Improving the Efficiency of the Extended Compact Genetic Algorithm

Thyago S. P. C. Duque, David E. Goldberg, Kumara Sastry
Illinois Genetic Algorithms Laboratory,
University of Illinois at Urbana Champaign,
Urbana, Illinois, USA
thyago@illigal.ge.uiuc.edu, deg@illigal.ge.uiuc.edu, kumara@kumarasastry.com

ABSTRACT

Evolutionary Algorithms are largely used search and optimization procedures that, when properly designed, can solve intractable problems in tractable polynomial time. Efficiency enhancements are used to turn them from tractable to practical.

In this paper we show preliminary results of two efficiency enhancements proposed for the Extended Compact Genetic Algorithm. First, a model building enhancement was used to reduce the complexity of the process from $O(n^3)$ to $O(n^2)$, speeding up the algorithm by 1000 times on a 4096 bits problem. Then, local-search hybridization was used to reduce the population size by at least 32 times, reducing the memory and running time required by the algorithm. These results draw the first steps toward a competent and efficient Genetic Algorithm.

Categories and Subject Descriptors:
G.1.6 [**Optimization**]: *Global optimization*

General Terms: Algorithms, Performance

Keywords: Estimation of Distribution Algorithms, ECGA, Model Building, Efficiency Enhancement

1. INTRODUCTION

Evolutionary Algorithms (EA) [2] have been successfully used in several different applications. Goldberg [3] presents a methodology to design competent Genetic Algorithms (GAs), which can solve hard problems accurately, efficiently and reliably, solving intractable problems in a tractable polynomial time. However, to solve large problems it is sometimes necessary to enhance the efficiency [6] of the algorithm.

The Extended Compact Genetic Algorithm (ECGA) [4] is a competent GA that uses the Marginal Product Model (MPM) to summarize important information on the population and to sample a new and possibly better population. An MPM has two components: a partition over the variables and a probability distribution over each partition. The Model Building (MB) process of the ECGA finds a parti-

tion that appropriately represents the population by greedily optimizing the Combined Complexity Criterion (CCC) using Algorithm 1. The ECGA repeats the Selection-MB-Sampling loop until a convergence criterion is satisfied.

This work proposes two efficiency enhancements for the ECGA as the first step toward a competent and efficient GA. It is motivated by the recent success of the Compact Genetic Algorithm (cGA) on solving a billion bit noisy optimization problem [5].

Algorithm 1 Model Building procedure in the ECGA

```
1. Assign each variable to an independent partition
2. Repeat until no improvement is possible:
3.     For each pair of partitions:
4.         Evaluate the CCC of the merged pair
5.     Merge the pair with smaller CCC, if any
```

---

2. EFFICIENCY ENHANCEMENTS FOR THE ECGA

Competent GAs can solve hard problems with polynomial resources; however, efficiency enhancements are necessary to solve large problems. This paper proposes two extensions to improve the performance of the ECGA: the MB relaxation and the hybridization of the ECGA with local search.

The mk-trap problem, which consists on the concatenation of m independent k-bit traps, was used as a benchmark function in this paper, using $k = 4$.

According to profiling information from gprof, the MB consumes 90% of the running time of the ECGA. For a chromosome of size ℓ, this process has complexity $O(\ell^3)$ since step 3 of algorithm 1 iterates $O(\ell^2)$ while step 2 iterates on additional $O(\ell)$. To improve this step we propose a relaxed MB presented on Algorithm 2, which has complexity $O(\ell^2)$ and still produce the same model for trap functions[1].

Algorithm 2 Improved greedy MB for the ECGA

```
1. Assign each variable to an independent partition
2. Repeat until no improvement is possible:
3.    Choose a partition (a) in a round-robin manner
4.    For each other partition (b):
5.        Evaluate the CCC when a and be are merged
6.    Merge the pair a and b that induced the smaller CCC
7.    With probability pb=0.95pb, break a partition
```

---

[1]An interesting future direction is to determine if the accuracy holds for other problems
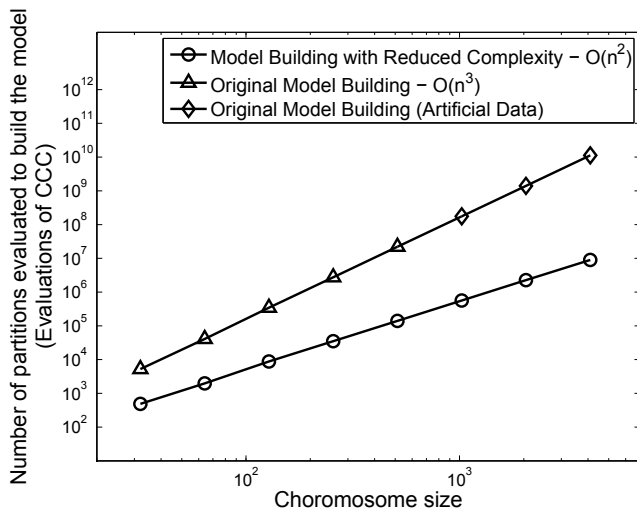
Figure 1: A comparison between Algorithm 1 and 2 shows a reduction from $O(\ell^3)$ to $O(\ell^2)$ in the number of evaluations of the CCC to build the model, resulting in speed-ups of 1000 times for a 4096 bits problem

Figure 1 compares Algorithm 1 and 2 on a log-log scale, showing the complexity reduction (concerning evaluations of the CCC) induced by the relaxation. The running time reduction was from about 10 times for a 32 bits problem to more than 1000 times for a 4096 bits problem.

The performance of the ECGA can be greatly improved by pre-processing the initial population using local search. In this paper, we use a binary bit-flip algorithm that processes each string sequentially, flipping a bit and keeping the change if fitness improves. The pre-processing improves the performance by removing all non locally-optimal subsolutions and consequently reducing the entropy of the building blocks (BBs). The smaller entropy reduces the population required to accurately build the model as shown by Figure 2. The local search also reduces the number of different BBs to be processed, reducing the mixing time [3].

Although the local search requires several function evaluations, the hybrid is no worse than the ECGA in this criterion, still reducing the population size and, consequently, memory and running time.

Due to space restrictions, most of the data and reasoning that led to these enhancements was omitted. Please refer to [1] for additional information.

3. CONCLUSION

This work presents a proof of principle on the importance of efficiency enhancements for practical GAs. We present two enhancements to the ECGA, a widely used competent genetic algorithm. These enhancements successfully improved performance of the algorithm, with speedups of more than 1000 times for large problems.

Although the results achieved by these enhancements are relevant, this work is only the first step toward a competent and efficient EA, capable of solving difficult large scale problems in practical time. Other efficiency enhancements,
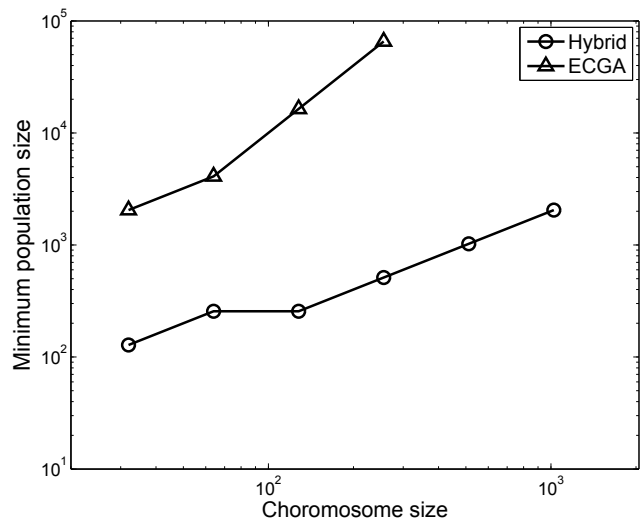

Figure 2: The Hybridization of the ECGA with a local search procedure reduces the minimum population size required by the algorithm, consequently reducing memory and running time

including parallelization, sporadic model building and chromosome compression could be useful next steps.

4. ACKNOWLEDGMENTS

5. REFERENCES

[1] T. Duque, D. Goldberg, and K. Sastry. Enhancing the Efficiency of the ECGA. IlliGAL Report No. 2008006, Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, 2008.

[2] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, 1989.

[3] D. E. Goldberg. *The Design of Innovation: Lessons from and for Competent Genetic Algorithms.* Kluwer Academic Publishers, 2002.

[4] G. Harik, F. Lobo, and K. Sastry. Linkage Learning via Probabilistc Modeling in the Extended Compact Genetic Algorithm (ECGA). *Scalable Optimization via Probabilistic Modeling, Studies in Computational Inteligence*, pages 39–61, 2006. (Also Illigal Report No. 99010).

[5] K. Sastry, D. Goldberg, and X. Llora. Towards billion-bit optimization via a parallel estimation of distribution algorithm. *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 577–584, 2007.

[6] K. Sastry, M. Pelikan, and D. Goldberg. Efficiency enhancement of estimation of distribution algorithms. *Scalable Optimization via Probabilistic Modeling: From Algorithms to Applications*, pages 161–185, 2006.