Takeover Times of Noisy Non-Generational Selection Rules that Undo Extinction

Günter Rudolph [1]

Abstract

The takeover time of some selection method is the expected number of iterations of this selection method until the entire population consists of copies of the best individual under the assumption that the initial population consists of a single copy of the best individual. We consider a class of non-generational selection rules that run the risk of loosing all copies of the best individual with positive probability. Since the notion of a takeover time is meaningless in this case these selection rules are modified in that they undo the last selection operation if the best individual gets extinct from the population. We derive exact results or upper bounds for the takeover time for three commonly used selection rules via a random walk or Markov chain model. The takeover time for each of these three selection rules is $O(n \log n)$ with population size n.

1. Introduction

The notion of the *takeover time* of selection methods used in evolutionary algorithms was introduced by Goldberg and Deb [1]. Suppose that a finite population of size n consists of a single best individual and $n-1$ worse ones. The takeover time of some selection method is the expected number of iterations of the selection method until the entire population consists of copies of the best individual. Evidently, this definition of the takeover time becomes meaningless if all best individuals may get extinct with positive probability. Therefore we study a specific modification of those selection rules: If the all best individual have been erased by erroneous selection then these selection rules undo this extinction by reversing the last selection operation. Here, we concentrate on non-generational selection rules. For such rules Smith and Vavak [2] numerically determined the takeover time or takeover probability based on a Markovian model whereas Rudolph [3] offered a theoretical analysis via the same Markovian model. This work is an extension of [2, 3] as the modified selection rules introduced here have not been considered yet.

Section 2 introduces the particular random walk model, which reflects our assumptions regarding the selection rules, and our standard machinery for determining the takeover time or bounds thereof. Section 3 is of preparatory nature as it contains several auxiliary results required in section 4 in which our standard machinery is engaged to provide the takeover times for our modifications of random replacement selection, noisy binary tournament selection, and "kill tournament" selection. Finally, section 5 relates our findings to results previously obtained for other selection methods.

[1] Department of Computer Science, University of Dortmund, Germany; e-mail: guenter.rudolph@udo.edu

2. Model

Let N_t denote the number of copies of the best individual at step $t \geq 0$. The random sequence $(N_t)_{t \geq 0}$ with values in $S = \{1, 2, \ldots, n\}$ and $N_0 = 1$ is termed a Markov chain if

$$\mathsf{P}\{\, N_{t+1} = j \,|\, N_t = i, N_{t-1} = i_{t-1}, \ldots, N_0 = i_0\} = \mathsf{P}\{\, N_{t+1} = j \,|\, N_t = i\} = p_{ij}$$

for all $t \geq 0$ and for all pairs $(i,j) \in S \times S$. Since we are only interested in non-generational selection rules the associated Markov chains reduce to particular random walks that are amenable to a theoretical analysis. These random walks are characterized by the fact that $|N_t - N_{t+1}| \leq 1$ for all $t \geq 0$ as a non-generational selection rule chooses—somehow—an individual from the population and decides—somehow—which individual should be replaced by the previously chosen one.

Two special classes of random walks were considered in [3] in this context. Here, we need another class reflecting our assumption that the selection rules undo a potential extinction of the best individual by reversing the last selection operation. This leads to a random walk with one reflecting and one absorbing boundary which is a Markov chain with state space $S = \{1, \ldots, n\}$ and transition matrix

$$P = \begin{pmatrix} r_1 & q_1 & 0 & \cdots & & & 0 \\ p_2 & r_2 & q_2 & 0 & \cdots & & 0 \\ 0 & p_3 & r_3 & q_3 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & p_{n-2} & r_{n-2} & q_{n-2} & 0 \\ 0 & \cdots & & 0 & p_{n-1} & r_{n-1} & q_{n-1} \\ 0 & \cdots & & & 0 & 0 & 1 \end{pmatrix}$$

with $p_i, q_i > 0$, $r_i \ge 0$, $p_i + r_i + q_i = 1$ for $i = 2, \ldots, n-1$ and $r_1 = 1 - q_1 \in (0,1)$. Notice that state n is the only absorbing state. The expected absorption time is $\mathsf{E}[\, T \mid N_0 = k\,]$ with $T = \min\{t \ge 0 : N_t = n\}$ and it can be determined as follows [4]. Let matrix Q result from matrix P by deleting its last row and column. If C is the inverse of matrix $A = I - Q$ with unit matrix I, then $\mathsf{E}[\, T \mid N_0 = k\,] = c_{k1} + c_{k2} + \cdots + c_{k,n-1}$ for $1 \le k < n$. Since $N_0 = 1$ in the scenario considered here, we only need the first row of matrix $C = A^{-1}$ which may be obtained via the adjugate of matrix A. This avenue was followed in Rudolph [5] for a more general situation. Using the result obtained in [5] (by setting $p_1 = 0$) we immediately get

$$c_{1j} = \frac{\displaystyle\sum_{k=0}^{n-j-1} \left(\prod_{u=j+1}^{n-k-1} p_u\right)\left(\prod_{v=n-k}^{n-1} q_u\right)}{\displaystyle\prod_{k=j}^{n-1} q_k} \tag{1}$$

for $1 \le j \le n-1$. Thus, the plan is as follows: First, derive the transition probabilities for a non-generational selection rule that fulfills our assumptions. This is usually easy. Next, these expressions are fed into equation (1) yielding c_{1j}. The result may be a complicated formula; in this case it will be bounded in an appropriate manner. Finally, we determine the sum

$$\mathsf{E}[\, T \mid N_0 = 1\,] = \sum_{j=1}^{n-1} c_{1j}$$

and we are done. For the sake of notational convenience we shall omit the conditioning $\{N_0 = 1\}$ and write simply $\mathsf{E}[\,T\,]$ for the expected takeover time.

3. Mathematical Prelude

In case of positive integers the Gamma function $\Gamma(\cdot)$ obeys the relationships $n\,\Gamma(n) = \Gamma(n+1) = n!$. For later purposes we need the following results:

Lemma 1 For $n \in \mathbb{N}$,

$$\sum_{k=0}^{n-1} \frac{\Gamma(n+k+1)}{\Gamma(k+1)} = \frac{\Gamma(2\,n+1)}{(n+1)\,\Gamma(n)}\,.$$

Proof: See [3], p. 905. ■

Lemma 2 Let $n \ge 2$ and $1 \le j \le n-1$. Then

$$S(n,j) = \frac{n^2\,\Gamma(n-j)\,\Gamma(n+j)}{\Gamma(j+1)\,\Gamma(2\,n-j+1)} \sum_{k=0}^{n-j-1} d_k \le \frac{1}{2} + \frac{1}{4\,n}$$

where

$$d_k = \frac{\Gamma(n+k+1)\,\Gamma(n-k)}{\Gamma(2\,n-k)\,\Gamma(k+1)}\,.$$

Proof: Due to lack of space we only offer a sketch of the proof. First show that $S(n,0) \ge S(n,j)$ for $j = 1, \ldots, n-1$ and $n \ge 2$. Since the bound

$$2\,S(n,0) = n\,\frac{\Gamma(n)^2}{\Gamma(2\,n)} \sum_{k=0}^{n-1} d_k \le 1 + \frac{1}{2\,n}$$

follows from [3], pp. 907-908, division by 2 yields the result desired. ■

Moreover, the nth harmonic number H_n can be bracketed by

$$\log n < H_n = \sum_{i=1}^{n} \frac{1}{i} < \log n + 1$$

for $n \ge 2$ and notice that

$$\sum_{i=0}^{n} a^{n-i}\, b^i = \frac{a^{n+1} - b^{n+1}}{a - b}$$

for $a \ne b$. Finally some notation: The set I_m^n denotes all integers between m and n (inclusive).

4. Analysis

4.1. Random Replacement Selection

Two individuals are drawn at random and the better one of the pair replaces a randomly chosen individual from the population. If the last best individual was erased by chance then the last selection operation is reversed. As a consequence, the transition probabilities of the associated Markov chain are $p_{nn} = 1$, $p_{11} = 1 - p_{12}$,

$$\begin{aligned} \forall i \in I_1^{n-1} : \; p_{i,i+1} &= \frac{i}{n}\left(2 - \frac{i}{n}\right)\left(1 - \frac{i}{n}\right) \\ \forall i \in I_2^{n-1} : \; p_{i,i-1} &= \left(1 - \frac{i}{n}\right)^2 \frac{i}{n} \end{aligned}$$

and $p_{ii} = 1 - p_{i,i-1} - p_{i,i+1}$. Since $p_i = p_{i,i-1}$, $q_i = p_{i,i+1}$ and

$$\begin{aligned} \prod_{v=n-k}^{n-1} q_v &= \frac{1}{n^{3k+1}} \frac{\Gamma(n+k+1)\,\Gamma(k+1)}{\Gamma(n-k)} \qquad (2) \\ \prod_{u=j+1}^{n-k-1} p_u &= \frac{1}{n^{3\,(n-k-j-1)}} \frac{\Gamma(n-j)^2}{\Gamma(j+1)} \frac{\Gamma(n-k)}{\Gamma(k+1)^2} \end{aligned}$$

one obtains

$$\sum_{k=0}^{n-j-1} \left(\prod_{u=j+1}^{n-k-1} p_u\right)\left(\prod_{v=n-k}^{n-1} q_u\right) =$$

$$\frac{1}{n^{3\,(n-j-1)+1}} \frac{\Gamma(n-j)^2}{\Gamma(j+1)} \sum_{k=0}^{n-j-1} \frac{\Gamma(n+k+1)}{\Gamma(k+1)} =$$

$$\frac{1}{n^{3\,(n-j-1)+1}}\,\frac{\Gamma(n-j)}{\Gamma(j+1)}\,\frac{\Gamma(2\,n-j+1)}{n+1} \qquad (3)$$

with the help of Lemma 1. Insertion of $k = n-j$ in equation (2) leads to

$$\prod_{v=j}^{n-1} q_v = \frac{\Gamma(2\,n-j+1)\,\Gamma(n-j+1)}{n^{3\,(n-j)+1}\,\Gamma(j)} \,. \qquad (4)$$

After insertion of equations (3) and (4) in equation (1) we have

$$c_{1j} = \frac{n^3}{n+1}\cdot\frac{1}{j}\cdot\frac{1}{n-j} = \frac{n^2}{n+1}\left(\frac{1}{j}+\frac{1}{n-j}\right)$$

and finally

$$\mathsf{E}[\,T\,] = \sum_{j=1}^{n-1} c_{1j} = \frac{2\,n^2}{n+1}\,H_{n-1}\,.$$

4.2. Noisy Binary Tournament Selection

Two individuals are drawn at random and the best as well as worst member of this sample is identified. The worst member replaces the best one with some replacement error probability $\alpha \in (0, \frac{1}{2})$, whereas the worst one is replaced by the best one with probability $1-\alpha$. Again, if the last best copy has been discarded then the last selection operation is reversed. Therefore the transition probabilities are as follows: $p_{nn} = 1$, $p_{12} = s_1\,(1-\alpha)$, $p_{11} = 1 - p_{12}$ and $p_{i,i+1} = s_i\,(1-\alpha)$, $p_{i,i-1} = s_i\,\alpha$, $p_{ii} = 1 - s_i$ for $i = 2,\ldots,n-1$. Here, s_i denotes the probability that the sample of two individuals contains at least one best as well as one worse individual from a population with $i = 1,\ldots,n-1$ copies of the best individual, i.e.,

$$s_i = 1 - \left(\frac{i}{n}\right)^2 - \left(1-\frac{i}{n}\right)^2 = 2\,\frac{i}{n}\left(1-\frac{i}{n}\right).$$

According to equation (1) we need

$$\begin{aligned}
\prod_{v=n-k}^{n-1} q_v &= \left(\frac{2\,(1-\alpha)}{n^2}\right)^k \frac{\Gamma(n)\,\Gamma(k+1)}{\Gamma(n-k)} \\
\prod_{u=j+1}^{n-k-1} p_u &= \left(\frac{2\,\alpha}{n^2}\right)^{2(n-j-1-k)} \frac{\Gamma(n-k)\Gamma(n-j)}{\Gamma(k+1)\Gamma(j+1)}
\end{aligned}$$

leading to

$$\sum_{k=0}^{n-j-1}\left(\prod_{u=j+1}^{n-k-1} p_u\right)\left(\prod_{v=n-k}^{n-1} q_u\right) =$$

$$\frac{2^{n-j-1}}{n^{2\,(n-j-1)}}\cdot\frac{\Gamma(n)\,\Gamma(n-j)}{\Gamma(j+1)}\sum_{k=0}^{n-j-1}\alpha^{n-j-1-k}(1-\alpha)^k =$$

$$\frac{2^{n-j-1}}{n^{2\,(n-j-1)}}\cdot\frac{\Gamma(n)\,\Gamma(n-j)}{\Gamma(j+1)}\cdot\frac{(1-\alpha)^{n-j}-\alpha^{n-j}}{1-2\,\alpha}\,. \qquad (5)$$

Since

$$\prod_{v=j}^{n-1} q_v = \left(\frac{2\,(1-\alpha)}{n^2}\right)^{n-j}\cdot\frac{\Gamma(n)\,\Gamma(n-j+1)}{\Gamma(j)} \qquad (6)$$

we get by inserting equations (5) and (6) into equation (1)

$$\begin{aligned}
c_{1j} &= \frac{n^2}{2}\,\frac{\Gamma(j)}{\Gamma(j+1)}\cdot\frac{\Gamma(n-j)}{\Gamma(n-j+1)}\cdot\frac{(1-\alpha)^{n-j}-\alpha^{n-j}}{(1-\alpha)^{n-j}(1-2\,\alpha)} \\
&= \frac{n^2}{2}\cdot\frac{1}{j}\cdot\frac{1}{n-j}\cdot\frac{1}{1-2\,\alpha}\cdot(1-r^{n-j})
\end{aligned}$$

where $r = \alpha/(1-\alpha)$ and finally

$$\mathsf{E}[\,T\,] = \sum_{j=1}^{n-1} c_{1j} =$$

$$\frac{n^2}{2}\cdot\frac{1}{1-2\,\alpha}\sum_{j=1}^{n-1}\frac{1}{j}\cdot\frac{1}{n-j}\cdot\left[1-\left(\frac{\alpha}{1-\alpha}\right)^{n-j}\right] =$$

$$\frac{n}{2\,(1-2\,\alpha)}\sum_{j=1}^{n-1}\left[\frac{1}{j}+\frac{1}{n-j}\right]\cdot\left[1-\left(\frac{\alpha}{1-\alpha}\right)^{n-j}\right] =$$

$$\frac{n}{1-2\,\alpha}\left[H_{n-1}-\frac{1}{2}\sum_{j=1}^{n-1}\frac{r^{n-j}}{j}-\frac{1}{2}\sum_{j=1}^{n-1}\frac{r^{n-j}}{n-j}\right] \le$$

$$\frac{n}{1-2\,\alpha}\left[H_{n-1}-\frac{1}{n}\sum_{j=1}^{n-1} r^j\right] =$$

$$\frac{n}{1-2\,\alpha}\left[H_{n-1}-\frac{1}{n}\cdot\frac{r-r^n}{1-r}\right] \le \frac{n\,H_{n-1}}{1-2\,\alpha}$$

where $r = \alpha/(1-\alpha) \in (0,1)$. The above bound is very accurate if α is not too close to $1/2$. For example, for

$$\alpha = \frac{1}{2} - \frac{1}{2\,n^k}$$

this bound yields $\mathsf{E}[\,T\,] \le n^{k+1}\,H_{n-1}$ for $k \ge 0$ whereas the worst case ($\alpha = 1/2$) reveals[1] that $\mathsf{E}[\,T\,] \le n^2\,H_{n-1}$ for $\alpha \in [0, 1/2]$. Moreover, notice that we get $\mathsf{E}[\,T\,] = n\,H_{n-1}$ in the best case ($\alpha = 0$) [3].

4.3. "Kill Tournament" Selection

This selection method proposed in [2] is based on two binary tournaments: In the first tournament the best individual is identified. This individual replaces the worst individual identified in the second tournament (the "kill tournament"). If the last best copy gets lost then the last selection operation is

[1] If $\alpha = 1/2$ then the entire derivation collapses to simple expressions leading to $\mathsf{E}[\,T\,] = n^2\,H_{n-1}$.

reversed. The transition probabilities are $p_{nn} = 1$, $p_{11} = 1 - p_{12}$,

$$\begin{aligned}\forall i \in I_1^{n-1} : \; p_{i,i+1} &= \frac{i}{n}\left(2 - \frac{i}{n}\right)\left[1 - \left(\frac{i}{n}\right)^2\right] \\ \forall i \in I_2^{n-1} : \; p_{i,i-1} &= \left(1 - \frac{i}{n}\right)^2 \left(\frac{i}{n}\right)^2\end{aligned}$$

and $p_{ii} = 1 - p_{i,i-1} - p_{i,i+1}$. As in the previous cases we require the products

$$\begin{aligned}\prod_{v=n-k}^{n-1} q_v &= \frac{\Gamma(n+k+1)\,\Gamma(k+1)\,\Gamma(2\,n)}{n^{4\,k+1}\Gamma(n-k)\,\Gamma(2\,n-k)} \\ \prod_{u=j+1}^{n-k-1} p_u &= \frac{1}{n^{4\,(n-j-1-k)}}\,\frac{\Gamma(n-k)^2\,\Gamma(n-j)^2}{\Gamma(j+1)^2\,\Gamma(k+1)^2}\end{aligned}$$

to evaluate the sum

$$\sum_{k=0}^{n-j-1}\left(\prod_{u=j+1}^{n-k-1} p_u\right)\left(\prod_{v=n-k}^{n-1} q_u\right) =$$

$$\frac{\Gamma(n-j)^2\,\Gamma(2\,n)}{n^{4\,(n-j-1)+1}\,\Gamma(j+1)^2}\sum_{k=0}^{n-j-1} d_k \tag{7}$$

with

$$d_k = \frac{\Gamma(n+k+1)\,\Gamma(n-k)}{\Gamma(2\,n-k)\,\Gamma(k+1)}\,.$$

Since

$$\prod_{v=j}^{n-1} q_v = \frac{\Gamma(2\,n-j+1)\,\Gamma(n-j+1)\,\Gamma(2\,n)}{n^{4\,(n-j)+1}\,\Gamma(j)\,\Gamma(n+j)} \tag{8}$$

insertion of (7) and (8) in (1) yields

$$\begin{aligned}c_{1j} &= \frac{n^4}{j\,(n-j)}\,\frac{\Gamma(n-j)\,\Gamma(n+j)}{\Gamma(j+1)\,\Gamma(2\,n-j+1)}\sum_{k=0}^{n-j-1} d_k \\ &\le \frac{n^2}{j\,(n-j)}\cdot\left(\frac{1}{2}+\frac{1}{4\,n}\right) \qquad (9) \\ &= \left(\frac{1}{j}+\frac{1}{n-j}\right)\left(\frac{n}{2}+\frac{1}{4}\right)\end{aligned}$$

and finally

$$\mathsf{E}[\,T\,] = \sum_{j=1}^{n-1} c_{1j} \le \left(n + \frac{1}{2}\right) H_{n-1}\,.$$

Here, the inequality in (9) follows from Lemma 2.

5. Summary

Now we are in the position to compare the takeover times of the selection methods considered here with those examined in [3]. Table 1 offers an overview of the takeover times of replace worst selection (RW), quaternary (QT), ternary (TT) and binary (BT) tournament selection with $\alpha = 0$ [3], and kill tournament "with undoing" (KT_u), random replacement selection "with undoing" (RR_u) and noisy binary tournament selection "with undoing" and replacement error α ($BT_u(\alpha)$). For fixed $\alpha < 1/2$ the takeover times of all non-generational selection rules considered here and in [3] are of order $O(n \log n)$. Consequently, it does not matter which selection rule is used, provided that the takeover time is actually a key figure of the selection pressure.

selection method	takeover time
QT	$\le \frac{1}{2}\,n\,H_{n-1}$
RW	$\le \frac{1}{2}\,n\,H_{2\,n-1}$
TT	$\frac{2}{3}\,n\,H_{n-1}$
BT	$n\,H_{n-1}$
KT_u	$\le (n+\frac{1}{2})\,H_{n-1}$
$BT_u(\frac{1}{4})$	$\le 2\,n\,H_{n-1}$
RR_u	$\frac{2\,n^2}{n+1}\,H_{n-1}$
$BT_u(\frac{1}{2})$	$n^2\,H_{n-1}$

Table 1: Survey of takeover times.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the *Collaborative Research Center "Computational Intelligence"* (`http://sfbci.uni-dortmund.de`).

References

[1] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms* (G. J. E. Rawlins, ed.), pp. 69–93, San Mateo (CA): Morgan Kaufmann, 1991.

[2] J. Smith and F. Vavak, "Replacement strategies in steady state genetic algorithms: Static environments," in *Foundations of Genetic Algorithms 5* (W. Banzhaf and C. Reeves, eds.), pp. 219–233, San Francisco (CA): Morgan Kaufmann, 1999.

[3] G. Rudolph, "Takeover times and probabilities of non-generational selection rules," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2000)* (D. Whitley *et al.*, eds.), pp. 903–910, San Fransisco (CA): Morgan Kaufmann, 2000.

[4] M. Iosifescu, *Finite Markov Processes and Their Applications*. Chichester: Wiley, 1980.

[5] G. Rudolph, "The fundamental matrix of the general random walk with absorbing boundaries," Technical Report CI-75 of the Collaborative Research Center "Computational Intelligence", University of Dortmund, October 1999.