ELSEVIER

Artificial Intelligence 140 (2002) 245–248

Artificial Intelligence

www.elsevier.com/locate/artint

Corrigendum

Erratum to: Drift analysis and average time complexity of evolutionary algorithms [Artificial Intelligence 127 (2001) 57–85] ☆

Jun He, Xin Yao *

School of Computer Science, The University of Birmingham, Birmingham B15 2TT, UK

Abstract

The proof of Theorem 6 in the paper by J. He and X. Yao [Artificial Intelligence 127 (1) (2001) 57–85] contains a mistake, although the theorem is correct [S. Droste et al., Theoret. Comput. Sci. 276 (2002) 51–81]. This note gives a revised proof and theorem. It turns out that the revised theorem is more general than the original one given an evolutionary algorithm with mutation probability $p_m = 1/(2n)$, using the same proof method as given by J. He and X. Yao [Artificial Intelligence 127 (1) (2001) 57–85].

1. Introduction

Wegener [3] pointed out that, in the proof of Theorem 6 in the paper [1, p. 69], the sum

$$1 + \frac{-1}{2!} + \frac{-2}{3!} + \cdots$$

should be 0, i.e., $c = 0$ rather than $c > 0$ as indicated in the proof [1, p. 69]. As a result, the method used to establish Theorem 6 cannot be used to prove the upper bound $\mathrm{O}(n \log(n))$, although the result given by the theorem is correct and was first established by Droste et al. using a different proof method [2].

Using the same proof method as we used previously, this note gives a revised proof and theorem for Theorem 6 in [1]. The only difference between the revised theorem and the

☆ PII of original article: S0004-3702(01)00058-3.

* Corresponding author.
E-mail address: x.yao@cs.bham.ac.uk (X. Yao).

0004-3702/02/$ – see front matter
PII: S0004-3702(02)00260-6

original one [1] (Theorem 6) is that the evolutionary algorithm (EA) used will now have mutation probability $p_m = 1/(2n)$, rather than $p_m = 1/n$. Interestingly, such a revision has led to a theorem that is more general than the original one. The new theorem presented in this note holds not only for the linear function, but also other unimodal functions as considered in [2].

2. The main result

A fitness function $f:\{0,1\}^n \to \mathbb{R}$ can be written as a polynomial [2]:

$$f(s_1 \dots s_n) = \sum_{I \subseteq \{1,\dots,n\}} c_f(I) \prod_{i \in I} s_i, \tag{1}$$

where coefficient $c_f(I) \in \mathbb{R}$.

Consider the following class of fitness functions: $f(s_1 \dots s_n)$ satisfies, for any $k = 1, \dots, n$ and fixed $s_1, s_{k-1}, s_{k+1}, \dots, s_n$,

$$f(s_1 \dots s_{k-1} 0 s_{k+1} \dots s_n) < f(s_1 \dots s_{k-1} 1 s_{k+1} \dots s_n). \tag{2}$$

In other words, if one "0" bit at any position flips into "1", the fitness will increase. $(1 \dots 1)$ is the unique maximum point. This is a class of unimodal functions, which includes the linear function, $f(x) = c_0 + \sum_{i=1}^{n} c_i s_i$, if all coefficients c_i are positives.

The $(1+1)$ EA is considered here (and $(2N + 2N)$ EAs without crossover can be analysed in a similar way). The mutation is characterised by a mutation rate $p_m > 0$ which specifies the probability of flipping each bit in a chromosome. The selection is to replace the parent if the offspring is not worse than it.

Theorem 1. *For any fitness function* (1) *satisfying* (2)*, the EA with mutation probability* $p_m = 1/(2n)$ *needs average* $\mathrm{O}(n \log n)$ *steps to reach the optimal solution.*

Proof. Define the distance function $d(x) = \sum_{i=1}^{n}(1 - s_i)$. Since $0 \leqslant d(x) \leqslant n$, we can divide $[0, d_{\max}]$ into $n+1$ intervals $d_0 < d_1 < d_2 < \cdots < d_n$ where $d_l = l$. We will use Theorem 3 in [1] to prove the result.

Assume at time $k \geqslant 0$, population ξ_k satisfies $d(\xi_k) > d_{l-1}$, where $l \in \{1, \dots, n\}$. Without the loss of generality, assume $d(\xi_k) = d_l$ (other cases can be proven in the same way), which implies that there are l "0" bits in x (where x is the best individual in ξ_k). Then

$$\begin{aligned} E\big[d(\xi_k) - d(\xi_{k+1})\big] = {} & E\big[\big(d(\xi_k) - d(\xi_{k+1})\big) I\big\{d(\xi_k) > d(\xi_{k+1})\big\}\big] \\ & + E\big[\big(d(\xi_k) - d(\xi_{k+1})\big) I\big\{d(\xi_k) < d(\xi_{k+1})\big\}\big]. \end{aligned}$$

First let's consider $E[(d(\xi_k) - d(\xi_{k+1})) I\{d(\xi_k) > d(\xi_{k+1})\} \mid d(\xi_k) = d_l]$. Let $\xi_k = x$. The probability of flipping one of l "0" bits in x while keeping its $n - l$ "1" bits unchanged is $C_l^1 \frac{1}{2n}(1 - \frac{1}{2n})^{n-l}$. Hence

$$\begin{aligned} & E\big[\big(d(\xi_k) - d(\xi_{k+1})\big) I\big\{d(\xi_k) > d(\xi_{k+1})\big\} \mid d(\xi_k) = d_l\big] \\ & \quad \geqslant C_l^1 \frac{1}{2n}\left(1 - \frac{1}{2n}\right)^{n-l} \geqslant \frac{l}{2n}\left(1 - \frac{1}{2n}\right)^{n-1}. \end{aligned}$$

Secondly, let's consider $E[(d(\xi_k) - d(\xi_{k+1}))I\{d(\xi_k) < d(\xi_{k+1})\} \mid d(\xi_k) = d_l]$. Let $\xi_k = x$. If event $I\{d(\xi_k) < d(\xi_{k+1})\}$ happened, then the following event I' must happen: at least one "0" bit in x must flip and if m "0" bits flip to 1, then at least $m+1$ "1" bits must flip to "0". So the probability of event $I\{d(\xi_k) < d(\xi_{k+1})\}$ happening is not more than that of event I'. The later event (I') can be divided into the following sub-events:

(1) m "0" bits ($1 \leqslant m \leqslant \min\{l, n-l-1\}$) in x become 1, and $m+1$ "1" bits become 0. The probability of this event happening is

$$\begin{aligned} & C_l^1\left(\frac{1}{2n}\right)C_{n-l}^2\left(\frac{1}{2n}\right)^2\left(1-\frac{1}{2n}\right)^{n-3} \\ &\quad + C_l^2\left(\frac{1}{2n}\right)^2 C_{n-l}^3\left(\frac{1}{2n}\right)^3\left(1-\frac{1}{2n}\right)^{n-5} + \cdots \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(\frac{2^{-2}}{2!}+\frac{2^{-3}}{3!}+\cdots\right) \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\frac{2^{-1}}{2!}. \end{aligned}$$

(2) m "0" bits ($1 \leqslant m \leqslant \min\{l, n-l-2\}$) in x become 1, and $m+2$ "1" bits become 0. The probability of this event happening is

$$\begin{aligned} & C_l^1\left(\frac{1}{2n}\right)C_{n-l}^3\left(\frac{1}{2n}\right)^3\left(1-\frac{1}{2n}\right)^{n-4} \\ &\quad + C_l^2\left(\frac{1}{2n}\right)^2 C_{n-l}^4\left(\frac{1}{2n}\right)^4\left(1-\frac{1}{2n}\right)^{n-6} + \cdots \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(\frac{2^{-3}}{3!}+\frac{2^{-4}}{4!}+\cdots\right) \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\frac{2^{-2}}{3!}. \end{aligned}$$

(3) m "0" bits ($1 \leqslant m \leqslant \min\{l, n-l-3\}$) become 1, and $m+3$ "1" bits become 0. The probability of this event happening is

$$\begin{aligned} & C_l^1\left(\frac{1}{2n}\right)C_{n-l}^4\left(\frac{1}{2n}\right)^4\left(1-\frac{1}{2n}\right)^{n-5} \\ &\quad + C_l^2\left(\frac{1}{2n}\right)^2 C_{n-l}^5\left(\frac{1}{2n}\right)^5\left(1-\frac{1}{2n}\right)^{n-7} + \cdots \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(\frac{2^{-4}}{4!}+\frac{2^{-5}}{5!}+\cdots\right) \\ & \leqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\frac{2^{-3}}{4!}. \end{aligned}$$

(4) And so on ....

Hence we get

$$E\big[\big(d(\xi_k) - d(\xi_{k+1})\big) I\big\{d(\xi_k) < d(\xi_{k+1})\big\} \mid d(\xi_k) = d_l\big]$$
$$\geqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(-1\frac{2^{-1}}{2!} - 2\frac{2^{-2}}{3!} - 3\frac{2^{-3}}{4!} - \cdots\right)$$
$$\geqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(-\frac{1}{2}\right).$$

So

$$E\big[d(\xi_k) - d(\xi_{k+1}) \mid d(\xi_k) = d_l\big] \geqslant \frac{l}{2n}\left(1-\frac{1}{2n}\right)^{n-1}\left(1-\frac{1}{2}\right) \simeq c\frac{l}{n},$$

where $c = \mathrm{e}^{-1/2}/4 > 0$ is a constant. In other words, condition 4 in [1] holds.

According to Theorem 3 in [1],

$$E[\tau] \leqslant c^{-1}\sum_{l=1}^{n}\frac{n}{l} = \mathrm{O}(n\log n). \qquad \square$$

3. Conclusion

The proof of Theorem 6 in [1] contains an error, although the theorem is correct. This note has revised the proof and removed the error. The revised theorem and its proof are presented in this note. As a result of such revision, a more general result has been obtained. Not only is the new theorem applicable to linear functions, as it was the case in [1], it is also true for more generic unimodal functions. It is worth mentioning that the proof method used to show the new result is the same as that used in [1], which is much simpler than other proof methods.

Acknowledgements

The authors are grateful to Professor Ingo Wegener for drawing our attention to the error in the proof of Theorem 6 in [1].

References

[1] J. He, X. Yao, Drift analysis and average time complexity of evolutionary algorithms, Artificial Intelligence 127 (1) (2001) 57–85.
[2] S. Droste, T. Jansen, I. Wegener, On the analysis of the $(1+1)$ evolutionary algorithms, Theoret. Comput. Sci. 276 (2002) 51–81.
[3] I. Wegener, Private communication, February 2002.