Analysis on Population Size and Neighborhood Recombination on Many-Objective Optimization

Naoya Kowatari[1], Akira Oyama[2], Hernán Aguirre[1], and Kiyoshi Tanaka[1]

[1] Faculty of Engineering, Shinshu University
4-17-1 Wakasato, Nagano, 380-8553 Japan
[2] Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency
{kowatari@iplab,ahernan,ktanaka}@shinshu-u.ac.jp,
oyama@flab.isas.jaxa.jp

Abstract. This work analyzes population size and neighborhood recombination in the context of many-objective optimization. Large populations might support better the evolutionary search to deal with the increased complexity inherent to high dimensional spaces, whereas neighborhood recombination can reduce dissimilarity between crossing individuals and would allow us to understand better the implications of a large number of solutions that are Pareto-optimal from the perspective of decision space and the operator of variation. Our aim is to understand why and how they improve the effectiveness of a dominance-based many-objective optimizer. To do that, we vary population size and analyze in detail convergence, front distribution, the distance between individuals that undergo crossover, and the distribution of solutions in objective space. We use DTLZ2 problem with $m = 5$ objectives in our study, revealing important properties of large populations, recombination in general, and neighborhood recombination in particular, related to convergence and distribution of solutions.

1 Introduction

Recently, there is a growing interest on applying multi-objective evolutionary algorithms (MOEAs) to solve many-objective optimization problems [1], where the number of objective functions to optimize simultaneously is considered to be more than three. It is well known that conventional MOEAs [2,3] scale up poorly with the number of objectives of the problem, which is often attributed to the large number of non-dominated solutions and the lack of effective selection and diversity estimation operators to discriminate appropriately among them, particularly in dominance-based algorithms. Selection, indeed, is a fundamental part of the algorithm and has been the subject of several studies, leading to the development of evolutionary algorithms that improve the performance of conventional MOEAs on many-objective problems [1]. However, finding trade-off solutions that satisfy simultaneously the three properties of convergence to the Pareto front, well spread, and well distributed along the front is especially difficult to achieve in many-objective problems and most search strategies for many-objective optimization proposed recently compromise one in favor of the other [1].

C.A. Coello Coello et al. (Eds.): PPSN 2012, Part II, LNCS 7492, pp. 22–31, 2012.

In addition to selection, detailed studies on the characteristics of many-objective landscapes, the effectiveness of operators of variation, and the effects of large populations are important to move forward in our understanding of evolutionary many-objective optimization in order to develop effective and efficient algorithms. From this standpoint, we have presented initial evidence that MOEAs can improve their performance on many-objective problems by using large populations and neighborhood recombination [4].

In this work, our aim is to understand why and how population size and neighborhood recombination increase the effectiveness of the algorithm. To study that, we choose NSGA-II [5] as our base algorithm and incorporate neighborhood recombination into it. We vary population size and analyze in detail convergence, front distribution, the distance between individuals that undergo crossover, and the distribution of solutions in objective space. We use DTLZ2 problem [6] with $m = 5$ objectives in our study.

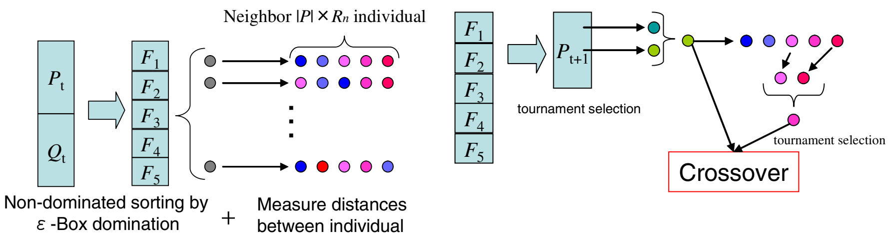
The motivation to look into large populations is that they might support better the evolutionary search to deal with the increased complexity inherent to high dimensional spaces. On the other hand, the motivation to study recombination is to understand better the implications of a large number of solutions that are Pareto-optimal from the perspective of decision space and the operator that make moves on it. A large number of non-dominated solutions could cause a large diversity of individuals in the instantaneous population and recombination of very dissimilar individuals could be too disruptive. Neighborhood recombination aims to reduce dissimilarity between crossing individuals.

Our study reveals important properties of large populations, recombination in general, and neighborhood recombination in particular, related to convergence and distribution of solutions.

2 Method

In many-objective problems the number of non-dominated solutions grows substantially with the number of objectives of the problem [7,8]. A side effect of this is that non-dominated solutions tend to cover a larger portion of objective and variable space compared to problems with fewer objectives [9]. The implications of a large number of non-dominated solutions have been studied in objective space, where selection operates. However, little is known about the implications in decision space, where recombination and mutation operate. It is expected that the large number of non-dominated solutions in many objective problems induce a large diversity of individuals in the instantaneous population. In which case recombination of very dissimilar individuals could be too disruptive affecting its effectiveness.

Neighborhood Recombination encourages mating between individuals located close to each other, aiming to reduce dissimilarity between crossing individuals and improve the effectiveness of recombination in high dimensional objective spaces. We choose NSGA-II as our base algorithm and incorporate neighborhood recombination into it. In this work, we leave untouched selection of NSGA-II,



(a) Neighborhood creation

(b) Mating for recombination

Fig. 1. Neighborhood Recombination

which uses a primary ranking based on dominance and a secondary ranking based on crowding distance. This would allow us to show and explain the effects of population size and recombination with a well known selection.

The main steps of Neighborhood Recombination are as follows. During the calculation of dominance relationships, the proposed method calculates the distance between individuals in objective space and keeps a record of the $|P| \times R_n$ closest neighbors of each individual. Note that when the ranked population of size $|P|+|Q|$ is truncated to form the new population of size $|P|$, some individuals would be deleted from the neighborhood of each individual. When individuals are selected for recombination, the first parent p_A is chosen from the parent population P using a binary tournament, while the second parent p_B is chosen from the neighborhood of p_A using another binary tournament. Then, recombination is performed between p_A and p_B. That is, between p_A and one of its neighbors in objective space. If all neighbors of individual p_A were eliminated during truncation, the second parent p_B is selected from the population P similar to p_A. Fig.1 illustrates the neighborhood creation and mating for recombination. In this work, we set the parameter that defines the size of the neighborhood of each individual to $R_n = 0.1$ $(10\%|P|)$.

3 Test Problem and Analysis Indicators

3.1 Test Problem

We study the behavior of the algorithms using the continuous function DTLZ2 [5], setting the number of objectives to $m = 5$ and the total number of variables to $n = m+9$. Problem DTLZ2 is designed in such a way that the Pareto-optimal front corresponds to a non-convex surface in objective space, which lies in the positive quadrants of the unit hyper-sphere, with Pareto-local fronts constructed parallel to it. To achieve this, the n variables of a solution $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ are classified in two subsets. The first $m-1$ variables $x_1, x_2, \cdots, x_{m-1}$, denoted $\boldsymbol{x}_{1:m-1}$, determine the position of solutions within the Pareto-local/optimal front, whereas the remaining $n-m+1$ variables $x_m, x_{m+1}, \cdots, x_n$, denoted $\boldsymbol{x}_{m:n}$, determine the distance of the Pareto-local front to the Pareto-optimal front. When $x_m = x_{m+1} = \cdots = x_n = 0.5$, the solution is located in the Pareto-optimal front. The m objective functions used in DTLZ2 are as follows

$$\left.\begin{array}{ll}
f_1(\boldsymbol{x}) & = (1 + g(\boldsymbol{x}_{m:n})) \prod_{i=1}^{m-1} cos(\frac{\pi}{2} x_i) \\
f_2(\boldsymbol{x}) & = (1 + g(\boldsymbol{x}_{m:n})) \left(\prod_{i=1}^{m-2} cos(\frac{\pi}{2} x_i) \right) sin(\frac{\pi}{2} x_{m-1}) \\
f_3(\boldsymbol{x}) & = (1 + g(\boldsymbol{x}_{m:n})) \left(\prod_{i=1}^{m-3} cos(\frac{\pi}{2} x_i) \right) sin(\frac{\pi}{2} x_{m-2}) \\
\vdots & \\
f_{m-1}(\boldsymbol{x}) & = (1 + g(\boldsymbol{x}_{m:n})) cos(\frac{\pi}{2} x_1) sin(\frac{\pi}{2} x_2) \\
f_m(\boldsymbol{x}) & = (1 + g(\boldsymbol{x}_{m:n})) sin(\frac{\pi}{2} x_1)
\end{array}\right\} \quad (1)$$

$$g(\boldsymbol{x}_{m:n}) = \sum_{i=m}^{n} (x_i - 0.5)^2 \quad (2)$$

3.2 Analysis Indicators

Proximity Indicator(I_p) [10]: Measures convergence of solutions by

$$I_p = \underset{\boldsymbol{x} \in P}{median} \left\{ \left[\sum_{i=1}^{m} (f_i(\boldsymbol{x}))^2 \right]^{\frac{1}{2}} - 1 \right\}, \quad (3)$$

where $\boldsymbol{x}$ denotes a solution in the population P. Smaller values of I_p indicate that the population P is closer to the Pareto-optimal front and therefore mean better convergence of solutions.

Mates Distance(D_c): Euclidian distances in variable space between pairs of solutions that undergo crossover are computed separately for the subsets of variables $\boldsymbol{x}_{1:m-1}$ and $\boldsymbol{x}_{m:n}$ that determine the position of solutions within the front and their distance to the Pareto-optimal front, respectively. Here we report the average distances $D_c(\boldsymbol{x}_{1:m-1})$ and $D_c(\boldsymbol{x}_{m:n})$ taken over all pairs of solutions that undergo crossover at a given generation.

Distribution of Solution in Objective Space (ψ_k): In order to observe where solutions of the parent population P are located in objective space, we classify each solution according to the number $k = 0, 1, \cdots, m-1$ of its objective values that are very small compared to the maximum objective value of the solution. More formally, The class k a solution belongs to is determined by

$$k = \sum_{i=1}^{m} \theta_i \quad (4)$$

$$\theta_i = \begin{cases} 1 & \text{if } f_i(\boldsymbol{x}) < f_{max}(\boldsymbol{x})/100 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where $f_{max}(\boldsymbol{x}) = max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x})\}$. Roughly, a solution belonging to class $k = 0$ is considered to be in the central region of objective space, whereas solutions belonging to class $k \geq 1$ are gradually considered to be in the edges of objective space and identify dominant resistant solutions. We report ψ_k, the number of solutions in population P belonging to class k.

Front Distribution: Shows the number of solutions per front obtained after applying non-dominated sorting to the combined population of parents P and offspring Q. Here, we report results for fronts $F_1 \sim F_5$.

4 Simulation Results and Discussion

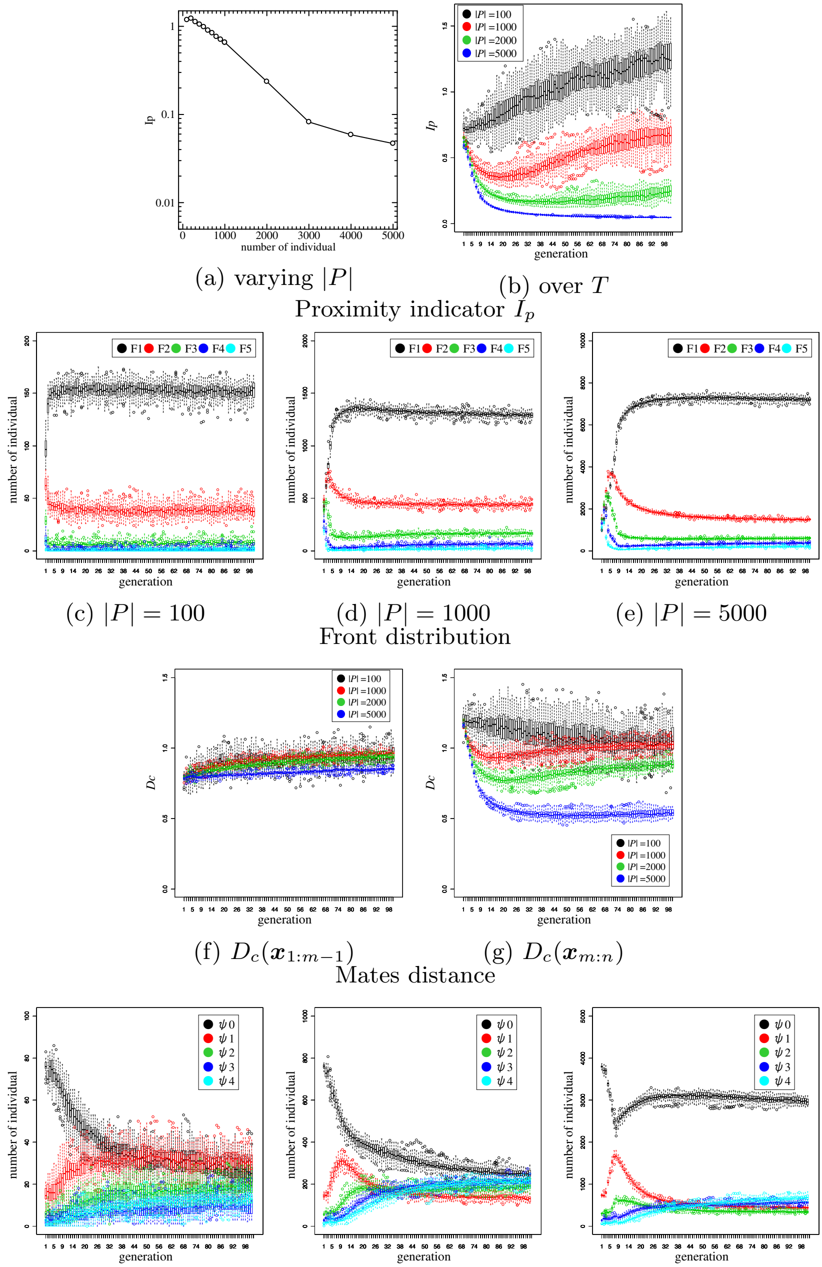
4.1 Preparation

In this work we use NSGA-II[6] as a base algorithm and include in its framework Neighborhood Recombination. We observe the behavior of conventional NSGA-II and NSGA-II with Neighborhood Recombination varying the population size from $|P| = 100$ to 5000 individuals. As genetic operators we use SBX crossover and Polynomial Mutation, setting their distribution exponents to $\eta_c = 15$ and $\eta_m = 20$, respectively. The parameter for the operators are crossover rate $p_c = 1.0$, crossover rate per variable $p_v = 0.5$, and mutation rate $p_m = 1/n$, where n is the number of variables. The maximum number of generations is fixed to $T = 100$. Here we report average results obtained running the algorithms 30 times.

4.2 Analysis Varying Population Size in NSGA-II

In this section we analysis the behavior of NSGA-II. Results are shown in Fig.2. First, we look at the convergence of the algorithm. Fig.2(a) shows I_p of Pareto-optimal solutions obtained in the final generation ($T = 100$) increasing the population size from $|P| = 100$ to 5000. It can be seen that I_p gets smaller by increasing population size $|P|$. That is, a larger population improves convergence of the algorithm. In order to investigate these results with more detail, Fig.2(b) shows the transition of I_p over the generations when the algorithm is set to population sizes $|P| = 100, 1000, 2000, 5000$. In the case of $|P| = 100$, it can be seen that I_p increases substantially. This indicates that the algorithm diverges from the Pareto-optimal front, rather than converge to it, as evolution proceeds. However, signs of convergence gradually appear by increasing population size $|P|$. Eventually, for $|P| = 5000$ no divergence is observed and I_p reduces to 0.05 with very small dispersion. Here an important conclusion is that population size is strongly correlated to the convergence ability of the algorithm.

Then, we analyze the front distribution over the generations. Results for the first five fronts $F_1, \cdots, F_5$ are shown in Fig.2(c)$\sim$(e) for population sizes $|P| = 100, 1000, 5000$, respectively. Note that the number of solutions in the first front $|F_1|$, obtained after applying non-dominated sorting to the combined population of parents and offspring of size $2|P|$, is larger than the size of the parent



(a) varying $|P|$ (b) over T

Proximity indicator I_p

(c) $|P| = 100$ (d) $|P| = 1000$ (e) $|P| = 5000$

Front distribution

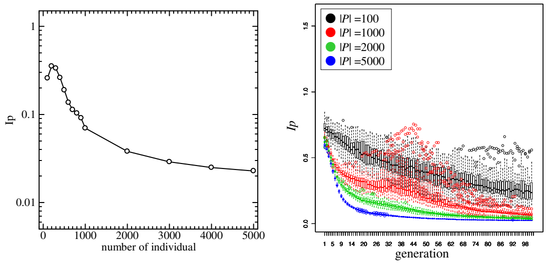
(f) $D_c(\boldsymbol{x}_{1:m-1})$ (g) $D_c(\boldsymbol{x}_{m:n})$

Mates distance

(h) $|P| = 100$ (i) $|P| = 1000$ (j) $|P| = 5000$

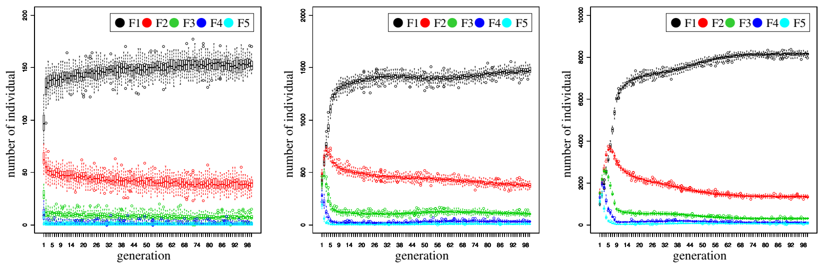
Distribution of solutions in objective space

Fig. 2. Results by NSGA-II



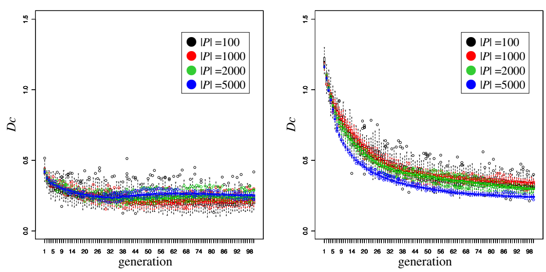
(a) over P (b) over T

Proximity indicator I_p



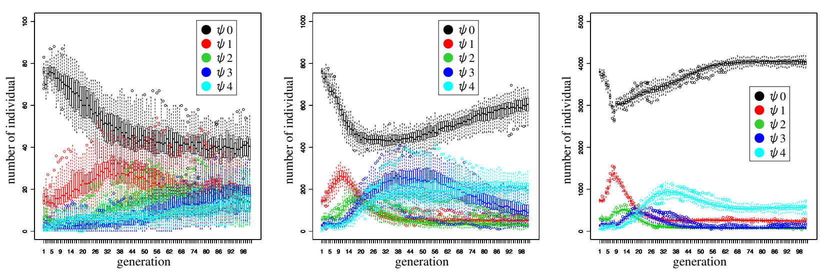
(c) $|P| = 100$ (d) $|P| = 1000$ (e) $|P| = 5000$

Front distribution



(f) $D_c(\boldsymbol{x}_{1:m-1})$ (g) $D_c(\boldsymbol{x}_{m:n})$

Mates distance



(h) $|P| = 100$ (i) $|P| = 1000$ (j) $|P| = 5000$

Distribution of solutions in objective space

Fig. 3. Results by NSGA-II with Neighborhood Recombination

population$|P|$ for most of the evolution. Especially when $|P| = 100$, the ratio $|F_1|/|P|$ is the highest and $|F_1|$ exceeds $|P|$ at a very early generation. When the population increases to $|P|$=1000 and 5000, the ratio $|F_1|/|P|$ reduces and it takes few more generations until $|F_1|$ exceeds the size of the parent population. Looking closely at Fig.2 (e) that shows results for $|P| = 5000$, we can see that the first 10 generations when $|F_1| < |P|$ is precisely the period where I_p reduces significantly, as shown in Fig.2 (b) $|P| = 5000$. If $|F_1| < |P|$ then the population is composed by solutions coming from two or more fronts, which means that parent selection can discriminate based on dominance ranking and not only on crowding distance as it is the case when $|F_1| > |P|$. These results suggest that a large enough random initial population is able to include lateral diversity (solutions from different fronts), which allows dominance-based selection to pull the population in the direction of the Pareto-optimal front.

Next, we look at $D_c(\boldsymbol{x}_{1:m-1})$ and $D_c(\boldsymbol{x}_{m:n})$, the average distances in decision space between individuals that undergo crossover. Note from Fig.2(f) that D_c $(\boldsymbol{x}_{1:m-1})$, computed on the subset of variables that determine the position within the front, is large at generation 0 and it tends to increase as the evolution proceeds. Note also that this trend becomes less pronounced as we increase the population size $|P|$. This shows the diversity of solutions and it is evidence that recombination takes place among very dissimilar solutions, raising questions about its effectiveness to help convergence. On the other hand, from Fig.2 (g) note that $D_c(\boldsymbol{x}_{m:n})$, computed on the subset of variables that determine the distance to the Pareto-optimal front, becomes smaller with the generations as we increase $|P|$. This reduction of $D_c(\boldsymbol{x}_{m:n})$ is expected if the population converge towards the Pareto-optimal front, which is located at $\boldsymbol{x}_m = \boldsymbol{x}_{m+1} = \cdots = \boldsymbol{x}_n = 0.5$.

Finally, we analyze the distributions of solutions in objective space $\psi_k(k = 0, 1, \cdots, 4)$ shown in Fig.2(h)$\sim$(j) for population sizes $|P| = 100, 1000, 5000$, respectively. Note that in the case of $|P| = 100, 1000$, the number of solutions ψ_0 are initially around 75% of the population size $|P|$, but after few generations it reduces to around 30% of $|P|$, showing that the number of solutions ψ_k, $k \geq 1$, close to the axis of the objectives functions increase significantly as evolution proceeds. On the other hand, when a population size $|P| = 5000$ is used, in few generations the number of solutions ψ_0 reduce from around 75% of $|P|$ to 50% of $|P|$, but then it rises and remains around 60% of $|P|$ until the last generation. This shows that larger populations are not easily pulled towards the extreme regions of objective space, caused by selection based on crowding distance that is at work for most generations since the size of the first front surpasses the population size ($|F_1| > |P|$).

4.3 Analysis of Neighborhood Recombination

In this section we analyze the behavior of NSGA-II with Neighborhood Recombination. Results are shown in Fig.3. Similar to the previous section, first we look at convergence. From Fig.3(a) note that the inclusion of neighborhood recombination reduces I_p drastically compared to NSGA-II for any population size, as shown in Fig.2(a). From Fig.3(b) it is important to note that, contrary to NSGA-II, no divergence of solutions is observed over the generations when neighborhood recombination is used, even for very small populations.

Next, we look at the front distribution shown in Fig.3(c)∼(e). Note that a similar trend to NSGA-II can be observed. However, when neighborhood recombination is used $|F_1|$ gradually increases with the number of generations, while in NSGA-II $|F_1|$ remained high but relatively constant.

Then, we analyze the distances among solutions that undergo crossover shown in Fig.3(f),(g). Looking at Fig.3(f), note that $D_c(x_{1:m-1})$ reduces substantially compared to NSGA-II. This is an effect of recombining individuals with one of its neighbors. By selecting a partner close in objective space we are increasing the likelihood of selecting one that is also close in variable space, unless the functions are highly non-linear. Most importantly, a short $D_c(x_{1:m-1})$ indicates that recombination takes place in less dissimilar individuals, which increases significantly the effectiveness of recombination for any population size as corroborated by the reduction of I_p shown above. Fig.3(g) shows that $D_c(\boldsymbol{x}_{m:n})$ shortens as the algorithm approaches better fronts, resembling the reduction of I_p.

Finally, looking at distribution of solutions ψ_k in objective space shown in Fig.3(h)∼(j), comparing with NSGA-II it can be seen that the number of solutions ψ_0 increases when neighborhood recombination is used, whereas the number of solutions ψ_k, $k \geq 1$, reduces. This shows that there are more solutions in the central region and fewer in the edges of objective space, even for very small populations. That is, a more effective recombination helps convergence and resists the pull of selection towards extreme regions of objective space.

5 Conclusions

In this work, we have studied the effects of population size and neighborhood recombination on the search ability of a dominance-based MOEA applied to many-objective optimization. We chose NSGA-II as our base algorithm and included in it an operator that keeps track of neighbors and recombine individuals that are close to each other in objective space. We varied population size and analyzed in detail convergence, front distribution, the distance between individuals that undergo crossover, and the distribution of solutions in objective space using problem DTLZ2 with $m = 5$ objectives.

Our results showed that population size is strongly correlated to the convergence ability of the algorithm. A large enough random initial population is able to include lateral diversity (solutions belonging to different fronts), which allows dominance-based selection to pull the population in the direction of the Pareto-optimal front. We also showed that small populations are easily pulled towards extreme regions of objective space by selection based on crowding distance and that larger populations gradually become more resistant to this effect.

We argued and presented evidence that recombination in many-objective optimization takes place on highly dissimilar individuals if no restriction is put to partner selection. Also, we verified that neighborhood recombination takes place in less dissimilar individuals and showed that this increases significantly the effectiveness of recombination for any population size, helping convergence and resisting the pull of selection towards extreme regions of objective space.

In this work we have focused mostly on population size and recombination. Selection is a fundamental part of the algorithm and there is ongoing work analyzing the effects of population size and recombination using improved selection mechanisms for many-objective optimization. However, due to space limitations, we shall report our finding elsewhere.

In the future, we would like to extend our analysis to other problems, increase the number of objectives and variables, and look at other ways to perform effective recombination in many-objective spaces. Also, determining an appropriate population size according to the number of objectives is an important area of research.

References

1. Ishibuchi, H., Tsukamoto, N., Nojima, Y.: Evolutionary Many-Objective Optimization: A Short Review. In: Proc. 2008 IEEE Congress on Evolutionary Computation, pp. 2424–2431. IEEE Press (2008)
2. Deb, K.: Multi-Objective Optimization using Evolutionary Algorithms. John Wiley & Sons, Chichester (2001)
3. Coello, C., Van Veldhuizen, D., Lamont, G.: Evolutionary Algorithms for Solving Multi-Objective Problems. Kluwer Academic Publishers, Boston (2002)
4. Kowatari, N., Oyama, A., Aguirre, H., Tanaka, K.: A Study on Large Population MOEA Using Adaptive Epsilon-Box Dominance and Neighborhood Recombination for Many-objective Optimization. In: Proc. Learning and Intelligent Optimization Conference, LION 6. LNCS. Springer (January 2012) (to appear)
5. Deb, K., Thiele, L., Laumanns, M., Zitzler, E.: Scalable Multi-Objective Optimization Test Problems. In: Proc. 2002 Congress on Evolutionary Computation, pp. 825–830. IEEE Service Center (2002)
6. Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II, KanGAL report 200001 (2000)
7. Aguirre, H., Tanaka, K.: Insights on Properties of Multi-objective MNK-Landscapes. In: Proc. 2004 IEEE Congress on Evolutionary Computation, pp. 196–203. IEEE Service Center (2004)
8. Aguirre, H., Tanaka, K.: Working Principles, Behavior, and Performance of MOEAs on MNK-Landscapes. European Journal of Operational Research 181(3), 1670–1690 (2007)
9. Sato, H., Aguirre, H.E., Tanaka, K.: Genetic Diversity and Effective Crossover in Evolutionary Many-objective Optimization. In: Coello Coello, C.A. (ed.) LION 5 2011. LNCS, vol. 6683, pp. 91–105. Springer, Heidelberg (2011)
10. Purshouse, R.C., Fleming, P.J.: Evolutionary Many-Objective Optimisation: An Exploratory Analysis. In: Proc. IEEE CEC 2003, pp. 2066–2073 (2003)