Convergence of a Recombination-Based Elitist Evolutionary Algorithm on the Royal Roads Test Function

Aram Ter-Sarkisov and Stephen Marsland

Department of Computer Science
Massey University, New Zealand
{a.ter-sarkisov,s.r.marsland}@massey.ac.nz
http://www.massey.ac.nz/seat

Abstract. We present analysis of performance of an elitist Evolutionary algorithm using a recombination operator 1-Bit-Swap on the Royal Roads test function. We derive complete, approximate and asymptotic convergence rates. Both complete and approximate models show the benefit of the size of the population and recombination pool when they are small and leveling out of this effect when limit conditions are applied. Numerical results confirm our findings.

Keywords: Evolutionary algorithms, computational complexity, probabilistic models.

1 Introduction

Evolutionary Algorithms (EA) are a set of heuristic optimization tools, that are well-suited to solve problems with poorly-understood landscape (black-box optimization). Despite rich history in application, theoretical analysis has been lagging behind. Although in the past few years a large amount of research in this area has evolved, it is mostly restricted to single-parent algorithms.

We analyze an elitist (hence +) $(\mu+\lambda)$ algorithm that operates on a population of solutions size μ and recombination pool size λ using a genetic operator called 1-Bit-Swap (1BS) and tournament selection function. Each generation a subset of best species α is saved (hence elitism).

We find three expressions for the expected runtime of $(\mu+\lambda)\text{EA}_{1BS}$: one exact, one approximate and the third one asymptotic. Asymptotic expression does not contain the variables for the size of the population and recombination pool due to cancellation, which is the leveling out effect: as the size of the population grows large enough, its effect relaxes. Approximate and asymptotic expressions are necessary, since the complete one doesn't seem to exist in the closed form.

An important idea, on which the derivations in this article are based, is the distribution of elite species in the population, α, which is assumed Uniform. It does not seem to be the case, that this approach was used in EA literature before.

D. Wang and M. Reynolds (Eds.): AI 2011, LNAI 7106, pp. 361–371, 2011.

1.1 Royal Roads Function

Royal Roads (RR) is a test function introduced in [1] and analyzed in [2], where a population-based Evolutionary Algorithm (EA) was found to have underperformed a simpler heuristic Randomized Local Search (RLS), which contradicted the theoretical findings in the same article. An interesting feature of RR is a plateau of fitness (i.e. same fitness for a large subset of genotype).

The notation we use in this article is different from the one in [1, 2]: a string length n is broken down into K consecutive bins numbered $\kappa_1, \kappa_2, \ldots, \kappa_K$ and the size/length of each bin is M (therefore $n = KM$). Fitness of each bin is equal to 0 if at least 1 bit is 0 and M if all bits are equal to 1.

Therefore, fitness of the string/parent is $0, M, 2M...n$. Additional notation is presented in Section 3.

1.2 Past Work

Unlike OneMax (Counting Ones), RR has seen less attention in EA literature, though in [3, 4] a variant of RR was analyzed and the upper bound of $O(n^6)$ was found for a version of RR in [4]. In [2] the bounds on convergence for RR were found to be $O(2^K \log N)$, N being the length of the string, K the length of schemata (length of the bin), up to a linear term tighter than for RLS ($O(2^K N \log N)$), although numerically RLS outperformed EA. Besides, this result is somewhat vague, since it doesn't involve population or recombination pool size in any way.

Most research in EA literature is focused on mutation-based single-species algorithms solving pseudo-boolean type functions, that includes OneMax and RR with some very sharp bounds derived for OneMax problem (e.g. $0.982n \log n$ in [5]). Recently a number of recombination-based algorithms were analyzed in [6, 7] for some cases where crossover can be provably effective.

Research on population-based algorithms (including EA) is more numerous, including that on OneMax test function, although $(1+1)$, $(\mu+1)$ and $(1+\lambda)$ setups are still more widespread. In [8] it was proven that the effect of population is problem-specific, i.e., increase in population size may not improve performance at all. Very recently, in [9], it was shown that population size $O(\log n)$ boosts performance and size $\Omega(\frac{n}{\log n})$ impairs the progress of the algorithm (on TrapZeros multimodal function) and reduces the probability of global convergence.

2 Analyzed Algorithm: $(\mu + \lambda)$EA$_{1BS}$

k-Bit-Swap genetic operator (KBS) was introduced in [10]. It contains some features of both mutation and uniform crossover and recombines information between two parents in a random manner. In this article we use 1-Bit-Swap (1BS), which picks exactly 1 bit from each parent uniformly at random.

Pseudocode for the analyzed algorithm is presented in Table 1 and is very simple both in outline and implementation.

Table 1. $(\mu + \lambda)\mathrm{EA}_{1BS}$

1	Initialize population size μ
	repeat for t generations:
2	select $\frac{\lambda}{2}$ pairs of parents from the population using Tournament selection
	repeat $\frac{\lambda}{2}$ times:
3a	select a bit at random in Parent 1
3b	select a bit at random in Parent 2
3c	swap values in the selected bits
4	after the recombination, keep α best species in the population,
	replace the rest with the best species from the pool

We use Tournament selection detailed below because it is fairly straightforward both in implementation and analysis.

- Select two species x_i, x_j uniformly at random
- if $f(x_i) = f(x_j)$, either x_i or x_j enters the pool at random
- else the species with better fitness enters the pool.

3 Model Setup and Assumptions

The main quantity we analyze in this article is the first hitting time of the global solution of the test problem:

$$\tau_A^{RR} = \min\{t \geq 0 : f(\alpha) = n\}$$

where A is the set of all possible populations that include a global solution. We want to find $\mathbf{E}\tau^{RR}_{(\mu+\lambda)EA_{1BS}}$, the expectation of this time parameter.

3.1 Improvement Process

We assume that each bin κ starts with an equal number of 0's and 1's, which implies that the starting fitness of the population is 0. To measure the progress of the algorithm we introduce, in addition to the fitness function, the auxiliary function OneMax (or counting Ones, for further reference see e.g. [11]) that we denote V_κ. Due to the nature of 1BS bins evolve in a sequence, i.e. two different bins cannot evolve at the same generation, therefore κ can be viewed as the 'active bin'.

We also assume that the starting auxiliary function of each bin, $\min V_\kappa = \frac{M}{2}$.

The successful event G is defined as evolution of at least one more elite species in the population. To avoid confusion, the number of bits equal to 0 left to swap in a bin we use l (they are numbered 0 through $\frac{M}{2} - 1$), and the bins left to fill we use κ numbered 0 through $K - 1$. We restrict our attention to elite pairs in the recombination pool, i.e. pairs in which both parents are currently-elite species. Following the process described in greater details in Section 2, the probability to select such a pair into a recombination pool is

$$P_{\mathrm{sel}}(\alpha) = \frac{\alpha^2(\alpha + 2(\mu - \alpha))^2}{\mu^4} = \frac{(\alpha(2\mu - \alpha))^2}{\mu^4}$$

Having selected the pair, the probability that as a results of swapping bits between them, a better species evolves is

$$P_{\text{swap}} = \frac{2(\frac{M}{2} - l)(\frac{n}{2} + \frac{\kappa M}{2} + l)}{n^2} = \frac{(M - 2l)(n + \kappa M + 2l)}{2n^2}$$

This probability comes from the fact that we want to select any 0 in bin κ in one of the parents and 1 anywhere in the other parent. Obviously as the number of 1's in both parents keeps growing, this probability grows too. In Section 4.1 we also extensively use the probability of failure:

$$P_{\text{F}} = 1 - P_{\text{swap}}$$

3.2 Population and Elitism Assumptions

This is a very important part of the paper. We assume that each generation currently-elite species in the population are distributed uniformly:

$$\alpha \sim Uniform\Big(\frac{1}{\mu}\Big)$$

This is a static model, i.e. this distribution does not change throughout the run of the algorithm. We also assume that the rate of elitism (number of species saved for the next generation) is *high enough*, that is, high enough to keep all elite species. We expect this result to yield a type of a lower bound, because this probability distribution assigns relatively high values to very high sizes of elite species. Say, in a real run the probability of have μ elite species in the population is much lower than $\frac{1}{\mu}$.

4 Derivation of the Expectation of Convergence Time

We present three main results: exact, approximate and asymptotic. The latter two are necessary, since, the complete one doe s not exist in he closed form.

4.1 Exact Expression

We start with introducing the probability of failure to improve V_κ (see also Appendix B):

$$P(G_0) = \sum_{j=0}^{\frac{\lambda}{2}} P(G_0|H_j) \sum_{\alpha=1}^{\mu} P(H_j|\alpha)P(\alpha) \quad (1)$$

where H_j is j'th elite pair in the recombination pool λ, α is the number of elite species in the population μ.

The probability to fail to improve a bit in a bin given l improvements so far is

$$
\begin{aligned}
P(G_{0l}) &= \frac{1}{\mu}\sum_{j=0}^{\frac{\lambda}{2}}\Big(\frac{2n^2-(M-2l)(n+\kappa M+2l)}{2n^2}\Big)^j\binom{\frac{\lambda}{2}}{j} \\
&\quad\cdot\sum_{\alpha=1}^{\mu}\Big(\frac{(\alpha(\alpha+2\mu(\mu-\alpha)))^2}{\mu^4}\Big)^j\Big(1-\frac{(\alpha(\alpha+2\mu(\mu-\alpha)))^2}{\mu^4}\Big)^{\frac{\lambda}{2}-j} \\
&= \frac{1}{\mu}\sum_{j=0}^{\frac{\lambda}{2}}P_{\mathrm{F}}^j\binom{\frac{\lambda}{2}}{j}\sum_{\alpha=1}^{\mu}(P_{\mathrm{sel}}(\alpha))^j(1-P_{\mathrm{sel}}(\alpha))^{\frac{\lambda}{2}-j} \\
&= \frac{1}{\mu}\sum_{\alpha=1}^{\mu}\sum_{j=0}^{\frac{\lambda}{2}}\binom{\frac{\lambda}{2}}{j}(P_{\mathrm{F}}P_{\mathrm{sel}}(\alpha))^j(1-P_{\mathrm{sel}}(\alpha))^{\frac{\lambda}{2}-j} \\
&= \frac{1}{\mu}\sum_{\alpha=1}^{\mu}(1-P_{\mathrm{sel}}(\alpha)P_{\mathrm{swap}})^{\frac{\lambda}{2}} \qquad (2)
\end{aligned}
$$

The last step is due to the Binomial expansion: $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$. Therefore,

$$P(G_l) = 1-P(G_{0l}) = 1-\frac{1}{\mu}\sum_{\alpha=1}^{\mu}(1-P_{\mathrm{sel}}(\alpha)P_{\mathrm{swap}})^{\frac{\lambda}{2}}$$

Expected time until improving the fitness of a bin κ is the sum of improvements over all values of the auxiliary function:

$$\mathbf{E}T_\kappa = \sum_{l=0}^{\frac{M}{2}-1}\frac{1}{P(G_l)} \qquad (3)$$

and, finally, summing over all κ from 0 to $K-1$ we obtain (since G depends on both l and κ)

$$
\begin{aligned}
\mathbf{E}\tau_{(\mu+\lambda)EA_{1BS}} &= \sum_{\kappa=0}^{K-1}\sum_{l=0}^{\frac{M}{2}-1}\frac{1}{P(G_{l,\kappa})} = \sum_{\kappa=0}^{K-1}\sum_{l=0}^{\frac{M}{2}-1}\frac{1}{1-\frac{1}{\mu}\sum_{\alpha=1}^{\mu}(1-(\alpha(2\mu-\alpha))^2P_{swap})^{\frac{\lambda}{2}}} \\
&= \sum_{\kappa=0}^{K-1}\sum_{l=0}^{\frac{M}{2}-1}\frac{1}{1-\frac{1}{\mu}\sum_{\alpha=1}^{\mu}(1-\frac{(\alpha((2\mu-\alpha)))^2(M-2l)(n+\kappa M+2l)}{2\mu^4n^2})^{\frac{\lambda}{2}}} \qquad (4)
\end{aligned}
$$

This derivation quite clearly shows the benefit of the population size due to $\frac{1}{\mu}$ term in front of the sum over α and μ^4 in the denominator of this sum. Also, increase in the size of λ leads to reduction the probability of failure.

We test this expression numerically for different values of n, μ, λ (see Appendix C). Unfortunately, this expression does not seem to exist in closed form, so we instead go ahead with finding an approximation to it in the next subsection.

4.2 Approximate and Asymptotic Expressions

$$\begin{aligned} P(G_{0l}) &= \frac{1}{\mu}\sum_{\alpha=1}^{\mu}\Big(1-\frac{(\alpha(\alpha+2(\mu-\alpha)))^2(M-2l)(n+\kappa M+2l)}{2\mu^4n^2}\Big)^{\frac{\lambda}{2}} \\ &= \frac{1}{\mu}\sum_{\alpha=1}^{\mu}\Big(1-\frac{(\alpha(2\mu-\alpha))^2(M-2l)(n+\kappa M+2l)}{2\mu^4n^2}\Big)^{\frac{\lambda}{2}} \\ &\leq \frac{1}{\mu}\sum_{\alpha=1}^{\mu} e^{-\frac{\lambda(\alpha(2\mu-\alpha))^2(M-2l)(n+\kappa M+2l)}{4\mu^4n^2}} \approx \frac{1}{\mu}\int_1^{\mu} e^{-\Big(\frac{\alpha(2\mu-\alpha)}{\sqrt{\gamma}}\Big)^2} d\alpha \end{aligned} \quad (5)$$

The last two steps in the summand were due to $\lim_{n\to\infty}(1-\frac{k}{n})=e^{-k}$ and the monotone nature of the summand. Note that $\gamma=\frac{4\mu^4n^2}{\lambda(M-2l)(n+\kappa M+2l)}$, and, assuming that $\mu=\lambda$, the upper bound on γ is $\frac{4\mu^4n^2}{\lambda(M-2l)(n+kM+2l)} < \frac{4\mu^4n^2}{3\lambda n} = O(\mu^3 n)$. although for monotonically decreasing functions, like the one we have got, by the integral test the sum is larger than the corresponding integral, for $\mu << n$ the sum is closely approximated by the integral.

Denote $I_1 = \int_1^{\mu} f(\alpha)d\alpha = \int_1^{\mu} e^{-\Big(\frac{\alpha(2\mu-\alpha)}{\sqrt{\gamma}}\Big)^2} d\alpha$. Expanding the integrand in Taylor series around $\alpha_0 = 1$ up to the second term, we get (since $f'(\alpha_0) = -\frac{4(2\mu-1)(\mu-1)}{e^{(\frac{2\mu-1}{\sqrt{\gamma}})^2}\gamma}$)

$$f(\alpha) \approx e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2} - \frac{4(2\mu-1)(\mu-1)(\alpha-1)}{e^{(\frac{2\mu-1}{\sqrt{\gamma}})^2}\gamma} \quad (6)$$

Therefore the integral turns into

$$\begin{aligned} I_1 = \int_1^{\mu} f(\alpha)d\alpha &\approx \int_1^{\mu}\Bigg(e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2} - \frac{4(2\mu-1)(\mu-1)(\alpha-1)}{e^{(\frac{2\mu-1}{\sqrt{\gamma}})^2}\gamma}\Bigg)d\alpha \\ &= e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}(\mu-1)\Big[1-\frac{2(2\mu-1)(\mu-1)^2}{\gamma}\Big] \end{aligned} \quad (7)$$

The probability of failure is approximately (with the assumptions specified above)

$$P(G_{0l}) \approx \frac{e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}(\mu-1)\Big[1-\frac{2(2\mu-1)(\mu-1)^2}{\gamma}\Big]}{\mu}$$

Accordingly, probability of a successful swap is

$$P(G_l) \approx 1 - \frac{e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}(\mu-1)\Big[1-\frac{2(2\mu-1)(\mu-1)^2}{\gamma}\Big]}{\mu}$$

Using the sum of expectations of Geometric random variables with different parameters (Appendix B), the expected time until filling a bin, i.e. improvement of the fitness function, is (we keep the γ substitution to simplify the notation)

$$\mathbf{E}T_\kappa = \sum_{l=0}^{\frac{M}{2}-1} \frac{\gamma}{\gamma - \gamma e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2} + 2(2\mu-1)(\mu-1)^2 e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}} \tag{8}$$

We do two approximations here, first the Riemannian sums approximation to obtain $[0,1]$ bounds on the integral, and then expand the integrand in Taylor series with 2 terms around midpoint to obtain a good approximation of the integral. The Riemannian sums approximation is defined by

$$\lim_{n\to\infty} \sum_{j=1}^{n} f(x_j) = n \int_0^1 f(nx)dx + o(n)$$

and γ is transformed accordingly:

$$\gamma = \frac{4\mu^4 n^2}{(M - 2(\frac{M}{2}-1)l)(n + \kappa M + (\frac{M}{2}-1)l)}$$

and

$$I_2 = \int_0^1 \frac{\gamma dl}{\gamma - \gamma e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2} + 2(2\mu-1)(\mu-1)^2 e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}} \tag{9}$$

therefore, the expected first hitting time until the evolution of the bin κ is (Taylor series expansion of the integrand is in the Appendix A due to its length).

$$\begin{aligned}\mathbf{E}T_\kappa &\approx \left(\frac{M}{2}-1\right)\int_0^1 \frac{\gamma dl}{\gamma - \gamma e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2} + 2(2\mu-1)(\mu-1)^2 e^{-(\frac{2\mu-1}{\sqrt{\gamma}})^2}} \\ &\approx \frac{4\mu^4 n^2(\frac{M}{2}-1)}{\lambda(\frac{M}{2}+1)(\frac{M}{2}+n+\kappa M-1)\left[\frac{2(2\mu-1)(\mu-1)^2}{\sigma_1} + \frac{4\mu^4n^2}{\lambda(\frac{M}{2}+1)\sigma_2} - \frac{4\mu^4n^2}{\lambda(\frac{M}{2}+1)\sigma_2\sigma_1}\right]}\end{aligned} \tag{10}$$

where

$$\sigma_1 = e^{\frac{\lambda(2\mu-1)^2(\frac{M}{2}+1)\sigma_2}{4\mu^4n^2}}$$

$$\sigma_2 = \frac{M}{2} + n + \kappa M - 1$$

It's easy to notice that σ_1 has a very interesting property (given $\mu = \lambda$):

$$\lim_{n\to\infty} e^{\frac{\lambda(2\mu-1)^2(\frac{M}{2}+1)(\frac{M}{2}+n+\kappa M-1)}{4\mu^4n^2}} = \lim_{n\to\infty} e^{\frac{M(M+n+KM)}{\mu n^2}} = \lim_{n\to\infty} e^{\frac{M}{\mu n}+O\left(\frac{M^2}{\mu n^2}\right)} = 1$$

which means, that for sufficiently large values of n and μ the second and the third terms in the square brackets cancel each other out, and the first term is just $2(2\mu-1)(\mu-1)^2$.

Finally, summing over all κ, the number of bins in the string, we get the approximation of the convergence time of the $(\mu + \lambda)$ algorithm on RR test function.

$$\begin{aligned}
\mathbf{E}\tau^{RR}_{(\mu+\lambda)EA_{1BS}} &\approx \frac{2\mu^4 n^2 (M-2)}{\lambda(M+2)(2\mu-1)(\mu-1)^2} \sum_{\kappa=0}^{K-1} \frac{1}{\frac{M}{2}+n-1+\kappa M} \\
&= \frac{2\mu^4 n^2 (M-2)}{\lambda M(M+2)(2\mu-1)(\mu-1)^2} \cdot \\
&\Big[\psi_0\Big(\frac{\frac{M}{2}+n-1+M+KM}{M}\Big) - \psi_0\Big(\frac{\frac{M}{2}+n-1}{M}\Big)\Big] \\
&\approx \frac{2\mu^4 n^2 (M-2)}{\lambda M(M+2)(2\mu-1)(\mu-1)^2} \log\Big(1+\frac{2KM}{M+2n}\Big) \qquad (11)
\end{aligned}$$

where ψ_0 is a Digamma function (see e.g. [12, 13]). In the derivation of the asymptotic expression for this bound, all population-related terms cancel out (since $\mu = \lambda$ and both numerator and denominator have the highest term μ^4), and the order of convergence is

$$\mathbf{E}\tau^{RR}_{(\mu+\lambda)EA_{1BS}} = O\left(\frac{n^2 \log\left(1+\frac{KM}{M+n}\right)}{M}\right) \qquad (12)$$

which seems to be a result comparable to those available in literature covering fitness functions with plateaus of fitness (e.g. [8, 11, 4]). Nevertheless for small μ we show in Appendix D that the effect of the population is beneficial, but converges to a constant as $\mu \to \infty$, thus the effect relaxes with the growth of the population size.

5 Conclusions and Future Work

We have derived three expressions for convergence of an elitist $(\mu+\lambda)$EA$_{1BS}$ on Royal Roads test function: exact, approximate and asymptotic. Both the exact and approximate expressions for the expected convergence time clearly show the benefit of increase in the population when the population is relatively small, asymptotically population effect is $O(1)$, which means that as the size keeps growing its effect relaxes.

An important assumption for the approximation of $\mathbf{E}\tau^{RR}_{(\mu+\lambda)EA_{1BS}}$ was that $\mu << n$, but we never specified the relation, unlike in [9]. This is something to look at in the future. Since the effect of the population is known to be problem-specific, we will be able to get good insights into it for unimodal functions with plateaus, such as Royal Roads.

Numerical results are consistent with our findings, with the computational results lower-bounded by theoretical. Since lower ratios of $\frac{\mu}{n}$ give sharper bounds, this may shed more light on the optimal population size for problems with function plateaus.

We have performed our analysis assuming Uniform distribution of elite species in the population, something noone seems to have done in EA literature before. This is a static approach to convergence (i.e. the distribution assumption does not change throughout the run of the algorithm). We would like to look at the dynamics of the elite species and their effect on the probability of success, $P(G_l)$ and expected convergence time.

References

1. Mitchell, M., Forrest, S., Holland, J.H.: The Royal Road for Genetic Algorithms: Fitness Landscapes and GA Performance. In: European Conference on Artificial Life, vol. 1, pp. 245–254 (1992)
2. Mitchell, M.: Introduction to Genetic Algorithms. Kluwer Academic Publishers (1996)
3. Watson, R.A., Jansen, T.: A Building Block Royal Road Where Crossover is Provably Essential. In: Genetic and Evolutionary Computing Conference (GECCO), pp. 1452–1459 (2007)
4. Storch, T., Wegener, I.: Real Royal Road Function for Constant Population Size. In: Cantú-Paz, E., Foster, J.A., Deb, K., Davis, L., Roy, R., O'Reilly, U.-M., Beyer, H.-G., Kendall, G., Wilson, S.W., Harman, M., Wegener, J., Dasgupta, D., Potter, M.A., Schultz, A., Dowsland, K.A., Jonoska, N., Miller, J., Standish, R.K. (eds.) GECCO 2003. LNCS, vol. 2724, pp. 1406–1417. Springer, Heidelberg (2003)
5. Doerr, B., Fouz, M., Witt, C.: Sharp Bounds by Probability-Generating Functions and Variable Drift. In: Genetic and Evolutionary Computing Conference (GECCO), pp. 2083–2090 (2011)
6. Doerr, B., Happ, E., Klein, C.: Crossover Can Provably be Useful in Evolutionary Computation (in press, 2011)
7. Kótzing, T., Sudholt, D., Theile, M.: How Crossover Helps in Pseudo-Boolean Optimization. In: Genetic and Evolutionary Computing Conference (GECCO), pp. 705–712 (2011)
8. He, J., Yao, X.: From an Individual to a Population: An Analysis of the First Hitting Time of Population-Based Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation 6(5), 495–511 (2002)
9. Chen, T., Tang, K., Chen, G., Yao, X.: A large population size can be unhelpful in evolutionary algorithms (in press, 2011)
10. Ter-Sarkisov, A., Marsland, S., Holland, B.: The k-Bit-Swap: A New Genetic Algorithm Operator. In: Genetic and Evolutionary Computing Conference (GECCO), pp. 815–816 (2010)
11. Chen, T., He, J., Sun, G., Chen, G., Yao, X.: A New Approach for Analyzing Average Time Complexity of Population-Based Evolutionary Algorithms on Unimodal Problems. IEEE Transactions on Systems, Man, and Cybernetics B 39(5), 1092–1106 (2009)
12. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions. United States Government Printing Office (1965)
13. Graham, R.L., Knuth, D.E., Patashnik, O.: Concrete Mathematics: A Foundation for Computer Science. Addison-Wesley Publishing Company (1995)

A Taylor Series Approximation of the Integrand

We give the expression for the Equation 10 here due to its length. It's Taylor series expansion of the integrand around midpoint of the interval (0.5)

$$\phi(l) \approx \frac{4\mu^2 n^2}{\lambda(\frac{M}{2}+1)\sigma_1\sigma_3} + s_3\left(\frac{s_3\left\{\frac{\frac{s_1}{\sigma_2\sigma_4} - \frac{s_2}{\mu^4 n^2 \sigma_2(\frac{M}{2}+1)}}{\sigma_3} - \frac{16(M-2)}{\sigma_2\sigma_3^2(\frac{M}{2}+1)}\right\} - s_3\frac{(16M-32)\sigma_5}{\sigma_4\sigma_3^2} + \varphi_1}{(\frac{M}{2}+1)\sigma_1^2\sigma_3}\cdot \right.$$
$$\left. + \frac{4(M-2)\sigma_5}{\sigma_1\sigma_3^2\sigma_4}\right)(l - \frac{1}{2})$$

where

$$\sigma_1 = \frac{2(\mu-1)(\mu-1)^2}{\sigma_2} + \frac{4\mu^2 n^2}{\lambda(\frac{M}{2}+1)\sigma_3} - \frac{4\mu^4 n^2}{\lambda\sigma_2\sigma_3(\frac{M}{2}+1)}$$

$$\sigma_2 = e^{\frac{\lambda(2\mu-1)(\frac{M}{2}+1)\sigma_3}{4\mu^4 n^2}},\ \sigma_3 = \frac{M}{2}+n+kM-1,\ \sigma_4 = \left(\frac{M}{2}+1\right)^2,\ \sigma_5 = n+kM-2$$

$$s_1 = 16(M-2),\ s_2 = 4\lambda(2\mu-1)^2((M-2)(\frac{M}{2}+1) - \sigma_3(M-2)),\ s_3 = \frac{\mu^4 n^2}{\lambda}$$

$$\varphi_1 = \frac{2(2\mu-1)^3(\mu-1)^2(2M+2n+2kM-nM-kM^2-4)}{s_3\sigma_2}$$

B Probability Theory

To derive expressions in Section 4.1, we extensively used properties of independent Geometric RVs that are not identically distributed, which is also known as Coupon collector's problem (see e.g. [13]): if $X_i \sim Geom(p_i)$ it expectation is $\mathbf{E}[X_i] = \frac{1}{p_i}$. Therefore, if $Y = \sum_{i=1}^n X_i$, $\mathbf{E}Y = \sum_{i=1}^n \mathbf{E}[X_i] = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_n}$

For Equation 1 we use the Law of total probability twice: first, conditioning on H_j, then on α:

$$P(A) = \sum_{i=1}^m P(A|B_i)P(B_i) = \sum_{i=1}^m P(A|B_i) \sum_{j=1}^n P(B_i|C_j)P(C_j)$$

C Numerical Results to Verify Equation 4

Column $\tilde{\tau}_{(\mu+\lambda)EA_{1BS}}$ was obtained by running the algorithm with different parameters 20 times, each run was 2000 generation each. The earliest achievement of the global minimum for each run was saved and then averaged over.

The results are very consistent in terms of exposing the effect of the population growth and are sharper for smaller ratios of $\frac{\mu}{n}$. Like we expected in the Assumptions section, theoretical bounds obtained are optimistic due to higher probabilities of observing high numbers of elite species.

Table 2. Theoretical and computational bounds for $(\mu+\lambda)EA^{RR}_{1BS}$

n	K	M	μ	λ	$\mathbf{E}\tau^{RR}_{(\mu+\lambda)EA_{1BS}}$	$\tilde{\tau}_{(\mu+\lambda)EA_{1BS}}$
32	4	8	4	4	145	315.3077
			10	10	72.4	268.2195
			20	20	44.2	192.2917
			30	30	34.5	173.5625
64	8	8	4	4	570.625	612.46
			10	10	279.88	497.93
			20	20	153.46	454.4681
			30	30	112.297	372.04
128	16	8	4	4	2264.36	1365
			10	10	1048	1239
			20	20	570.44	1091.5
			30	30	401.99	949.4

D Effect of the Population

We rewrite Equation 11 in order to factor out terms involving μ, taking for simplicity $\mu = \lambda$.

$$\mathbf{E}\tau_{(\mu+\lambda)EA_{1BS}} = \varphi(n, M, K)\frac{\mu^4}{\lambda\mu(2\mu-1)(\mu-1)^2} = \varphi(n, M, K)\frac{\mu^3}{2\mu^3 - 5\mu^2 + 4\mu - 1}$$

For small values of μ this expression lies between 0.5 and 1 and quickly converges to 0.5, so $\mathbf{E}\tau_{(\mu+\lambda)EA_{1BS}} = O(\varphi(n, M, K))$ because asymptotically

$$\frac{\mu^3}{2\mu^3 - 5\mu^2 + 4\mu - 1} = O(1)$$

This explains the benefit of the growth of the population when it is small and its leveling out for larger values