Complexity of Model Learning in EDAs: Multi-structure Problems

Hadi Sharifi
Iran University of Science and Technology
hsharifi@comp.iust.ac.ir

Amin Nikanjam
K.N.Toosi University of Technology
nikanjam@kntu.ac.ir

Hossein Karshenas
University of Isfahan
hosseinkarshenas@gmail.com

Negar Najimi
K.N.Toosi University of Technology
ngr.najimi@gmail.com

ABSTRACT

Many of the real-world problems can be decomposed into a number of sub-problems for which the solutions can be found easier. However, proper decomposition of large problems remains a challenging issue, especially in optimization, where we need to find the optimal solutions more efficiently. Estimation of distribution algorithms (EDAs) are a class of evolutionary optimization algorithms that try to capture the interactions between problem variables when learning a probabilistic model from the population of candidate solutions. In this paper, we propose a type of synthesized problems, specially designed to challenge this specific ability of EDAs. They are based on the principal idea that each candidate solution to a problem may be simultaneously interpreted by two or more different structures where only one is true, resulting in the best solution to that problem. Of course, some of these structures may be more likely according to the statistics collected from the population of candidate solutions, but may not necessarily lead to the best solution. The experimental results show that the proposed benchmarks are indeed difficult for EDAs even when they use expressive models such as Bayesian networks to capture the interactions in the problem.

Categories and Subject Descriptors

G.1.6 [**Optimization**]: Global Optimization

General Terms

Algorithms, Performance, Theory

Keywords

Linkage learning, Model building, Hard problems, Effectiveness

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
ACM 978-1-4503-2881-4/14/07.
http://dx.doi.org/10.1145/2598394.2598479.

1. INTRODUCTION

Estimation of Distribution Algorithms (EDAs) [1] are a class of evolutionary algorithms. The main characteristic of EDAs is building a probabilistic model according to the population of promising solutions. The learned model is supposed to stochastically encode the dependencies between the variables of problem. This model is later used to generate new solutions that are incorporated into the population. In this way, EDAs may overcome the traditional problem of linkage learning.

In this paper we present a new class of benchmark problems for discrete EDAs. A well-known EDA is employed to solve this problem and the results are analyzed. Our experiments reveal that this class of problems is very hard for EDAs. Actually EDA is deceived when trying to solve these problems and defeated in learning the correct model of the problem.

2. THE MULTI-STRUCTURE PROBLEMS

The problems we introduce in this section are formed by combining two different structures from which only one results in the global optimum. To solve these problems, it is not only necessary to learn the interactions between variables, what EDAs are capable of, but also it is important to encode the interactions constituting the correct model. The problems get even harder when it is more likely to estimate the model leading to the local optimum, resulting in a special kind of deception in the model learning phase of EDAs. In other words, learning the incorrect model misleads EDAs from the global optimum. The optimal fitness value in these problems is achieved just by interpreting the input string using the correct structure.

The first two problems are built from combining $m-k$ trap and onemax functions. The first function uses a control bit within the input string to decide on interpreting the rest of the variables:

$$MSP1(x) = \begin{cases} \alpha + f_{trap(m,k)}(x_2, \ldots, x_n) & x_1 = 1 \\ (n-1) - f_{onemax}(x_2, \ldots, x_n) & x_1 = 0 \end{cases} \quad (1)$$

where α is a constant determining the global optima of the problem. In the second function, no control bit is used and the two structures are directly compared together when applied on the whole string:

$$MSP2(x) = \max\left\{\alpha + f_{trap(m,k)}(x),\right.$$
$$\left.(n-1) - f_{onemax}(x)\right\} \quad (2)$$

where α is defined the same as before.

The next problem combines two trap functions with different orders of interaction:

$$MSP3(x) = \max\left\{\alpha + f_{trap(m_1,k_1)}(x),\right.$$
$$\left.f_{trap(m_2,k_2)}(\bar{x})\right\} \quad (3)$$

where α is defined the same as before, $\bar{x}$ is the negation of string x and we have $k_1 \neq k_2$ and $m_1k_1 = m_2k_2$.

3. EXPERIMENTAL RESULTS

To study the complexity of the introduced problems, in this section the results of applying a well-known Bayesian EDA, BOA [2], on increasing sizes of these problems are presented. Parameters of algorithm are as recommended by the authors of the original paper. The algorithm stops after finding the optimum solution or elapsing 100 generations. For each problem size, the population size is determined using the bisection method [3] with 30 independent successful runs. All the results are averaged over these 30 successful runs. The experiments are performed on a 2.8GHz dual-core Intel Pentium 4 computer with 4GB of memory. The problem instances considered for the experiments use 3-bit and 5-bit trap groups. The α parameter is set to 1. Figures 1 and 2 show the average number of fitness evaluations for each problem before reaching the optimum solution.

The results show that the number of fitness evaluations required to solve MSP1 increases much faster than $m-k$ trap function for both $k=3$ and $k=5$. MSP2 is much harder and the required computational time to solve this problem grows very rapidly, thus preventing the algorithm to efficiently solve larger instances of this problem. The difficulty of the third problem, MSP3, depends on which of the constituting structures lead to the global optimum. As presented in the results, the problem is harder than normal $m-k$ trap function only when the structure formed over larger groups specifies the global optimum (here the structure related to 5-bit trap groups). Note that when the simpler structure determines the global optimum, a structure-learning EDA, like BOA, is not deceived by the second structure. For example, considering the trap function, the structure with k=3 is simpler than the one with k=5 because it is formed over smaller groups of variables. In this case the algorithm can solve the problem with a performance very close to that of $m-k$ trap functions or even simpler (see MSP3($k_1=3$) and MSP3($k_1=5$) in figures 1 and 2). This means that the algorithm is able to identify the correct structure to be used for solution sampling and ignores the deceptive one.

4. CONCLUSIONS

In the literature, some problems were proposed to be hard for EDAs, tackling the model-building capability of these algorithms. In this paper, we introduce a new concept to synthesize hard problems for EDAs which are called multi-structure problems (MSPs). In the proposed problems, the fitness of an input string can be evaluated by two different structures, one of which is simpler comparing to the other when considering the complexity of model learning. The

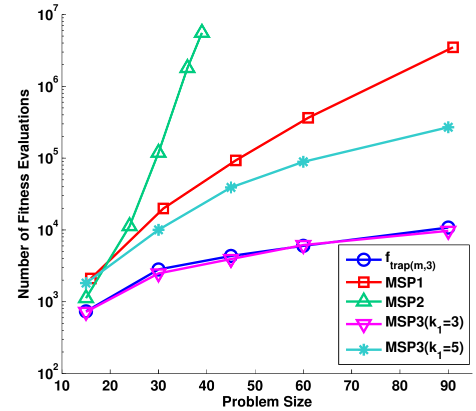

Figure 1: Average number of fitness evaluations required to find the global optimum of the problems with $k=3$

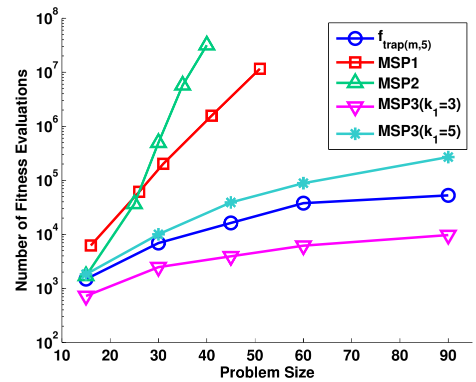

Figure 2: Average number of fitness evaluations required to find the global optimum of the problems with $k=5$

optimal fitness value is obtained using the structure with higher complexity. Learning the correct model is the key to reliably reach the global optimum in these problems. A well-known Bayesian EDA has been examined for solving three version of these problems. The experimental results revealed that solving MSPs are indeed very hard for EDA. Actually EDA tends to learn the simpler model, causing them to mislead from the global optimum. We believe that these problems present the idea of deceptive model building.

5. REFERENCES

[1] P. Larrañaga and J. A. Lozano, editors. *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[2] M. Pelikan, D. E. Goldberg, and E. Cantú-Paz. BOA: The Bayesian optimization algorithm. In *Genetic and Evolutionary Computation Conference (GECCO)*, volume 1, pages 525–532, Orlando, Florida, USA, July 1999. Morgan Kaufmann Publishers.

[3] K. Sastry. Evaluation-relaxation schemes for genetic and evolutionary algorithms. Master's thesis, University of Illinois at Urbana-Champaign, 2001.