Runtime Analysis of Abstract Evolutionary Search with Standard Crossover

Tina Malalanirainy
University of Exeter
United Kingdom
tm425@exeter.ac.uk

Alberto Moraglio
University of Exeter
United Kingdom
a.moraglio@exeter.ac.uk

ABSTRACT

The Convex Search Algorithm (CSA) is a generalization across representations of Evolutionary Algorithms (EAs) with crossover and no mutation. The Standard Evolutionary Search Algorithm (SESA) is a more accurate generalization of EAs with crossover and no mutation, using a standard two-parents crossover. This work extends the runtime analysis of the CSA on quasi-concave landscapes [4] to the SESA. We instantiate the analysis to binary strings and integer vectors endowed with the Hamming distance and the Manhattan distance. We find that the SESA requires a larger population size to converge to a global optimum; resulting in a larger runtime upper bound than the CSA. Empirical studies on LeadingOnes confirmed the existence of a smallest population size above which both algorithms are guaranteed to find the global optimum. Below this threshold, the SESA is less successful than the CSA.

CCS CONCEPTS

• **Theory of computation → Design and analysis of algorithms**; Mathematical optimization; Discrete optimization; Optimization with randomized search heuristics; **Evolutionary algorithms**; **Randomness, geometry and discrete structures**; Random search heuristics; Theory and algorithms for application domains; *Theory of randomized search heuristics*;

KEYWORDS

Runtime analysis, Population size lower bound, Crossover

ACM Reference Format:
Tina Malalanirainy and Alberto Moraglio. 2019. Runtime Analysis of Abstract Evolutionary Search with Standard Crossover. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion), July 13–17, 2019, Prague, Czech Republic.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3319619.3321959

1 MOTIVATION

EAs with crossover and no mutation have been generalized across representations as CSA [2]. Nevertheless, the CSA is not common in practice as it makes use of a population based operator called convex hull recombination. Indeed, most crossover operators are generalized across representations by geometric crossover; which picks an offspring in the geodesic segment of two parents [1, 3]. Hence, a more accurate algorithm that generalizes EAs with crossover and no mutation across representations, should make use of geometric crossover instead of convex hull recombination. The runtime of an EA with crossover and no mutation across representations has only been analysed in [4] for a general EA with only a convex hull recombination. We extend the result across representations of [4] by considering EAs with only a geometric crossover.

GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3321959

2 METHOD

When both the geometric crossover and the convex hull recombination sample the convex hull of the parent population, the runtime analysis of the CSA on quasi-concave landscapes in [4] can be readapted. This is done by considering the worst case uniform offspring distribution on quasi-concave landscapes; for an EA with only a geometric crossover. Recall from [4], that a level set contains individuals whose fitness are at least a fixed threshold value. Furthermore, the level sets of a quasi-concave landscape are always geodesically convex. A parent population P' covers a level set L entirely, when the convex hull $co(P')$ of P' equals the level set L. The probability that the convex hull of k individuals sampled uniformly at random from P' covers the level set L is denoted $P_M^{\text{cov}}(k)$. The runtime analysis is done in four steps that are summarized below.

(i) Estimate a lower bound on the probability that the next parent population covers a level set that is strictly higher than the level set covered by the current parent population,
(ii) Estimate a lower bound on the probability that a strictly higher level set is always covered at each generation,
(iii) Estimate the smallest population size for which the lower bound of Step ii is greater than or equals 0.5,
(iv) Estimate the worst runtime corresponding to the smallest population size obtained in Step iii.

3 RESULTS

Both the geometric crossover and the convex hull recombination sample the same set of offspring in metric spaces where the set $Seg(A)$, of the union of all the segments that can be made out of the points of A, is geodesically convex for any subset A. A counter example where the sets $Seg(A)$ and $co(A)$ do not coincide is illustrated in Figure 2. The union of all the segments that can be made out of the points of A is $Seg(A) = \{x_1, y_1, x_2, y_2, x, y\}$. Let us consider the segment $[x, y]$ whose extremes are points of $Seg(A)$. We have $[x, y] = \{x_1, y_1, x_2, y_2, x, y, z\}$. We can see that $z \in [x, y]$ but

Table 1: Lower bound μ_0 on the population size for which the algorithm finds a global optimum with probability at least 0.5.

$(\{0,1\}^n, \mathrm{HD})$	$(\{0,1,\ldots,d-1\}^n, \mathrm{HD})$	$(\{0,1,\ldots,d-1\}^n, \mathrm{MD})$
$\frac{m_0^2 \log_2[(4q+6)n]}{r}$	$\frac{d m_0^2 \ln[2(q+2)dn]}{r}$	$\frac{d m_0^2 \ln[4(q+2)n]}{r}$

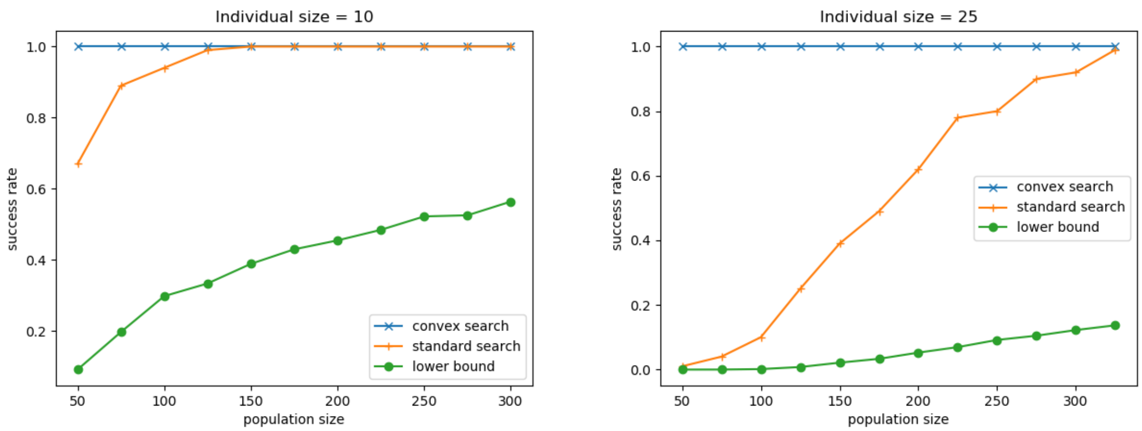

Figure 1: Average success rates for the Convex Search Algorithm (CSA) and the Standard Evolutionary Search Algorithm (SESA) on LeadingOnes against the average lower bound on the success probability.

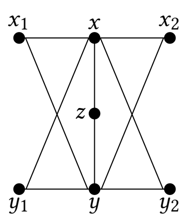

Figure 2: The set $Seg(A)$ where $A = \{x_1, x_2, y_1, y_2\}$ is not a convex set

$z \notin Seg(A)$ though $x, y \in Seg(A)$. Thus, the set $Seg(A)$ is not geodesically convex and therefore does not coincide with $co(A)$.

We also found that the offspring distribution corresponding to the geometric crossover is not uniform on the convex hull of the parent population, as for the case of the convex hull recombination. However, by considering the worst case where the offspring are uniformly distributed with the least possible probability, we were able to readapt the analysis of the CSA for the SESA. The lower bound of Step ii is given by

$$b(m,l) = \left[P_M^{\mathrm{cov}}\left(\frac{\mu r}{m^2}\right)\right]^{q+1} - q\exp\left[-\frac{\mu r}{2}\left(\frac{l^2-1}{l^2}\right)^2\right], \tag{1}$$

where μ is the population size, the parameter m (resp. l) is the largest (resp. smallest) size of the parent population at each generation, $r < 1$ is the smallest ratio of the sizes of two consecutive level sets, and $q+1$ is the total number of distinct level sets. Hence, Equation (1) is bounded below by

$$b = \min_{1 \le l \le m \le \mu} b(m,l). \tag{2}$$

Let m_0 and l_0 be solutions of Equation (2). A lower bound μ_0 on the population size for which (2) is at least 0.5 can be computed for $(\{0,1\}^n, \mathrm{HD})$, $(\{0,1,\ldots,d-1\}^n, \mathrm{HD})$, and $(\{0,1,\ldots,d-1\}^n, \mathrm{MD})$ using m_0. In this case, the SESA finds a global optimum within $2\mu_0 q$ fitness evaluations. Table 1 gives a summary of the different values of μ_0 for each metric space. Empirical results on LeadingOnes shown in Figure 1, confirmed that the SESA requires a larger population size to find a global optimum with probability at least 0.5.

4 CONCLUSION

Deriving solutions for Equation (2) is not straightforward, as it is a function of two variables. Consequently, the bounds on the population size shown in Table 1 can not be computed unless a solution to Equation (2) is found. Therefore, the results above are only useful in comparing the worst success rates of the SESA and the CSA on quasi-concave landscapes. Moreover, the current analysis of the SESA on quasi-concave landscapes is restricted to metric spaces where the set $Seg(A)$ is geodesically convex for all subsets A. A different analysis needs to be devised for metric spaces where a non geodesically convex set $Seg(A)$ exists for some subset A.

REFERENCES

[1] Alberto Moraglio. 2008. *Towards a geometric unification of evolutionary algorithms.* Ph.D. Dissertation. University of Essex.
[2] Alberto Moraglio. 2011. Abstract convex evolutionary search. In *Proceedings of the 11th workshop proceedings on Foundations of genetic algorithms.* ACM, 151–162.
[3] Alberto Moraglio and Riccardo Poli. 2004. Topological interpretation of crossover. In *Genetic and Evolutionary Computation Conference.* Springer, 1377–1388.
[4] Alberto Moraglio and Dirk Sudholt. 2017. Principled design and runtime analysis of abstract convex evolutionary search. *Evolutionary computation* 25, 2 (2017), 205–236.