Combining Model-Based EAs for Mixed-Integer Problems[1]

Krzysztof L. Sadowski [a] Dirk Thierens [a] Peter A. N. Bosman [b]

[a] *Utrecht University, The Netherlands*
[b] *CWI Amsterdam, The Netherlands*

1 Introduction

A key characteristic of Mixed-Integer (MI) problems is the presence of both continuous and discrete problem variables. In this paper, we study the design of an algorithm that integrates the strengths of LTGA [2] [3] and iAMaLGaM [1]: state-of-the-art model-building Evolutionary Algorithms (EAs) designed for discrete and continuous search spaces, respectively. We wish to study if making use of the model building and learning abilities of both these algorithms can be applied to MI problems while retaining excellent scale-up behavior. The model-building nature of these algorithms allows us to consider black-box problems where no prior information about a problem structure is known. How difficult is it to achieve a proper evaluation balance and adequate scalability as the problem size increases? Is it even possible to solve dependent problems where continuous variables interact with the discrete ones, while using integrated but independently learning models?

2 Problems

We use well-established problems (see full paper for definitions) and adapt them into the MI setting. We consider different combinations of discrete and continuous problems where the contributions of the discrete and continuous parts are kept independent through addition. Minimization is assumed.

In the full paper, we introduce five functions, F_1 - F_5 which represent different types of variable dependencies. Variables in F_1 are fully independent. Only continuous variables are dependent in F_2. Only discrete variables are dependent in F_3. In F_4 both sub-spaces are dependent. The F_5 benchmark includes cross-domain dependencies between the continuous and discrete variables. It is an additively decomposable specific combination of the deceptive trap function with the rotated ellipsoid (see full paper for details).

3 Results

For the analysis of $F_1 - F_4$, we consider different problem lengths l. For each problem size, we consider different proportions of variables used with 5, $0.25l$, $0.5l$, $0.75l$ and $l-5$ continuous variables (remaining variables are discrete). Success criterion is solving a problem 29/30 times with the precision of 10^{-10}.

Heat maps in Figure 1 show that more evaluations are required for the same problem sizes as the composition of the problem shifts towards more continuous variables. Moreover, benchmarks which contain dependencies within the continuous sub-space, F_2 and F_4, require larger number of evaluations than F_1 or F_3. Population sizes are also affected by the problem composition. In F_3 and F_4 we observe much larger population size requirements, as the landscape of these functions includes discrete

[1]The full paper has been accepted for publication in the *13th International Conference on Parallel Problem Solving from Nature (PPSN'14).*"

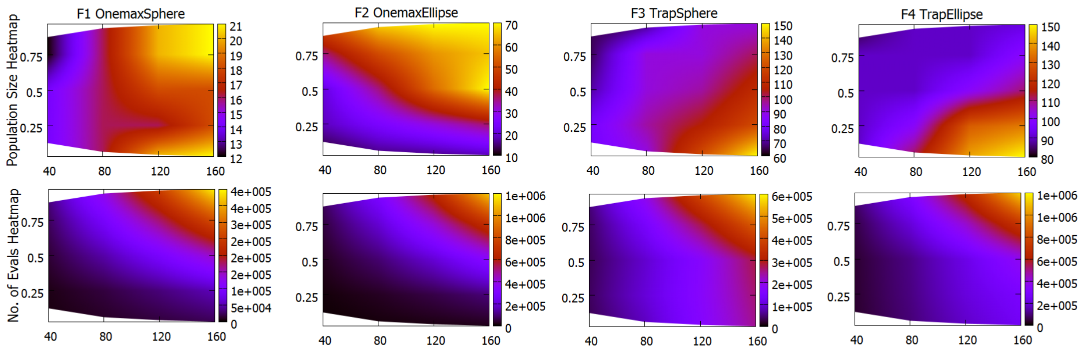

Figure 1: Heat Maps representing the population sizes (top row) and evaluations (bottom row) needed for different variable compositions. Horizontal axis represents the problem length, the vertical axis is the fraction of continuous variables ($l_c/(l_c + l_d)$) in the problem

variable dependencies. This shows that in addition to problem length, the composition of the problem and variable dependencies are a big factor for efficiency in terms of evaluations and population sizes. Scalability analysis in the full paper verified that the results exhibit polynomial scalability on the tested MI problems. Results on F_5 in the full paper show that the hybrid algorithm we propose is in fact capable of solving this dependent benchmark, however not in all cases. When the fitness contribution of the discrete variables is scaled down to very small values, it causes their initial fitness contributions to be very small, resulting in the algorithm prematurely converging on sub-optimal solutions. As this contribution is scaled up, the problem becomes simpler and requires smaller population sizes and less evaluations in order to be solved.

4 Conclusions

Mixed-Integer problems introduce many optimization challenges which do not arise in purely real or discrete optimization problems. Obtaining a proper balance in exploration of model information for different types of variables, varying variable ratios and additional overhead or fitness contribution scaling are some of the important issues which should be taken into account when solving MI problems. Our algorithm achieved polynomial scale-up behavior on the tested benchmarks. We showed that a well-balanced algorithm can solve some cases of even very dependent mixed-integer problems, despite having independent model learning methods for the discrete and continuous sub-spaces. The results provide a good foundation and motivation for further work in mixed-integer landscapes with model building EAs.

References

[1] Peter A. N. Bosman, Jörn Grahl, and Dirk Thierens. Enhancing the Performance of Maximum-Likelihood Gaussian EDAs Using Anticipated Mean Shift. In *Parallel Problem Solving from Nature — PPSN X*, volume 5199 of *LNCS*, pages 133–143, 2008.

[2] Dirk Thierens. The linkage tree genetic algorithm. In *Parallel Problem Solving from Nature — PPSN XI*, volume 6238 of *LNCS*, pages 264–273, Berlin, 2010. Springer–Verlag.

[3] Dirk Thierens and Peter A N Bosman. Optimal mixing evolutionary algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '11, pages 617–624, New York, NY, USA, 2011. ACM.