Why Is Parity Hard for Estimation of Distribution Algorithms?

David Coffin
Department of Computer Science, UCL
Gower Street
London WC1E 6BT
+44 (0) 207 679 0427
d.coffin@cs.ucl.ac.uk

Robert E. Smith
Department of Computer Science, UCL
Gower Street
London WC1E 6BT
+44 (0) 207 679 4115
robert.elliott.smith@gmail.com

ABSTRACT
We describe a *k*-bounded and additively separable test problem on which the hierarchical Bayesian Optimization Algorithm (hBOA) scales exponentially.

Categories and Subject Descriptors
G.1.6 [**Numerical Analysis**]: Optimization – *global optimization.*

General Terms: Algorithms, Theory.

Keywords
Parity, hierarchical Bayesian Optimization Algorithm.

1. INTRODUCTION
Functions exist that contain no lower order dependencies, although they contain high-order dependencies that significantly affect an individual's overall fitness. In the binary domain, the most extreme form of this problem is parity.

$$\text{parity}(S) = \begin{cases} C_{even} & \text{if } bc(S) \text{ is even} \\ C_{odd} & \text{otherwise} \end{cases}$$

where *bc* is a bit count (unitation) function over bit string *S*.

The CPF is defined below, where *m* is the number of concatenated sub-functions, *k* is the size, in bits, of each sub-function, and S_i denotes the *i*'th bit in bit string *S*.

$$\text{CPF}(S) = \sum_{i=0}^{m} \text{parity}(S_{ik} \ldots S_{ik+k})$$

We used non-overlapping concatenated test functions, with *k*=5, c_{odd}=5 and c_{even}=0, for all experiments. For the CPF, 2^{l-m} strings are optimal – an exponentially decreasing proportion of the search space.

As an example of a state-of-the-art EDA, we measured the number of evaluations required for hBOA to find any one of these optima, following the methodology and parameter settings of Pelikan et al. in [1]. In particular we used the bisection method to determine population size, with assumed convergence in a maximum of *n*
generations. Once the population size was determined using the bisection method, we ran hBOA 100 times for each problem size.

GECCO'07, July 7–11, 2007, London, England, United Kingdom.
ACM 978-1-59593-697-4/07/0007.

2. RESULTS
Figure 1 is a semi-log plot of the mean number of evaluations hBOA required to find any individual with optimal fitness on the CPF problem together with the population sizes. Results indicate that hBOA scales exponentially ($1.49e^{0.123n}$, R2=0.942).

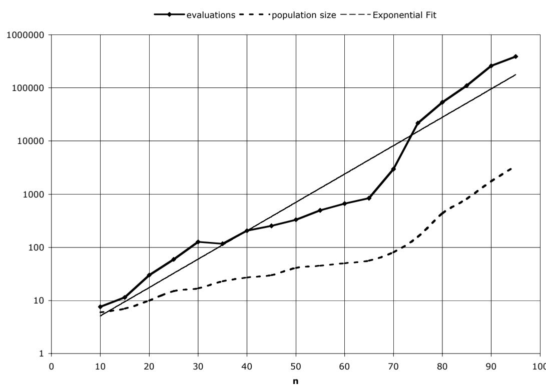

Figure 1: A semi-log scalability plot of hBOA on the CPF.

3. ANALYSIS
Given the expected selection from a random population, for a *k*-sized parity function the joint probabilities equal the products of the marginal probabilities for any sub-set of less than *k* variables. This implies that there is no there is no correlation of any kind between any pair of variables within a parity function. In particular, there is no mutual information.

Although there are dependencies between the variables in a parity sub-function, an EDA will not be expected to detect them in a random population unless it considers modeling *k*-wise dependencies. If the EDA fails to model these dependencies when constructing its first probabilistic model, then it will assume each bit is independent, and those bits will be randomly distributed in the subsequent generation as well, and all future generations by induction. All the EDAs we are aware of begin building their probabilistic models by looking for pair-wise dependencies.

4. REFERENCES
[1] Pelikan, M., Sastry, K., Butz, M. V., and Goldberg, D. E. *Hierarchical boa on random decomposable problems.* Technical Report 2006002, IlliGAL, 2006.