Population Size and Scalability in the AεSεH Evolutionary Many-objective Algorithm

Hernán Aguirre
Faculty of Engineering, Shinshu University
4-17-1 Wakasato, Nagano, 380-8553 JAPAN
ahernan@shinshu-u.ac.jp

Arnaud Liefooghe
Université Lille 1, LIFL, UMR CNRS 8022, France
Inria Lille-Nord Europe, France
arnaud.liefooghe@lifl.fr

Sébastien Verel
Université Nice Sophia-Antipolis, France
Inria Lille-Nord Europe, France
verel@i3s.unice.fr

Kiyoshi Tanaka
Faculty of Engineering, Shinshu University
4-17-1 Wakasato, Nagano, 380-8553 JAPAN
ktanaka@shinshu-u.ac.jp

ABSTRACT

This work studies the effects of population size on performance scalability of the Adaptive ε-Sampling and ε-Hood evolutionary many-objective algorithm.

Categories and Subject Descriptors

I.2.8 [**ARTIFICIAL INTELLIGENCE**]: Problem Solving, Control Methods, and Search—*Heuristic methods*

General Terms

Algorithms, Design, Performance, Verification

Keywords

Population size, Scalability, Many-objective optimization

1. INTRODUCTION

The population size greatly influences the dynamics of a multi-objective evolutionary algorithm. However, its effects on large dimensional objectives spaces are not well understood, particularly its correlation to the fidelity of selection to retain optimal solutions and not drop them in favor of inferior solutions that appear non-dominated in the population.

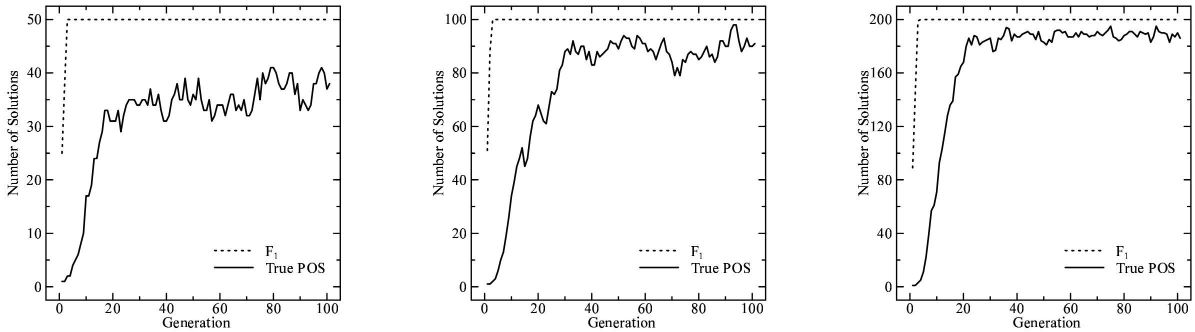
In this work we study the effects of population size on performance scalability of the Adaptive ε-Sampling and ε-Hood evolutionary many-objective algorithm (AεSεH) [1]. We are particularly interested on understanding the relationship between the size of the Pareto optimal set, a characteristic of the many-objective problem at hand, the population size the algorithm uses, and the ability of the algorithm to retain true Pareto optimal solutions in its population, which is directly correlated to the effectiveness of survival selection and the maintenance of selection pressure to find new optimal solutions.

In our study we use a MNK-landscape [2] randomly generated with $m = 5$, objectives, $n = 20$ bits, and $k = 1$ epistatic bit. The motivation to use small landscapes with minimum non-linearity is that it should be relatively simple for the algorithm to hit the optimal set. In addition, it is feasible to enumerate these small landscapes. So, we can easily verify the ability of the algorithm to retain optimal solutions in the population. Precisely, we analyze the dynamics of the algorithm observing the number of true Pareto optimal solutions found at each generation and their accumulated number found during evolution.

GECCO'13 Companion, July 6–10, 2013, Amsterdam, The Netherlands.
ACM 978-1-4503-1964-5/13/07.

2. THE AεSεH ALGORITHM

AεSεH [1] is an elitist evolutionary many-objective algorithm that applies ε-dominance principles both for survival selection and for clustering and mating solutions located close by in objective space. Firstly, during extinctive selection AεSεH computes dominance among solutions and eliminates dominated ones. Then, it applies *ε-sampling* to the large set of non-dominated solutions. Here, randomly sampled non-dominated solutions survive and those ε-dominated by the samples are eliminated. The aim is to get a set of surviving solutions spaced according to the distribution implicit in the mapping function $\boldsymbol{f}(\boldsymbol{x}) \mapsto^{\epsilon} \boldsymbol{f}^{'}(\boldsymbol{x})$ used for ε-dominance. After survival selection, in AεSεH there is not an explicit ranking that could be used to bias mating. Rather, the algorithm uses a procedure called *ε-hood creation* to cluster solutions in objective space. This method is also based on ε-dominance. Here, a randomly sampled solution from the surviving population and its ε-dominated solutions determine a neighborhood. Mating for recombination is implemented by the procedure *ε-hood mating*, where neighborhoods formed by *ε-hood creation* are considered to be elements of a list. To select two mates, first a neighborhood from the list is specified deterministically in a round-robin schedule. Then, two individuals are selected randomly within the specified neighborhood, so that an individual will recombine with other individual that is located close by in objective space. The motivation to restrict mating is to enhance the effectiveness of recombination in many-objective problems, where the difference in variable space between individuals in the population is expected to be larger than in multi-objective problems and therefore more disruptive for recombination. The number of sampled solutions by *ε-sampling* and the number of neighborhoods created by *ε-hood creation* depend on the value of ε used for ε-dominance, ε_s and ε_h (≥ 0), respectively. The algorithm adapts both parameters at each generation, so that the number of solu-



(a) $|\mathcal{P}| = 50$ (b) $|\mathcal{P}| = 100$ (c) $|\mathcal{P}| = 200$

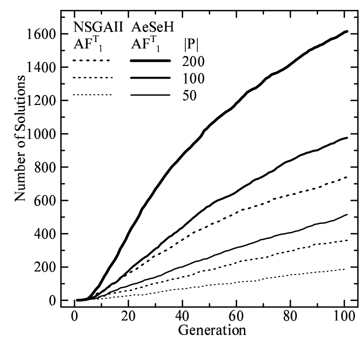
Figure 1: Non-dominated solutions F_1 and true Pareto optimal solutions F_1^T in the population $\mathcal{P}$.

tions sampled with ε_s is close to the size of the population and the number of neighborhoods created with ε_h is close to a number specified by the user. AεSεH in this work uses two point crossover with rate $pc = 1.0$ and bit flip mutation with rate $pm = 1/n$. The reference neighborhood size is set to 20 individuals. The mapping function $\boldsymbol{f}(\boldsymbol{x}) \mapsto^{\epsilon} \boldsymbol{f}'(\boldsymbol{x})$ used for ε-dominance in *ε-sampling* and *ε-hood creation* is additive $f_i' = f_i + \varepsilon$, $i = 1, 2, \cdots, m$.

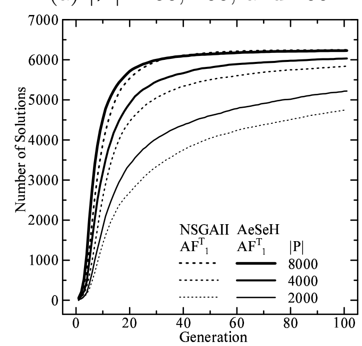
3. RESULTS AND DISCUSSION

Fig.1 shows the number of non-dominated solutions $|F_1|$ and the number of true Pareto optimal solutions $|F_1^T|$ (True POS) in the population after survival selection over the generations, running AϵSϵH for 100 generations with three small population sizes $|\mathcal{P}| = 50$, 100 and 200. Note that after few generations all solutions in the population are non-dominated, $|F_1| = |\mathcal{P}|$. However, note that the number of true Pareto optimal solutions is smaller than the number of non-dominated solutions, $|F_1^T| < |F_1|$, and fluctuates up and down. This shows that optimal solutions are dropped from the population when survival selection is applied, especially in very small populations. Fig.2 (a) shows the accumulated number of true Pareto optimal solutions $|AF_1^T|$ found by AεSεH during evolution for the same populations, together with those found by NSGA-II for comparison. Note that a substantially larger number of true Pareto optimal solutions are accumulated by AεSεH than by NSGA-II. This is because AεSεH retains more optimal solutions than NSGA-II, which in turns keeps a stronger selection pressure that enables the algorithm to find more new optimal solutions.

Fig.2 (b) shows AF_1^T by AεSεH and NSGA-II using population sizes that correspond approximately to 1/3, 2/3 and 4/3 of the true Pareto optimal solutions set POS^T of the landscape. Note that when a population larger than $|POS^T| = 6265$ is used both algorithms perform similarly in terms of the accumulated number of true Pareto optimal solutions found. This is because such a huge population can cover a broad region of objective space, eliminating the effect of dropping optimal solutions, which enhances selection. However, for smaller populations, AεSεH is more effective and efficient due to its ability to retain optimal solutions.



(a) $|\mathcal{P}| = 50$, 100, and 200



(b) $|\mathcal{P}| = 1/3$, 2/3 and 4/3 of $|POS^T| = 6265$

Figure 2: Accumulated number of true Pareto optimal solutions AF_1^T by AεSεH and NSGA-II.

4. CONCLUSIONS

We showed that optimal solutions are dropped from the population during survival selection in many-objective optimization. Larger populations reduce this undesired effect and improve the performance of the optimizer. AεSεH uses ε-sampling to induce a uniform distribution of the surviving solutions. This enhances selection fidelity when small populations are used and improves the effectiveness of the algorithm.

5. REFERENCES

[1] H. Aguirre, A. Oyama, and K. Tanaka, "Adaptive ε-Sampling and ε-Hood for Evolutionary Many-Objective Optimization", EMO Conference, LNCS, Springer, pp. 322–336, vol. 7811, 2013.

[2] H. Aguirre and K. Tanaka, "Insights on Properties of Multi-objective MNK-Landscapes", *Proc. 2004 IEEE Congress on Evolutionary Computation*, IEEE Service Center, pp.196–203, 2004.