The Benefit of Recombination in Noisy Evolutionary Search

Tobias Friedrich
Hasso Plattner Institute
Potsdam, Germany

Timo Kötzing
Hasso Plattner Institute
Potsdam, Germany

Martin S. Krejca
Hasso Plattner Institute
Potsdam, Germany

Andrew M. Sutton
Hasso Plattner Institute
Potsdam, Germany

ABSTRACT

Practical optimization problems frequently include uncertainty about the quality measure, for example due to noisy evaluations. Thus, they do not allow for a straightforward application of traditional optimization techniques. In these settings, randomized search heuristics such as evolutionary algorithms are a popular choice because they are often assumed to exhibit some kind of resistance to noise. Empirical evidence suggests that some algorithms, such as estimation of distribution algorithms (EDAs) are robust against a scaling of the noise intensity, even without resorting to explicit noise-handling techniques such as resampling.

In this paper, we want to support such claims with mathematical rigor. We introduce the concept of *graceful scaling* in which the run time of an algorithm scales polynomially with noise intensity. We study a monotone fitness function over binary strings with additive noise taken from a Gaussian distribution. We show that myopic heuristics cannot efficiently optimize the function under arbitrarily intense noise without any explicit noise-handling. Furthermore, we prove that using a population does not help. Finally we show that a simple EDA called the Compact Genetic Algorithm can overcome the shortsightedness of mutation-only heuristics to scale gracefully with noise. We conjecture that recombinative genetic algorithms also have this property.

This extended abstract summarizes our work "The Benefit of Recombination in Noisy Evolutionary Search," which appeared in *Proceedings of International Symposium on Algorithms and Computation (ISAAC)*, 2015, pp. 140–150.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

Keywords

Evolutionary algorithms, run time analysis, noisy optimization

GECCO'16 Companion July 20-24, 2016, Denver, CO, USA

ACM ISBN 978-1-4503-4323-7/16/07.

DOI: http://dx.doi.org/10.1145/2908961.2930953

1. INTRODUCTION

Evolutionary algorithms are widely used for solving real-world optimization problems in uncertain environments. In many practical situations, the evaluation of the objective function is not deterministic, but has a large stochastic component. In these scenarios, evolutionary algorithms must somehow filter the fitness signal from the noise. If the noise intensity is relatively small, this poses little to no problems to selection. However, as the noise intensity grows, the signal becomes more obscured, and the picture is no longer as clear.

In [2], we address the dependence of optimization time on noise intensity (measured as the variance). Specifically, we ask whether there is some kind of threshold point in the noise intensity at which the noise becomes too high for efficient optimization, or is it possible for algorithms to be somehow robust to scaling of the noise level?

In the first survey article regarding evolutionary algorithms in noisy environments, Beyer [1] pointed out that the case of *extreme noise* can be easily treated: there is no useful selection information and so an EA will essentially perform a random walk. We prove that this is indeed the case, at least when an algorithm is myopic in the sense that it only makes local changes. However, we also contend that some algorithms, such as estimation of distribution algorithms, can leverage the underlying fitness signal in such a way that their behavior never defaults to the diffusion-like behavior observed for myopic algorithms.

To formally characterize how search heuristics can exhibit robustness to noise intensity, we introduce the concept of *graceful scaling*; intuitively, a search heuristic scales gracefully with noise if (polynomially) more noise can be compensated by (polynomially) more resources.

2. OUR RESULTS

We consider centered Gaussian noise with variance σ^2 and use OneMax as the underlying fitness function. Already such a seemingly simple setting poses difficulties to the analysis of evolutionary algorithms, as these algorithms are not developed with the analysis in mind. We first prove that there is indeed a noise intensity threshold for myopic algorithms. For simple hillclimbers like RLS and the (1+1) EA, the threshold is already as low as a constant. However, we also show that a population cannot help, and the $(\mu{+}1)$ EA using *any polynomial population size* of at least $\omega(1)$ and a noise intensity of $\sigma^2 = \omega(n^2)$ is provably inefficient. This implies that the algorithm cannot scale gracefully for Gaussian noise.

On the other hand, we also investigate a simple estimation of distribution algorithm (EDA) known as the Compact Genetic Algorithm (cGA). The working principles of the cGA are in stark contrast to the (μ+1) EA. Rather than relying on local, myopic mutation operations, the cGA maintains an underlying product distribution that reflects an estimate of the ideal *allele frequencies* a true population would have. In each iteration, it compares two individuals drawn from this product distribution and updates the frequencies based on tournament selection. The only parameter used by the cGA is an update step-size $1/K$ that governs how much the marginal probabilities change in each iteration.

This approach allows the cGA to smooth the noise sufficiently without having to resort to explicit noise-handling strategies. We prove that as long as K is sufficiently large, that is $K = \omega(\sigma^2\sqrt{n}\log n)$, then the cGA scales gracefully with Gaussian noise. Specifically, we prove that after $\mathcal{O}(K\sigma^2\sqrt{n}\log Kn)$ many iterations, the cGA will with high probability have converged the all-1-string, as desired. In other words, there is no threshold point in noise intensity at which the algorithm begins to perform poorly.

This result gives insight into how the cGA can filter the signal out of the noise efficiently. Even under intense noise (i.e., a large variance), as long as the allele frequency update $1/K$ is small enough, selection errors due to misclassification by the noisy function are not penalized greatly, and overall the allele frequencies *drift* in the right direction. This can be contrasted with mutation-only approaches for which such selection errors become fatal in the sense that progress in the correct direction is no longer visible to the algorithm.

We conjecture that evolutionary algorithms that explicitly use recombination can also leverage similar mechanisms to exhibit graceful scaling, however we do not prove this in the current paper. A step in that direction was recently made by Prügel-Bennett, Rowe, and Shapiro [3] who proved that a generational evolutionary algorithm using uniform crossover needs only $\mathcal{O}(n\log^2 n)$ function evaluations to optimize OneMax with additive noise of variance $\sigma^2 = n$.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 618091 (SAGE).

References

[1] H.-G. Beyer, "Evolutionary Algorithms in Noisy Environments: Theoretical Issues and Guidelines for Practice," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, no. 2–4, pp. 239–267, 2000.

[2] T. Friedrich, T. Kötzing, M. S. Krejca, and A. M. Sutton, "The benefit of recombination in noisy evolutionary search," in *Proceedings of International Symposium on Algorithms and Computation (ISAAC)*, 2015, pp. 140–150.

[3] A. Prügel-Bennett, J. Rowe, and J. Shapiro, "Run-time analysis of population-based evolutionary algorithm in noisy environments," in *Proceedings of Foundations of Genetic Algorithms (FOGA)*. ACM, 2015, pp. 69–75.