Modeling and Analysis of Genetic Algorithm with Tournament Selection *

Anton V. Eremeev

Omsk Branch of Sobolev Institute of Mathematics,
13 Pevtsov st. 644099, Omsk, Russia.

Abstract. In this paper we propose a mathematical model of a simplified version of genetic algorithm (GA) based on mutation and tournament selection and obtain upper and lower bounds on expected proportion of the individuals with the fitness above certain threshold. As an illustration we consider a GA optimizing the bit-counting function and a GA for the vertex cover problem on graphs of a special structure. The theoretical estimates obtained are compared with experimental results.

1 Introduction

In this paper we propose a mathematical model of a simplified version of the genetic algorithm (GA) based on mutation and selection operators and evaluate the probability distribution of its population. We study the GA with s-tournament selection operator which randomly chooses s individuals from the previous population and selects the best one of them (see e.g. [4, 10]).

The predictions of GA behavior are based on some a priori known parameters of the mutation operator. Using the proposed model we obtain upper and lower bounds on expected proportion of the individuals with fitness above certain thresholds. These bounds resemble the well-known Schema Theorem (see e.g. [3]), but the notion of schema is not used in this model. Instead of schemata here we consider the sets of genotypes with the fitness bounded from below.

In this framework we analyze the GA optimizing the bit-counting function and a GA for the vertex cover problem on graphs of a special structure. Finally the theoretical predictions are compared with the results of computational experiments.

Let the optimization problem consist in maximization of the goal function $f : X \to \mathbf{R}$, where X is the space of solutions. The GA searches for the optimal and sub-optimal solutions using a population of individuals, where each individual consists of genotype g and phenotype $x(g) \in X$. Here g is a fixed length string of genes $g^1, g^2, \ldots, g^n$, and the genes $g^i, i = 1, 2, ..., n$ are the symbols of a certain alphabet A (the alphabet $A = \{0, 1\}$ is the most commonly used). The function $x(g)$ maps genotype g to X, thus defining the representation of solutions in GA.

* The research was supported in part by the Russian Fund for Basic Research Grant 97-01-00771.

In each iteration the GA constructs a new population on the basis of the previous one. The search process is guided by the values of nonnegative fitness function $\Phi(g) = \phi(f(x(g)))$, which defines the fitness of the individual with genotype g. Here $\phi : R \to R$ is a monotone function, which is usually used for tuning the GA for solving a particular class of problems.

The individuals of the population may be ordered according to the sequence in which they are generated, thus the population may be considered as a vector of genotypes $\Pi^t = (g_1^{(t)}, g_2^{(t)}, ..., g_N^{(t)})$, where N is the size of population, which is constant during the run of the GA, and t is the number of the current iteration. In this paper we consider the populational version of GA, where all individuals of a new population are generated independently from each other with identical probability distribution depending on the existing population only.

Each individual is generated through selection of a parent genotype by means of selection operator, and modification of this genotype in mutation operator. During the mutation a subset of genes in the genotype string g is randomly altered. In general the mutation operator may be viewed as a random variable $Mut(g) \in A^n$ (where A^n is the set of all genotype strings) with the probability distribution depending on g. The most frequently used type of this operator randomly changes each gene of g with a fixed mutation probability p_m.

The genotypes of the initial population Π^0 are generated with some a priori chosen probability distribution. The stopping criterion is usually the limit of maximum iterations t_{max}. The result is the best solution generated during the run. In this paper we will consider a GA with the following scheme.

1. Generate the initial population Π^0.
2. For $t := 0$ to $t_{max} - 1$ do
   2.1. For $k := 1$ to N do
      Choose the parent genotype g from Π^t using s-tournament selection.
      Mutate g.
      Add $g_k^{(t+1)} = Mut(g)$ to the population Π^{t+1}.

The s-tournament selection operator randomly chooses s individuals from the previous population and selects the best one of them. In literature there are other standard operators, besides the tournament selection, in use. For example, the proportional selection where the selection probability is proportional to the fitness of individuals [3], or the truncation selection, where the parents are randomly chosen from $T\%$ best individuals of the previous population [7].

The GA may be considered as a Markov chain in a number of ways. For example, the states of the chain may correspond to different vectors of N genotypes that constitute the population Π^t. In this case if the genotype consists of n genes from the alphabet $A = \{0, 1\}$ then the number of states in the Markov chain is 2^{nN}. With the help of this model for the GA with proportional selection the following result is obtained [9].

Theorem 1. *Let x^t be the best solution found by GA until the step t; $f^* = \max\{f(x(g)) : g \in \{0,1\}^n\}$, and the probability of changing a gene in mutation is $p_m \in (0,1)$. Then $\lim_{t\to\infty} P\{f(x^t) = f^*\} = 1$.*

It is also proven that if this GA worked for infinite number of iterations, the genotypes corresponding to the optimal solutions would be lost and found infinitely often (see [9]). The similar results may be obtained for the GAs with tournament selection.

Another model representing the GA as a Markov chain has been proposed in [8], where all populations which differ only in the ordering of individuals are considered to be equivalent. In case if $A = \{0, 1\}$, each state of this Markov chain may be represented by a vector of 2^n components, where the proportion of each genotype in the population is indicated by the corresponding coordinate. In the framework of this model, M.Vose and collaborates have obtained a number of general results concerning the emergent behavior of GA [8, 11, 12].

The major difficulties in practical application of these models to real optimization problems are connected with the necessity to use the a priori information on fitness value of each genotype. Besides that, the size of the transition matrix grows exponentially as the length of genotype increases. In this paper we will consider one of the possible ways to handle these difficulties. The way we will use is the grouping of the states of population into larger classes.

2 The Approximating Model

The model proposed here represents the GA with tournament selection as it was introduced earlier. Here the information about the fitness of separate individuals is not used explicitly. Instead, the model makes use of certain a priori known parameters of probability distribution of the mutation operator described below.

Assume that there are d level lines of the fitness function fixed such that $\Phi_0 = 0 < \Phi_1 < \Phi_2 \ldots < \Phi_d$. The number of level lines and the fitness values corresponding to them may be chosen arbitrarily, but they should be relevant to the given problem and the mutation operator to yield a meaningful model. Let us introduce the following sequence of subsets of the set A^n:

$$H_i = \{g : \Phi(g) \geq \Phi_i\}, \quad i = 0, \ldots, d.$$

Due to nonnegativity of the fitness function, H_0 equals to the set of all genotypes. Besides that, for the sake of convenience, let us define the set $H_{d+1} = \emptyset$.

Now suppose that for all $i = 0, ..., d$ and $j = 1, ..., d$ the a priori lower bounds α_{ij} and upper bounds β_{ij} on mutation transition probability from subset $H_i \backslash H_{i+1}$ to H_j are known, i.e. for every $g \in H_i \backslash H_{i+1}$ holds $\alpha_{ij} \leq P\{Mut(g) \in H_j | g\} \leq \beta_{ij}$, where $P\{Mut(g) \in H_j | g\} = \sum_{g' \in H_j} P\{Mut(g) = g' | g\}$.

Let $\mathbf{A}$ denote the matrix with the elements α_{ij} where $i = 0, ..., d$, and $j = 1, ..., d$. The similar matrix of upper bounds β_{ij} is denoted by $\mathbf{B}$. Let the population on iteration t be represented by the *population vector* $z^{(t)} = (z_1^{(t)}, z_2^{(t)}, \ldots, z_d^{(t)})$ where $z_i^{(t)} \in \mathbf{R}$ is the proportion of genotypes from H_i within the population Π^t. The vector $z^{(t)}$ is a random vector, where $z_i^{(t)} \geq z_{i+1}^{(t)}$ for $i = 1, ..., d-1$ since $H_{i+1} \subseteq H_i$. Let $P\{g^{(t)} \in H_j\}$ be the probability that an individual, which is added after selection and mutation into Π^t, has a genotype from H_j for $j = 0, ..., d$, and $t > 0$ (according to the scheme of the algorithm this probability is identical for all genotypes of Π^t, i.e. $P\{g^{(t)} \in H_j\} = P\{g_1^{(t)} \in H_j\} = ... = P\{g_N^{(t)} \in H_j\}$). Let E denote the mathematical expectation. Then it is not difficult to obtain the following:

Proposition 2. *$E[z_i^{(t)}] = P\{g^{(t)} \in H_i\}$ for all $t > 0, i = 1, ..., d$.*

3 Bounding the Expectation of Population Vector

In this section the lower and upper bounds on proportion of genotypes from H_i for all $i = 1, \dots, d$ will be considered. Let $P_{ch}(S, z)$ denote the probability that the genotype, chosen by the tournament selection from the current population with vector z, belongs to a subset $S \subseteq A^n$. Note that if the current population is represented by the vector $z^{(t)} = z$, then a genotype obtained by selection and mutation would belong to H_j with conditional probability

$$P\{g^{(t+1)} \in H_j | z^{(t)} = z\} = \sum_{i=0}^{d} \sum_{g \in H_i \setminus H_{i+1}} P\{Mut(g) \in H_j | g\} P_{ch}(\{g\}, z) \geq$$

$$\sum_{i=0}^{d} \alpha_{ij} \sum_{g \in H_i \setminus H_{i+1}} P_{ch}(\{g\}, z) = \sum_{i=0}^{d} \alpha_{ij} P_{ch}(H_i \setminus H_{i+1}, z). \tag{1}$$

Given the tournament size s we obtain the following selection probabilities: $P_{ch}(H_i, z^{(t)}) = 1 - (1 - z_i^{(t)})^s$, and, consequently, $P_{ch}(H_i \setminus H_{i+1}, z) = P_{ch}(H_i, z) - P_{ch}(H_{i+1}, z) = (1 - z_{i+1})^s - (1 - z_i)^s$. This leads to the inequality:

$$P\{g^{(t+1)} \in H_j | z^{(t)} = z\} \geq \sum_{i=0}^{d} \alpha_{ij} ((1 - z_{i+1})^s - (1 - z_i)^s).$$

Let $Z_N = \{z \in \mathbf{R}^d : z_i \in \{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\}, z_i \geq z_{i+1}\}$ be the set of all possible vectors of population which consists of N individuals. Then using the total probability formula we obtain the following bound on unconditional probability:

$$P\{g^{(t+1)} \in H_j\} \geq \sum_{z \in Z_N} \sum_{i=0}^{d} \alpha_{ij} ((1 - z_{i+1})^s - (1 - z_i)^s) P\{z^{(t)} = z\} =$$

$$\sum_{i=0}^{d} \alpha_{ij} E[(1 - z_{i+1}^{(t)})^s - (1 - z_i^{(t)})^s]. \tag{2}$$

The Proposition 2 yields that $E[z_j^{(t+1)}] = P\{g^{(t+1)} \in H_j\}$. Consequently since $(1 - z_{d+1}^{(t)})^s = 1$ and $(1 - z_0^{(t)})^s = 0$,

$$E[z_j^{(t+1)}] \geq \alpha_{dj} - \sum_{i=1}^{d} (\alpha_{i,j} - \alpha_{i-1,j}) E[(1 - z_i^{(t)})^s]. \tag{3}$$

Denote $I_j^+(\mathbf{A}) = \{1 \le i \le d : \alpha_{i,j} - \alpha_{i-1,j} \ge 0\}$ and $I_j^-(\mathbf{A}) = \{1 \le i \le d : \alpha_{i,j} - \alpha_{i-1,j} < 0\}$. Let us apply the Jensen inequality to those terms of sum (3) for which $i \in I_j^-(\mathbf{A})$ (in view of the fact that under the expectation sign the function on $z_i^{(t)}$ is convex). The terms where $i \in I_j^+(\mathbf{A})$ can be bounded using the simple estimate $(1-z_i^{(t)})^s \le 1-z_i^{(t)}$. This leads to the following proposition.

Proposition 3. *The components $E[z_j^{(t+1)}]$ of expected next population vector are bounded below for all $j = 1, ..., d$ and $t \ge 0$ as follows*

$$E[z_j^{(t+1)}] \ge \alpha_{dj} - \sum_{I_j^+(\mathbf{A})} (\alpha_{ij} - \alpha_{i-1,j})(1 - E[z_i^{(t)}]) - \sum_{I_j^-(\mathbf{A})} (\alpha_{ij} - \alpha_{i-1,j})(1 - E[z_i^{(t)}])^s. \tag{4}$$

Note that if the probability $P\{Mut(g) \in H_j | g\}$ for all $i = 0, ..., d$, and $j = 1, ..., d$ does not depend on choice of $g \in H_i \backslash H_{i+1}$, then we may assume that $\alpha_{ij} = \beta_{ij} = P\{Mut(g) \in H_j | g \in H_i \backslash H_{i+1}\}$ for all i and j. In this case the mutation operator will be called a *step mutation operator with respect to the sequence of subsets* $H_0, H_1, ..., H_d$ (or a step mutation operator for short).

In order to emphasize the fact that $\gamma_{ij} = P\{Mut(g) \in H_j | g \in H_i \backslash H_{i+1}\}$ is not the transition probability used in the Markov chains (which is not considered in this paper) sometimes we will call γ_{ij} *the threshold transition probability.* The matrix Γ will detone here the matrix of threshold transition probabilities of a step mutation operator: $\Gamma = \mathbf{A} = \mathbf{B}$.

If the tournament size $s = 1$, then the selection has a uniform distribution and its operation does not depend on the fitness of individuals. It is easy to see that in this special case (4) becomes tight for the GA with step mutation operator. Our aim in this section is to obtain a lower bound on $E[z^{(t)}]$ for arbitrary t if the expectation of the initial population vector $E[z^{(0)}]$ is given. In order to do this let us introduce the following definition.

A matrix $\mathbf{M}$ with elements $m_{ij}, i = 0, ..., d$, and $j = 1, ..., d$ will be called *monotone* if $m_{i-1,j} \le m_{i,j}$ for all i, j from 1 to d.

The matrix of bounds on transition probabilities is monotone if for any $j = 1, ..., d$ the genotypes from any subset H_i have the bound on transition probability to H_j not less than the bounds of the genotypes from the subsets $H_{i'}$ for all $i' < i$. Obviously, for any mutation operator the monotone bounds exist. (For example $\mathbf{A} = \mathbf{0}$ where $\mathbf{0}$ is a zero matrix and $\mathbf{B} = \mathbf{U}$ where $\mathbf{U}$ is the matrix with all elements equal 1). The problem may be connected only with the absence of bounds which are sharp enough to evaluate the mutation operator properly. Imposing the assumption of monotone bounds on (4) we derive the following :

Proposition 4. *If the matrix $\mathbf{A}$ is monotone, then for any tournament size $s \ge 1$ and $j = 1, \ldots, d$*

$$E[z_j^{(t+1)}] \ge \alpha_{0j} + \sum_{i=1}^{d} (\alpha_{i,j} - \alpha_{i-1,j}) E[z_i^{(t)}]. \tag{5}$$

Let $\mathbf{W}$ be a $(d \times d)$ matrix with the elements $w_{ij} = \alpha_{ij} - \alpha_{i-1,j}$; $\mathbf{I}$ is the identity matrix of the same size, and vector $\alpha = (\alpha_{01}, ..., \alpha_{0d})$.

Theorem 5. *If the matrix* $\mathbf{A}$ *is monotone and* $\alpha_{dj} - \alpha_{0j} < 1$ *for* $j = 1, ..., d$, *then for all* $t \geq 0$

$$E[z^{(t)}] \geq E[z^{(0)}]\mathbf{W}^t + \alpha(\mathbf{I} - \mathbf{W})^{-1}(\mathbf{I} - \mathbf{W}^t). \tag{6}$$

Proof. Let us consider a sequence of d-dimensional vectors $y^{(0)}, y^{(1)}, ..., y^{(t)}, ...$, where $y^{(0)} = E[z^{(0)}]$, $y^{(t+1)} = y^{(t)}\mathbf{W} + \alpha$. Note that the right-hand side of (5) will not increase if the components of $E[z^{(t)}]$ are substituted with their lower bounds. Thus by induction on t we get: $E[z^{(t)}] \geq y^{(t)}$ for any t.

Consider the vector norm $||z|| = \max_j |z_j|$ in $\mathbf{R}^d$ and the matrix norm $||W|| = \max_j \sum_{i=1}^d |w_{ij}|$ corresponding to it. Under the conditions of this theorem we have $w_{ij} = \alpha_{ij} - \alpha_{i-1,j} \geq 0$, and $||W|| = \max_j \sum_{i=1}^d w_{ij} = \max_j(\alpha_{dj} - \alpha_{0j}) < 1$. Therefore by the properties of linear operators we conclude that the matrix $(\mathbf{I} - \mathbf{W})^{-1}$ exists. Using the induction on t we obtain the identity: $y^{(t)} = y^{(0)}\mathbf{W}^t + \alpha(\mathbf{I} - \mathbf{W})^{-1}(\mathbf{I} - \mathbf{W}^t)$, which leads to (6). □

Note that in most of the GA implementations an arbitrary given genotype may be produced with a non-zero probability as a result of mutation, and the corresponding Markov chain is ergodic (see e.g. [9]). In this case the condition $\alpha_{dj} - \alpha_{0j} < 1$ is obviously satisfied for all j, if the bounds are properly chosen.

As it follows from the proof of the Theorem 5, $||W^t|| \leq ||W||^t < 1$ and (6) approaches $\alpha(\mathbf{I} - \mathbf{W})^{-1}$ when t tends to infinity, thus the limit of this bound does not depend on distribution of the initial population. Also let us note that the inequality (6) turns into equation in the case of a step mutation operator and the tournament size $s = 1$.

By reasoning similar to the proof of Proposition 3 we obtain the following:

Proposition 6. *The components of expected next population vector are bounded above as follows*

$$E[z_j^{(t+1)}] \leq \beta_{dj} - \sum_{I_j^-(\mathbf{B})} (\beta_{ij} - \beta_{i-1,j})(1 - E[z_i^{(t)}]) - \sum_{I_j^+(\mathbf{B})} (\beta_{ij} - \beta_{i-1,j})(1 - E[z_i^{(t)}])^s \tag{7}$$

for $j = 1, \ldots, d$, *and if the matrix* $\mathbf{B}$ *is monotone then*

$$E[z_j^{(t+1)}] \leq \beta_{dj} - \sum_{i=1}^{d} (\beta_{ij} - \beta_{i-1,j})(1 - E[z_i^{(t)}])^s. \tag{8}$$

By means of iterative application of (8) the components of the vectors $E[z^{(t)}]$ may be bounded up to arbitrary t, starting from the expectation of the initial population vector $E[z^{(0)}]$. The nonlinearity in the right-hand side of (8), however, creates an obstacle for obtaining an analytical bound similar to the bound (6) of Theorem 5. This problem could be tackled using the contractive transformations theory or by approximation of the corresponding difference equation with the help of differential one, but this investigation is outside the scope of this paper.

Note that all of the bounds obtained up to this point do not include the population size and they are valid for arbitrary N. The Theorem 7 in the following section will show that the right-hand side of (8) reflects the asymptotic behavior of population under step mutation operator as $N \to \infty$.

3.1 GA with a Step Mutation Operator

If the GA uses a step mutation operator, the probability distribution of the next population is completely determined by the vector of the current population. In this case the GA may be viewed as a Markov chain with the states corresponding to the elements of Z_N.

Note that in general the population vectors are random values depending on N. In order to express this fact in notation let us denote the proportion of genotypes from H_i in population Π^t by $z_i^t(N)$.

Theorem 7. *Let the GA use a step mutation operator with monotone threshold transition matrix Γ, and let the genotypes of the initial population be identically distributed. Assume that the sequence of d-dimensional vectors $y^{(0)}, y^{(1)}, ..., y^{(t)}, ...$ is defined as follows:*

$$y^{(0)} = E[z^{(0)}(N)], \qquad y_j^{(t+1)} = \gamma_{dj} - \sum_{i=1}^{d}(\gamma_{ij} - \gamma_{i-1,j})(1 - y_i^{(t)})^s \tag{9}$$

for $j = 1, ..., d$ and $t \geq 0$. Then $E[z^{(t)}(N)] \underset{N\to\infty}{\longrightarrow} y^{(t)}$ at any iteration t.

Proof. Since the GA is based on a step mutation operator,

$$E[z_j^{(t+1)}(N)] = \gamma_{dj} - \sum_{i=1}^{d}(\gamma_{ij} - \gamma_{i-1,j})E[(1 - z_i^{(t)}(N))^s].$$

Consequently if it is proven that

$$\lim_{N\to\infty}\left(E[(1 - z_i^{(t)}(N))^s] - (1 - E[z_i^{(t)}(N)])^s\right) = 0, \tag{10}$$

then the convergence of $E[z^{(t)}(N)]$ to $y^{(t)}$ will mean that $E[z^{(t+1)}(N)]$ converges to $y^{(t+1)}$ as $N \to \infty$. In this case the statement of the theorem follows by induction on t for arbitrary finite t.

In order to prove (10) let us fix $t \geq 0$ and consider the sequence of independent identically distributed random variables $\xi_1^i, \xi_2^i, ..., \xi_N^i$, where $\xi_l^i = 1$ if the genotype of l-th individual in population $\Pi^{(t)}$ belongs to H_i, and $\xi_l^i = 0$ otherwise.

Using the law of large numbers, for any $i = 1, ..., d$ and $\varepsilon > 0$ we obtain

$$P\left\{\left|\frac{\sum_{l=1}^{N}\xi_l^i}{N} - E[\xi_1^i]\right| < \varepsilon\right\} \underset{N\to\infty}{\longrightarrow} 1.$$

Note that $\sum_{l=1}^{N} \xi_l^i / N = z_i^{(t)}(N)$, and besides that, in view of Proposition 2, $E[\xi_1^i] = P\{\xi_1^i = 1\} = E[z_i^{(t)}(N)]$ (in case if $t = 0$ this equality also holds, since the genotypes of the initial population are identically distributed). Consequently, $P\left\{\left|z_i^{(t)}(N) - E[z_i^{(t)}(N)]\right| < \varepsilon\right\} \underset{N\to\infty}{\longrightarrow} 1$ for any $\varepsilon > 0$. Hence, by continuity of the function $(1-x)^s$, it follows that

$$P\left\{\left|(1 - z_i^{(t)}(N))^s - (1 - E[z_i^{(t)}(N)])^s\right| \geq \varepsilon\right\} \underset{N\to\infty}{\longrightarrow} 0.$$

Let us denote $\mathrm{F}_N(x) = P\left\{(1 - z_i^{(t)}(N))^s - (1 - E[z_i^{(t)}(N)])^s < x\right\}$. Then

$$\lim_{N\to\infty}\left(E\left[(1 - z_i^{(t)}(N))^s\right] - (1 - E[z_i^{(t)}(N)])^s\right) = \lim_{N\to\infty}\int_{-\infty}^{\infty} x\, d\mathrm{F}_N(x) \leq$$

$$P\left\{\left|(1 - z_i^{(t)}(N))^s - (1 - E[z_i^{(t)}(N)])^s\right| \geq \varepsilon\right\} + \lim_{N\to\infty}\int_{|x|<\varepsilon} \varepsilon\, d\mathrm{F}_N(x) \underset{N\to\infty}{\longrightarrow} \varepsilon.$$

for arbitrary $\varepsilon > 0$, hence (10) holds. □

Corollary 8. *Suppose that a step mutation operator with monotone threshold transition matrix is given. Let $z^{(t)}(N)$ be a population vector of GA with tournament size s, and let $\hat{z}^{(t)}(N)$ represent a population of GA with tournament size $\hat{s} \geq s$.*

If the individuals of initial populations are identically distributed, and $E[\hat{z}_i^{(0)}(N)] \geq E[z_i^{(0)}(N)]$ for $i = 1, ..., d$, then for any iteration t and $i = 1, ..., d$ $E[\hat{z}_i^{(t)}(N)] \geq E[z_i^{(t)}(N)]$ holds, provided that N is big enough.

4 Some Applications of the Model

First we shall consider the simple *bit-counting* fitness function $\Phi : \{0,1\}^n \to \{0, ..., n\}$, which equals the number of 1's in the binary string of genotype. For this problem a number of versions of GA with truncation selection have been considered in [1, 5] and other papers. Suppose that the GA uses the standard mutation operator, changing every gene with probability p_m. Let the subsets $H_0, ..., H_d$ be defined by the level lines $\Phi_0 = 0, \Phi_1 = 1, ..., \Phi_d = d$ and $d = n$. It is easy to see that in this case the GA has a step mutation operator. The matrix Γ for this operator could be obtained using the result from [1], but here we shall consider this example as a special case of the following problem.

Let the representation of the problem admit a decomposition of the genotype string into d nonoverlapping substrings (called *blocks* here) in such a way that the fitness function Φ equals to the number of blocks for which a certain property **K** holds [2]. Let m be the number of blocks and let $K(g, \lambda) = 1$ if **K** holds for the block λ of genotype g, and $K(g, \lambda) = 0$ otherwise (here $\lambda = 1, ..., m$).

[2] These functions are a special case of the additively decomposed functions, where the elementary functions are boolean and substrings are nonoverlapping (see e.g. [6]).

Suppose that during mutation, any block for which **K** did not hold gets the property **K** with probability $\tilde{r}$, i.e. $P\{K(Mut(g),\lambda)=1|K(g,\lambda)=0\}=\tilde{r}$ for $\lambda=1,...,m$. On the other hand, assume that a block with the property **K** keeps this property during mutation with probability r, i.e. $P\{K(Mut(g),\lambda)=1|K(g,\lambda)=1\}=r; \lambda=1,...,m$. Let the subsets $H_0,...,H_d$ correspond to the level lines $\Phi_0=0,\Phi_1=1,...,\Phi_d=d$ again. In this case the element γ_{ij} of threshold transition matrix Γ equals the probability to obtain a genotype containing j or more blocks with property **K** after mutation of a genotype which contained i blocks with this property. It is not difficult to see that for this mutation operator the threshold transition probabilities $\gamma_{ij}=P\{Mut(g)\in H_j|g\in H_i\backslash H_{i+1}\}$ for $i=0,...,d, j=1,...,d$ are given by the following expression:

$$\gamma_{ij}=\sum_{k=0}^{d-i}\binom{d-i}{k}\tilde{r}^k(1-\tilde{r})^{d-i-k}\sum_{\nu=0}^{\min\{i,i-j+k\}}\binom{i}{\nu}(1-r)^{\nu}r^{i-\nu}. \qquad (11)$$

Verifying the definition of monotone matrix we obtain the following:

Proposition 9. *If $r\geq\tilde{r}$ then the matrix Γ defined by (11) is monotone.*

Now for the bit-counting function the matrix Γ is obtained, assuming that $\tilde{r}=(1-r)=p_m$, $d=n$. Obviously, this operator is monotone if $p_m\leq 0.5$.

The formula (11) may be also used for finding the threshold transition matrices of some other optimization problems with a "regular" structure. As an example we consider the vertex cover problem (VCP) on graphs of a special structure. In general, the vertex cover problem is formulated as follows.

Let $G=(V,E)$ be a graph with a set of vertices V and the edge set E. A subset $C\subseteq V$ is called a vertex cover of G if every edge has at least one endpoint in C. The vertex cover problem is to find a vertex cover C^* of minimal cardinality.

Let G_d be a graph consisting of d disconnected triangle subgraphs. Obviously, each triangle is covered optimally by two vertices and the redundant cover consists of three vertices. In spite of the simplicity of this problem, it is proven in [13] that the working time of some well-known algorithms of branch and bound type grows exponentially on d if they are applied to the VCP on graph G_d.

Suppose that the VCP is handled by the GA with *non-binary* representation (see e.g. [2]): each gene $g^i\in\{0,1\}, i=1,...,|E|$ corresponds to an edge of G_d, assigning one of its endpoints which has to be included in the cover. The phenotype $C=x(g)$ is a cover, containing all vertices which are assigned by at least one of the genes. Let the mutation operator alter each gene with probability p_m. The natural way to choose the fitness function in this case is to assume $\Phi(g)=|V|-|x(g)|$. Then for G_d the fitness $\Phi(g)$ coincides with the number of optimally covered blocks in $C=x(g)$. Let the genes representing the same triangle constitute a single block, and let the property **K** imply that a block is optimally covered. Then by looking at all possible ways of producing the redundant covers of the triangle subgraph one can see that $\tilde{r}=1-p_m^3-(1-p_m)^3$, and $r=1-p_m(1-p_m)^2-p_m^2(1-p_m)$. Using (11) we obtain the threshold transition matrix for this mutation operator. It is easy to verify that in this case the inequality $r \geq \tilde{r}$ holds for any mutation probability p_m, and therefore the operator is always monotone.

4.1 Computation Experiments

This section presents some experimental results in comparison with the theoretical estimates obtained above. Here we will consider the application of GA to the VCP on graph G_d as it was described before. The proportion of the optimal genotypes in the population for GAs with different population size is presented in Fig.1. Here $d = 8$, $p_m = 0.01$, $s = 2$ and $z^{(0)} = 0$ (i.e. the initial population consists of genotypes that define a cover where each triangle subgraph is covered redundantly).

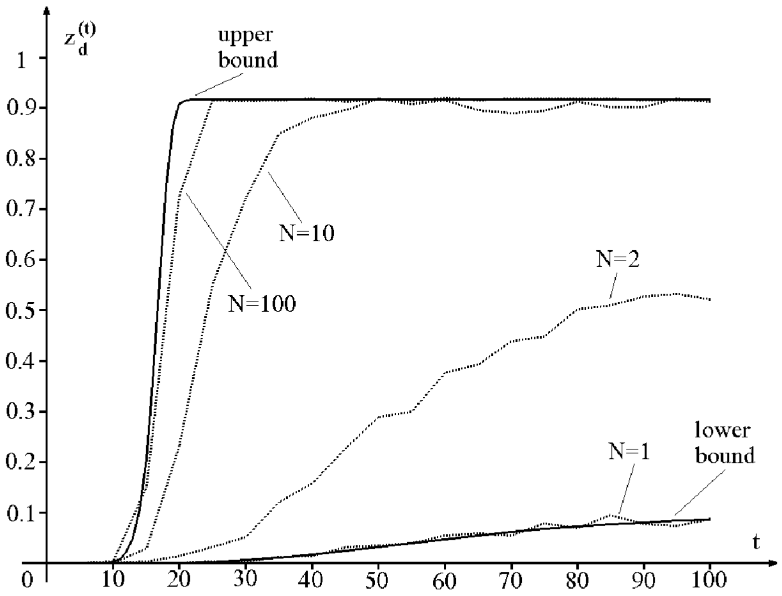

Fig. 1. Average proportion of optimal solutions as a function of the iteration number. $d = 8$, $p_m = 0.01$, $s = 2$.

The computational results are shown in dotted lines. The solid lines correspond to the lower and upper bounds given by (6) and (8). This plot shows that the upper bound (8) provides a good approximation to the value of $z_d^{(t)}$ obtained experimentally, even if the population size is not very large. The rest of the components of $z^{(t)}$ demonstrated a similar tendency.

Another series of experiments was carried out in order to compare the behavior of GAs with different tournament sizes. Figure 2 presents the results for the GA with $p_m = 0.1$, $N = 100$ and $z^{(0)} = 0$ solving the VCP on graph G_6.

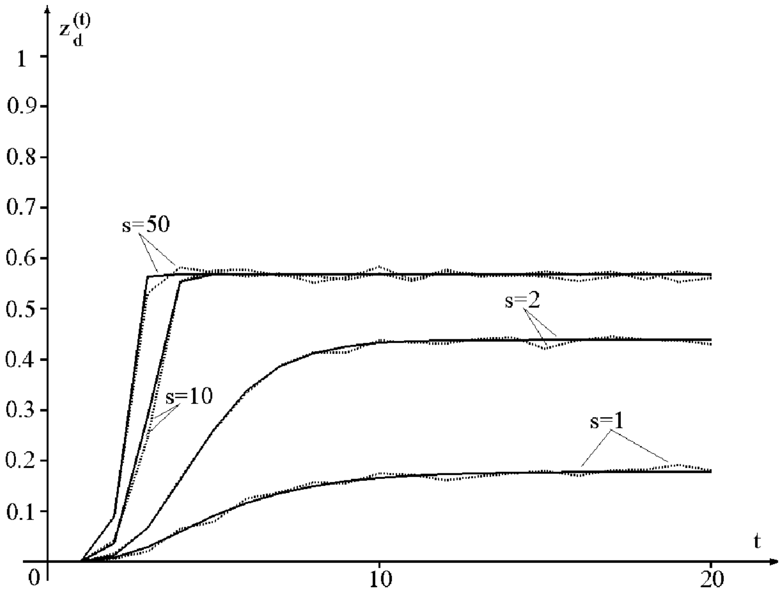

Fig. 2. Average proportion of optimal solutions as a function of the iteration number. $d = 6$, $p_m = 0.1$, $N = 100$.

This plot demonstrates the increase in proportion of the optimal genotypes caused by the extension of the tournament size, which is consistent with Theorem 7 and Corollary 8.

5 Conclusions

In this paper we presented a new model of GA with tournament selection and obtained the upper and lower bounds on proportion of "good" genotypes in population. The tournament selection is frequently used in applications, however theoretically it has not been investigated as much as the proportional selection.

The model may be applied to the GAs with different fitness functions and mutation operators. The adequacy of the bounds depends on the tightness of the monotone bounds on mutation transition probabilities. The analysis of interconnections between the mutation operator, the problem coding, and the goal function is separated from the rest of the GA model. This analysis of mutation ought to provide some coarse graining of the fitness landscape, and the usefulness of the model will depend on the precision of this information.

The bounds obtained provide the estimates for the speed of spreading of "good" genotypes in the population of GA. In particular, it is proved for the GA, which uses a step mutation operator with monotone threshold transition matrix, that the expected components of population vector reach the lower bounds if the population size $N = 1$, and they tend to the upper bounds if $N \to \infty$.

Further research is expected to involve the investigation of the crossover operator and applications of the model to more complex optimization problems.

Acknowledgment
The author would like to thank Prof. Boris A. Rogozin and Sergey A. Klokov for the helpful comments on this paper.

References

1. Bäck, T.:The interaction of mutation rate, selection, and self-adaptation within a genetic algorithm. In: Proc. of Parallel Problem Solving from Nature (PPSN II) Männer, R. and Manderick, B. (eds.). North Holland (1993) 85–94
2. Eremeev, A.V.: A genetic algorithm with a non-binary representation for the set covering problem. In: Proc. of Operations Research (OR'98). Springer-Verlag (1999) 175–181
3. Goldberg, D.E.: Genetic algorithms in search, optimization and machine learning. Addison Wesley Publishing Company (1989)
4. Goldberg, D.E., Korb, B. and Deb, K.: Messy genetic algorithms: motivation, analysis, and first results. Complex Systems, **3** (5) (1989) 493–530
5. Mühlenbein, H.: How genetic algorithms really work: I. Mutation and hillclimbing. In: Proc. of Parallel Problem Solving from Nature (PPSN II) Männer, R. and Manderick, B. (eds.). North Holland (1993) 15–26
6. Mühlenbein, H., Mahnig, T. and Rodriguez, A. O.: Schemata, distributions and graphical models in evolutionary optimization. To Appear in Journal of Heuristics.
7. Mühlenbein, H. and Schlierkamp-Voosen, D.: The science of breeding and its application to the breeder genetic algorithm (BGA). Evolutionary Computation, **1** (1) (1993) 335–360
8. Nix, A. and Vose, M. D.: Modeling genetic algorithms with Markov chains. Annals of Mathematics and Artificial Intelligence, **5** (1992) 79–88
9. Rudolph, G.: Convergence analysis of canonical genetic algorithms. IEEE Transactions on Neural Networks, **5** (1) (1994) 96–101
10. Thierens, D. and Goldberg, D.E.: Convergence models of genetic algorithm selection schemes In: Proc. of Parallel Problem Solving from Nature (PPSN III) Davidor, Y., Schwefel, H.- P. and Männer, R. (eds.). Springer-Verlag (1994) 117–129
11. Vose, M. D.: Modeling simple genetic algorithms. Evolutionary Computation, **3** (4) (1995) 453–472
12. Vose, M. D. and Wright, A.H.: The Walsh transform and the theory of the simple genetic algorithm. Genetic Algorithms for Pattern Recognition. Pal, S. K., Wang, P. P. (eds.) (1995) 25–44
13. Zaozerskaya, L.A.: Investigation and solving of some classes of integer programming problems on the basis of the regular partitions. PhD. thesis, Omsk Branch of Sobolev Institute of Mathematics, SB RAS. (1998)