents i and j,

$$r_{i,j}(0) = \frac{(1-\mu)^{\gamma}}{2} \left[\eta^{|i|} \left(1 - \chi + \frac{\chi}{\gamma - 1} \sum_{k=1}^{\gamma-1} \eta^{-\Delta_{i,j,k}} \right) + \eta^{|j|} \left(1 - \chi + \frac{\chi}{\gamma - 1} \sum_{k=1}^{\gamma-1} \eta^{\Delta_{i,j,k}} \right) \right].$$

Rewriting the above expression we get

$$r_{i,j}(0) = \frac{(1-\mu)^{\gamma}}{2} \left[\eta^{|i|}(1-\chi) + \eta^{|j|}(1-\chi) + \frac{\chi}{\gamma-1} \left(\sum_{k=1}^{\gamma-1} \eta^{|i| - \Delta_{i,j,k}} + \sum_{k=1}^{\gamma-1} \eta^{|j| + \Delta_{i,j,k}} \right) \right].$$

Case 1. $\mu \le (K-1)/K$.

Consider the following inequalities with $\eta \le K - 1$ (i.e., $\mu \le (K-1)/K$):

$$\left(\frac{\eta}{K-1} \right)^{|i|} (1-\chi) \ge \left(\frac{\eta}{K-1} \right)^{\gamma} (1-\chi)$$

since $|i| \le \gamma$, (3)

$$\left(\frac{\eta}{K-1} \right)^{|j|} (1-\chi) \ge \left(\frac{\eta}{K-1} \right)^{\gamma} (1-\chi)$$

since $|j| \le \gamma$, (4)

$$\frac{\chi}{\gamma-1}\sum_{k=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|i|-\Delta_{i,j,k}} \geq \chi\left(\frac{\eta}{K-1}\right)^{\gamma}$$

since $|i|-\Delta_{i,j,k} \leq \gamma$ (5)

and

$$\frac{\chi}{\gamma-1}\sum_{k=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|j|+\Delta_{i,j,k}} \geq \chi\left(\frac{\eta}{K-1}\right)^{\gamma}$$

since $|j|+\Delta_{i,j,k} \leq \gamma$. (6)

Combining (3), (4), (5) and (6), and simplifying we get

$$r_{i,j}(0) \geq \frac{(1-\mu)^{\gamma}}{2(K-1)}2\eta^{\gamma}.$$

Hence, since $\eta = \mu/(1-\mu)$,

$$r_{i,j}(0) \geq \left(\frac{\mu}{K-1}\right)^{\gamma}.$$

Case 2. $\mu \geq (K-1)/K$.

Consider the following inequalities for $\eta \geq K-1$ (i.e., $\mu \geq (K-1)/K$:

$$\left(\frac{\eta}{K-1}\right)^{|i|}(1-\chi) \geq \left(\frac{\eta}{K-1}\right)^{0}(1-\chi)$$

since $|i| \geq 0$, (7)

$$\left(\frac{\eta}{K-1}\right)^{|j|}(1-\chi) \geq \left(\frac{\eta}{K-1}\right)^{0}(1-\chi)$$

since $|j| \geq 0$, (8)

$$\frac{\chi}{\gamma-1}\sum_{k=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|i|-\Delta_{i,j,k}} \geq \chi\left(\frac{\eta}{K-1}\right)^{0}$$

since $|i|-\Delta_{i,j,k} \geq 0$ (9)

and

$$\frac{\chi}{\gamma-1}\sum_{k=1}^{\gamma-1}\left(\frac{\eta}{K-1}\right)^{|j|+\Delta_{i,j,k}} \geq \chi\left(\frac{\eta}{K-1}\right)^{0}$$

since $|j|+\Delta_{i,j,k} \geq 0$. (10)

Combining (7), (8), (9) and (10), and simplifying we get

$$r_{i,j}(0) \geq \frac{(1-\mu)^{\gamma}}{2}2\left(\frac{\eta}{K-1}\right)^{0} = (1-\mu)^{\gamma}.$$

□

Combining Lemmas 3.1 and 3.2 we get

$$Q_{i,j} \geq \left(\min\left\{\frac{\mu}{K-1}, 1-\mu\right\}\right)^{n\gamma}. \quad (11)$$

Below we show that the bounds obtained in Lemma 3.2 are tight, therefore (11) yields the minimum element of Q.

Case 1. $\mu \leq (K-1)/K$.

Consider the probability of producing string 0 from parents $N-1$ and $N-1$. Since both parents are the same, applying crossover yields $N-1$ as the offspring. Hence, the only way we can get a 0 is to mutate all the bits to 0. Then, the probability of mutating all the bits to 0 is $(\mu/(K\text{-}1))^{\gamma}$. Thus $r_{N-1,N-1} = (\mu/(K\text{-}1))^{\gamma}$. So M has a minimum element with the value $(\mu/(K\text{-}1))^{\gamma}$. Hence, using Lemma 3.1 yields $(\mu/(K\text{-}1))^{n\gamma}$. Note that this corresponds to moving from a population of all $N-1$s to a population of all zeros.

Case 2: $\mu \geq (K-1)/K$.

Consider the probability of producing 0 from parents 0 and 0. Applying crossover will yield 0 as an offspring. So the only way we can get a zero after recombination is by not mutating any of the digits. Then, the probability of getting a 0 from parents 0 and 0 is $(1-\mu)^{\gamma}$. Thus $r_{N-1,N-1} = (1-\mu)^{\gamma}$. Hence, using this minimum value in Lemma 3.1 yields $(1-\mu)^{n\gamma}$.

Using Theorem 3.1 we derive an upper bound on the number of iterations which is presented in the following theorem.

Theorem 3.2. *The number of iterations sufficient to see all populations of a cardinality K GA with probability* δ *is given by*

$$\bar{t} = \left\lceil \max_j \left\{ \frac{\ln(1-\delta)}{\ln(1-\min[(\mu/(K-1))^{n\gamma}, (1-\mu)^{n\gamma}])} \right\} \right\rceil.$$

Proof. Combining Eq. (2) and Theorem 3.1. yields the desired result.

□

It is easy to show that the minimum for $\bar{t}$ is attained for $n = 1$, and $\mu = (K-1)/K$.

Theorem 3.2 raises an interesting question for a very practical problem. Given a problem whose domain can be represented with a coding of cardinality K, what is the best choice of coding? That is, should one use binary coding or a higher cardinality coding for a given problem. We summarize the answer to this question in a theorem.

Theorem 3.3. *If $\bar{t}$ iterations guarantee that an optimal solution has been found with probability δ by using binary coding, to guarantee the same worst-case level of confidence in $\bar{t}$ or fewer iterations by using $K = 2^x$ cardinality encoding, the mutation rate of the higher cardinality GA (μ_K) has to satisfy*

$$\mu_K \geq \mu_2^x (2^x - 1),$$

where $\mu_2 \leq 0.5$ is the mutation rate used in binary encoding.

Proof. Given $K = 2^x, \gamma, n$, the binary encoding can be accomplished by using a string of length $x\gamma$. The desired number of iterations for the binary case will then be given by

$$\bar{t}_2 = \frac{\ln(1-\delta)}{\ln(1-\mu_2^{nx\gamma})}.$$

The same quantity for the general case would be given by

$$\bar{t}_K = \frac{\ln(1-\delta)}{\ln\left(1-(\mu_K/(2^x-1))^{n\gamma}\right)}.$$

Solving the inequality $\bar{t}_K \leq \bar{t}_2$ for μ_K in terms of μ_2 yields the desired result. □

Aytug and Koehler (1994) have shown that $\bar{t}_2$ attains its minimum value for $\mu_2 = 0.5$, and $n = 1$. Theorem 3.3 suggests that when an optimal mutation rate ($\mu_2 = 0.5$) is used for a binary GA a high cardinality GA should operate with a higher mutation rate, $\mu_K = (K-1)/K$ (i.e., where $\bar{t}_K$ is minimized), to complete the search in the same amount of time. It also suggests that when a binary GA has its optimal number of iterations a 2^x cardinality GA can not run faster regardless of the mutation rate used.

However, the above behavior does not hold when arbitrary mutation rates are used for the binary case. Given

$$\mu_2 < \left(\frac{1}{2^x-1}\right)^{\frac{1}{x-1}}, x > 1, \tag{12}$$

μ_K need not be greater than μ_2 so that $t_K \leq t_2$. To see this pick $\mu_2 = 0.2$ for $x = 2$, which definitely satisfies (12). Using Theorem 3.3 and solving for μ_K yields $\mu_K \geq 0.12$. So, we can achieve a lower number of iterations with a 2^x cardinality GA even when using lower mutation rates compared to the binary GA. However, when (12) is not satisfied μ_K needs to be higher than μ_2.

4. Summary

In this paper we have shown that results of Vose and Liepins (1991), Aytug and Koehler (1994) and Bhattacharrya and Koehler (1994) are applicable to GAs with 2^x cardinalities. We have derived an easily obtainable upper bound for the worst case run-time complexity of 2^x cardinality GAs and have shown that in the worst case the high cardinality GA uses only mutation as its search operator. We have also established a theoretical basis for comparing worst case performances of a binary and 2^x cardinality coding of the same problem. One interesting result is the fact that in the worst case GAs with 2^x cardinality alphabets do not always need higher mutation rates to "converge" at the same rate as binary GAs. There is a clear trade off between the cardinality of a GA and mutation rate of the other choice of coding.

Research is under way to tighten the bound on the run-time complexity of GAs. We also plan to generalize the results of this paper and future results to GAs operating on alphabets with arbitrary cardinalities.

References

Aytug, H., and Koehler, G.J. (1994), "Stopping criteria for finite length genetic algorithms", *ORSA Journal on Computing*, forthcoming.

Bhattacharrya, S., and Koehler, G.J. (1994), "An analysis of non-binary genetic algorithms with cardinality 2_x", Working Paper, Dept. of Management, REHN Hall Southern Illinois University at Carbondale, Carbondale, IL 62901.

Davis, L. (1991a), *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York.

Davis, T. (1991b), "Toward an extrapolation of the simulated annealing convergence theory onto the simple genetic algorithm", Dissertation presented to the University of Florida.

Eschelman, L.J., and Schaffer, J.D. (1992), "Real-coded genetic algorithms and interval schemata", in: L.D. Whitley (ed.), *Foundations of Genetic Algorithms* 2, Morgan Kaufmann, 187–201.

Forrest, S., and Mitchell, M. (1993), "What makes a problem hard for a genetic algorithm? Some anomalous results and their explanation", *Machine Learning* 13, 285–319.

Goldberg, D.E. (1990), "Real-coded genetic algorithms, virtual alphabets, and blocking", Technical Report IlliGAL No. 90001, University of Illinois at Urbana-Champaign.

Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Adison Wesley.

Grefenstette, J.J. (1991), "Conditions for implicit parallelism", in: G.J.E. Rawlins (ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann Publishers.

Grefenstette, J.J. (1992), "Deception considered harmful", in: L.D. Whitley (ed.), *Foundations of Genetic Algorithms* 2, Morgan Kaufmann, 75–91.

Muhlenbein, H. (1991), "Evolution in time and space – The parallel genetic algorithm", in: G.J.E. Rawlins (ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann Publishers, 316–337.

Nix, A., and Vose, M.D. (1992), "Modeling genetic algorithms with markov chains", *Annals of Mathematics and Artificial Intelligence* 5, 79–88.

Suzuki, J., (1993), "A Markov chain analysis on a genetic algorithm", *Proceedings of the Fifth International Conference on Genetic Algorithms*, University of Illinois at Urbana Champaign, 146–153.

Vose, M.D., and Liepins, G.E. (1991), "Punctuated equilibria in genetic search", *Complex Systems* 5/1, 31–44.

Vose, M.D. (1993), "A critical evaluation of the schema theorem," Working Paper, Computer Science Department, The University of Tennessee, Knoxville, TN 37996-1301.

Wright, A.H. (1991), "Genetic algorithms for real parameters optimization", in: G.J.E. Rawlins (ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann Publishers, 205–218.