https://doi.org/10.20965/jaciii.2010.p0618

Paper:

Basic Study on Assembling of Objective Functions in Many-Objective Optimization Problems

Shun Otake, Tomohiro Yoshikawa, and Takeshi Furuhashi

Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
E-mail: {otake@cmplx.cse., yoshikawa@cse., furuhashi@cse.}nagoya-u.ac.jp
[Received November 26, 2009; accepted May 25, 2010]

Genetic Algorithms (GAs) have been widely applied to Multiobjective Optimization Problems (MOPs), called MOGA. A set of Pareto solutions in MOPs having plural fitness functions are searched, then GA is applied in a multipoint search. MOGA performance decreases with the increasing number of objective functions because solution space spreads exponentially. An effective MOGA search is an important issue in many objective optimization problems. One effective approach is assembling objective functions and reducing their number, but appropriate assembly and the number of objective functions to be assembled has not been studied sufficiently. Our purpose here is to determine the effects of assembling objective functions by studying assembly effects when MOGA is applied to a simplified Nurse Scheduling Problem (sNSP) in two types of assembly based on objective function meaning and correlation coefficients.

Keywords: nurse scheduling problem, many-objective optimization problem, assembling of objective functions, visualization

1. Introduction

In applying Genetic Algorithms (GAs) [1] to Multiobjective Optimization Problems (MOPs) [2, 3], or MOGA [4], a set of Pareto solutions is searched for that is superior to other solutions in at least one objective function, in MOPs if objective functions are a trade-off. The GA is effective because it features a multipoint search and can search Pareto solutions in one trial. MOGA performance, however, decreases with an increase in the number of objective functions because solution space spreads exponentially [5–7]. Effective searching in Many-objective Optimization Problems (MaOPs) – MOPs with four or five objective functions or more – has become an important issue. Approaches such as parallel distributed processing, which improves computation time, or assembling objective functions and decreasing evaluation space, can ensure an effective search in such problems. Parallel distributed processing, e.g., the master-slave and island models, reportedly effectively decreases computational time and conserves solutions [8–10].

Because appropriately assembling objective functions has yet to be studied, we determined the effects of assembling objective functions to find appropriate assembly and the number of objective functions to be assembled in MOPs.

We studied the effects of assembling objective functions, whether they were indeed assembled, on searching genotype and evaluation space using visualization [11]. We applied Nondominated Sorting Genetic Algorithm-II (NSGA-II) [12], a well-known MOP optimization, to a simplified Nurse Scheduling Problem (sNSP) [13, 14] using two types of assembly based on objective function meaning and correlation coefficients. We compared differences in searching area and the superiority of solutions – for whether objective functions of sNSP are assembled – using visualization.

2. Visualization

We used visualization proposed by H. Ishiguro et al. [11], to visualize evaluation values and genes in the same space by defining the distance between genes, evaluation values, and gene and evaluation values. The distance between genes i, j, encoded in binary, is defined by Hamming distance (G_{ij}) and between evaluation values i, j by Euclidean distance (E_{ij}), and the distance between gene i and evaluation value j by their dissimilarity (D_{ij}). Dissimilarity is defined as follows:

$$D_{ij} = 1 - \frac{\sum_{k=1}^{N_P} \left(G_{ik} - \overline{G_i}\right)\left(E_{jk} - \overline{E_j}\right)}{\sqrt{\sum_{k=1}^{N_P} \left(G_{ik} - \overline{G_i}\right)^2}\sqrt{\sum_{k=1}^{N_P} \left(E_{jk} - \overline{E_j}\right)^2}}, \quad (1)$$

where

$$\overline{G_i} = \frac{1}{N_p}\sum_{k=1}^{N_P} G_{ik},$$

$$\overline{E_i} = \frac{1}{N_p}\sum_{k=1}^{N_P} E_{ik}$$

N_P is the number of visualized solutions.

G^*_{ij}, E^*_{ij}, and D^*_{ij} are calculated by normalizing G_{ij}, E_{ij}, and D_{ij} with their average and variance for use and used in the objective function of Multidimensional Scaling (MDS) [15] in Eq. (2), where g^*_{ij}, e^*_{ij}, and d^*_{ij} are the distance in visualization space. By minimizing objective function $\mathscr{F}$ in Eq. (2), genes and evaluation values are plotted in the same space, while each distance in original space is kept as is, making it possible to grasp relationships of similarities between genotype and evaluation space, because in visualization space, the distance between genes and evaluation values means the distance in genotype or evaluation space and the distance between gene and evaluation means the dissimilarity of space.

$$\mathscr{F} = \frac{\sum_{i<j}^{N_P}(G^*_{ij}-g^*_{ij})^2 + \sum_{i<j}^{N_P}(E^*_{ij}-e^*_{ij})^2 + \sum_{i,j}^{N_P}(D^*_{ij}-d^*_{ij})^2}{\sum_{i<j}^{N_P}(G^*_{ij})^2 + \sum_{i<j}^{N_P}(E^*_{ij})^2 + \sum_{i,j}^{N_P}(D^*_{ij})^2},$$

$$G^*_{ij} = \frac{G_{ij}-\overline{G}}{\sigma_G},$$

$$\overline{G} = \frac{1}{{N_p}^2}\sum_{i=1}^{N_P}\sum_{j=1}^{N_P} G_{ij},$$

$$\sigma_G{}^2 = \frac{1}{{N_p}^2}\sum_{i=1}^{N_P}\sum_{j=1}^{N_P}(G_{ij}-\overline{G})^2 \quad \ldots\ldots\ldots\ldots \quad (2)$$

3. Simplified Nurse Scheduling Problem

The Nurse Scheduling Problem (NSP) is a combinatorial optimization problem for scheduling monthly nursing work, e.g., day, night, midnight, and holiday shifts, while meeting certain conditions. We combined these four shifts and nurse schedules for ten nurses and one week are designed to simplify this simplified nurse scheduling problem.

Conditions include shift sequence, the number of nurses needed per shift, required nursing skills, and the "fairness" of shifts. Insufficiency in work conditions is used as an objective function and NSGA-II applied. Objective functions of sNSP are shown in **Table 1**. All objective functions except for the 4th are degrees of insufficiency for each condition, and the 4th is sufficiency. Obj_4 is multiplied by -1 to minimize objective functions. In actual schedules, no solution satisfies all conditions or is extremely difficult to search for, so the NSP becomes a multiobjective optimization problem.

4. Experiments

NSGA-II is applied to sNSP as shown above and search performance compared, with one in which objective functions are assembled based on meanings or correlation coefficients and with the other not assembled. Pareto solutions acquired each way are visualized and the effect of assembly determined. For GA parameters, solutions numbered 5000, the crossover rate was 1.0, and the mutation rate was (1/gene length). Settings such as coding, generation of initial populations, crossover, and mutation are the same as in [11].

Table 1. Objective functions in sNSP.

Obj_1	number of violations for necessary number of nurses in shifts
Obj_2	number of violations for necessary skill of nurses in shifts
Obj_3	number of prohibited shift sequences
Obj_4	number of recommended shift sequences
Obj_5	number of excessive night and midnight shifts
Obj_6	number of insufficient for fairness of holidays
Obj_7	number of insufficient for fairness of consecutive holidays
Obj_8	number of miss-allocation for continuous 2 holidays in a week

4.1. Experiment 1: Meaning-Based Assembly

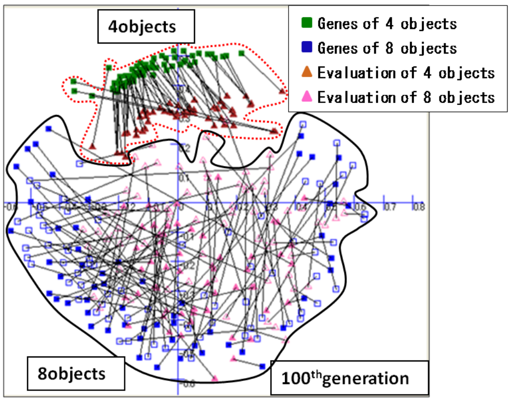
Assembly based on meaning is divided into 4 objective functions, the same as in [11], and is used as in Eq. (3). Obj_1 to Obj_8 were assembled into F_1, F_2, F_3 and F_4 based on four meanings – shifts in a day, shift sequences, fairness and holidays. As in [11], Obj_3 was multiplied by 2 because it is the objective function for prohibited shift sequences and more important than Obj_4 for recommended shift sequences in F_2.

$$\begin{aligned}
\min \quad & F_1(\textit{shifts in a day}) = Obj_1 + Obj_2\\
\min \quad & F_2(\textit{shift sequences}) = 2\times Obj_3 + Obj_4\\
\min \quad & F_3(\textit{fairness}) = Obj_5 + Obj_6 + Obj_7\\
\min \quad & F_4(\textit{holidays}) = Obj_8 \quad \ldots\ldots \quad (3)
\end{aligned}$$

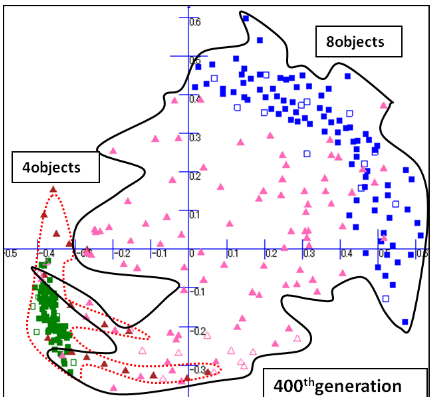
Figures 1(a)–(c) are visualization results at **(a)** 100^{th}, **(b)** 400^{th} and **(c)** 1000^{th} generations. These show only Pareto solutions, searched for four and eight objective functions. The superiority of solutions was evaluated based on eight objective functions. Triangles and squares in figures are inferior solutions. In all figures, squares and triangles are genes and evaluation values. Related genes and evaluation values are connected by lines. If an efficient search is done by assembling objective functions, inferior solutions under four objective functions are fewer than under eight objectives, keeping Pareto solutions broad in the visualization result.

In **Fig. 1(a)**, at the 100^{th} generation, many inferior solutions exist for the eight objective functions, showing that the search in eight objective functions was done to spread and acquire diversified Pareto solutions rather than to acquire better ones. Assembling objective functions worked well to search for superior solutions. **Fig. 1(a)** also shows that searches under four and eight objective functions were done in different areas, because genes are plotted far from each other. Similarly, under four objective functions with a search area narrower than eight, they evolved faster.

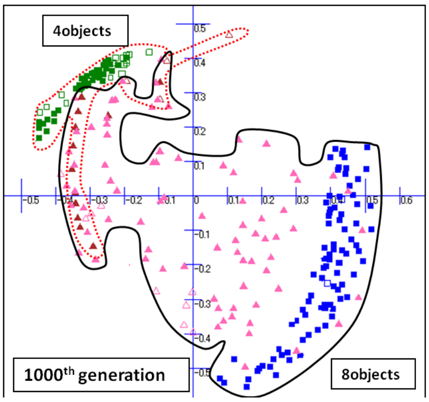
Figure 1(b), at the 400^{th} generation, shows that the number of inferior solutions searched for under eight ob-



(a) 100th generation

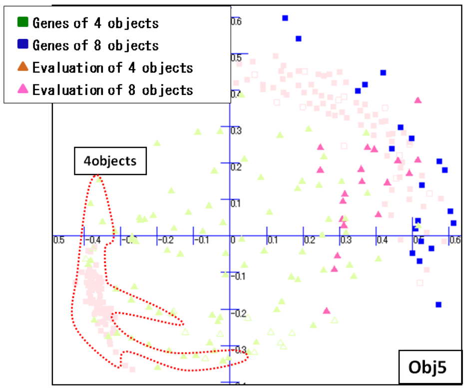


(b) 400th generation

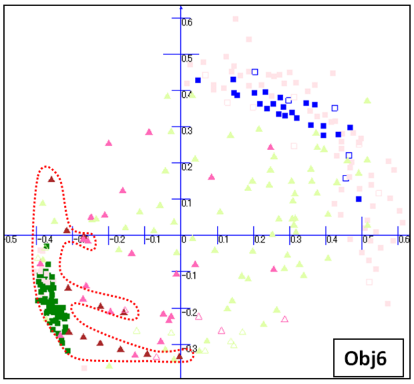


(c) 1000th generation

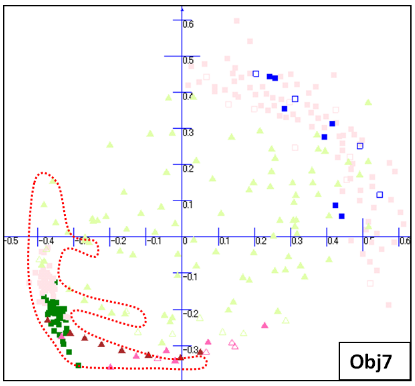
Fig. 1. Visualization result by MDS.



(a) Obj_5



(b) Obj_6



(c) Obj_7

Fig. 2. Best solutions in each objective function.

jective functions decreased and that under four objective functions increased compared to the 100th generation and the area covering evaluation values searched for under four objective functions outlined with a dotted line, and that under eight objective functions outlined with a solid line, which partially overlap. Note that lines between genes and evaluation values are not drawn. At this generation, the search under eight objective functions caught up with that under four, and, compared to that at the 100th generation, genes searched for under eight objective functions are plotted in a narrow area, while that under eight objective functions searches solutions is superior to that searched for under four objective functions at the 1000th generation because, in **Fig. 1(c)**, hardly any inferior solutions are searched for under eight objective functions, in contrast to the large number found under four objective functions – quite difficult from **Fig. 1(a)**.

Figures 2 (a)–(c) show the best solutions at the 400th generation, in which the search under eight objective functions caught up with that under four assembled objectives in Obj_5 to Obj_7, which compose F_3 under four objective functions. By **Fig. 2**, among solutions searched for under eight objective functions, some best solutions exist for each objective function. In contrast, among solutions searched for under four objective functions, no best solution exists for the 5th objective function, although some best solutions are superior for the 6th and 7th objective functions. This implies that Obj_5 and Obj_6 and Obj_5 and Obj_7 are related trade-off. For trade-off between Obj_5 and Obj_6 and between Obj_5 and Obj_7, the search was done to improve Obj_6 and Obj_7 instead of Obj_5 in F_3, and solutions did not evolve to improve Obj_5.

Correlation coefficients between objective functions were calculated using Pareto solutions at the 5000th generation searched for with eight objective functions. Correlation coefficients among Obj_5 through Obj_7 are shown in **Table 2**, showing that correlation values between Obj_5 and Obj_6 and Obj_5 and Obj_7 are negative, so trade-offs appear to exist between them.

These results show that assembly into four objective functions narrows search areas and is effective for progress in the search, especially in early generations, but no assembly yields more superior and varied solutions at large generations if computational time is sufficient. Even if functions have similar meanings, they may be trade-offs, biasing the evolution.

4.2. Experiment 2: Correlation-Based Assembly

Search performance was compared in assembling functions related to positive and negative correlations. Obj_6 and Obj_8, whose correlation coefficient was the largest, i.e., 0.97, were assembled and a search under seven objective functions (i) (Eq. (4)) was done. Similarly, Obj_2 and Obj_5, whose correlation coefficient was the smallest, i.e., -0.73, were assembled and another search under seven objective functions (ii) (Eq. (5)) done.

$$\begin{aligned} &\min \quad F_i = Obj_i \ (1 \le i \le 5) \\ &\min \quad F_6 = Obj_6 + Obj_8 \\ &\min \quad F_7 = Obj_7 \end{aligned} \qquad (4)$$

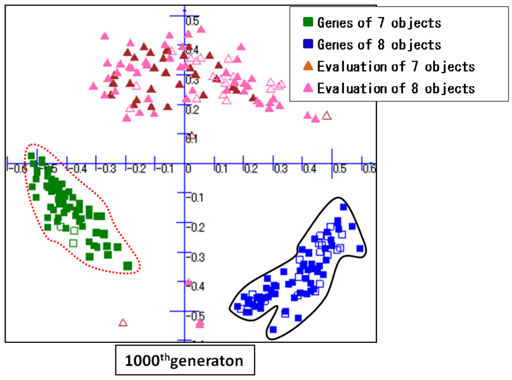
$$\begin{aligned} &\min \quad F_1 = Obj_1 \\ &\min \quad F_2 = Obj_2 + Obj_5 \\ &\min \quad F_i = Obj_i \ (3 \le i \le 4) \\ &\min \quad F_i = Obj_{(i+1)} \ (5 \le i \le 7) \end{aligned} \qquad (5)$$

Figure 3 shows Pareto solutions searched for at the 1000th generation. **Fig. 3(a)** shows only Pareto solutions searched for under seven objective functions (i) and that under eight ones in **Fig. 1(c)**. Similarly, **Fig. 3(b)** shows only those searched for under seven objective functions (ii) and those in **Fig. 1(c)**.

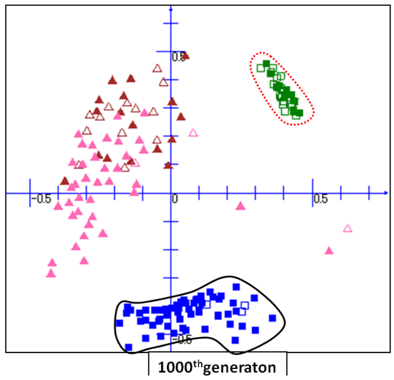
Unlike **Fig. 1(c)**, **Fig. 3(a)** shows many inferior solutions searched for under eight objective functions, and the search under seven objective functions obtained Pareto solutions superior to those under eight objective functions. For gene distribution, little difference existed between seven and eight objective functions, and both searched for broader area and a variety of solutions, and for different

Table 2. Correlation matrix.

	Obj_5	Obj_6	Obj_7
Obj_5	1		
Obj_6	-0.49	1	
Obj_7	-0.32	0.26	1



(a) Obj_6 and Obj_8(correlation: 0.97)



(b) Obj_2 and Obj_5(correlation: -0.73)

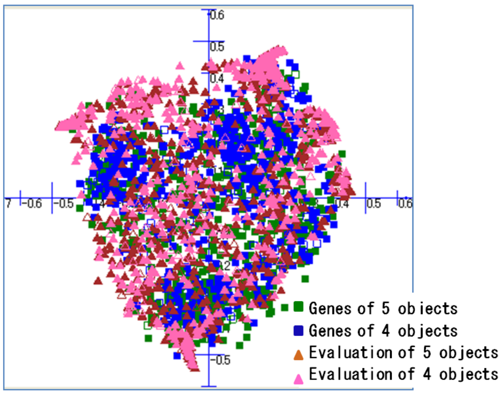
Fig. 3. Assembling objective functions.

areas in genotype space but the same evaluation space. **Fig. 3(b)** shows many inferior solutions searched for under seven objective functions (ii), and that extremely narrow space was searched for more than for eight objective functions.

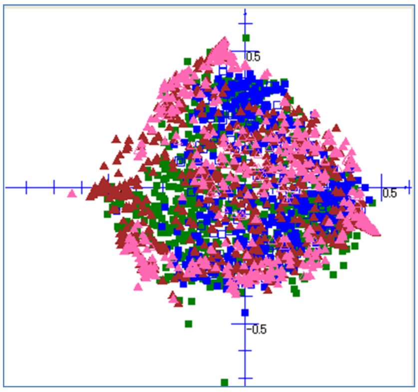
In conclusion, correlation coefficients are useful in assembling objective functions, and search performance is made more efficient by assembling objective functions with high positive correlation coefficients. Convergence at an early stage becomes faster by assembling objective functions correlated negatively, but the variety of genes decreases and the performance of the search may deteriorate at a later stage.

4.3. Experiment 3: Assembling Objective Knapsack Problem Functions

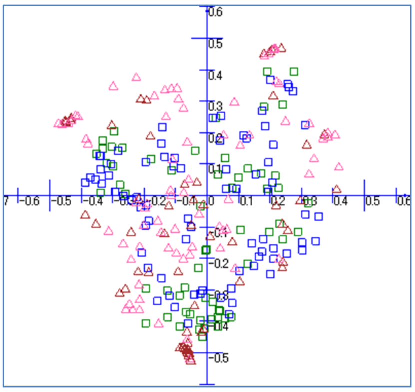
We examined the effects of assembling objective functions based on correlation coefficients in a five objective 0/1 knapsack problem. As in Section 4.2, two pairs of ob-



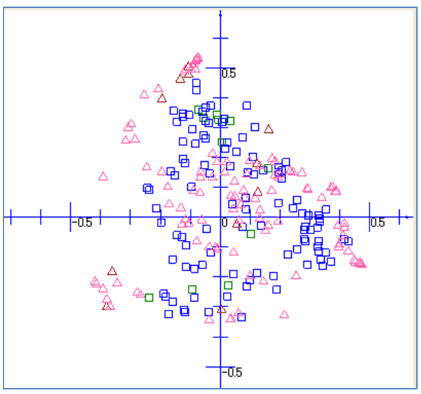
(a) High-positive correlation



(b) High-negative correlation



(c) High-positive correlation (inferior solutions)



(d) High-negative correlation (inferior solutions)

Fig. 4. Assembling objective functions in knapsack problem.

jective functions were assembled with related high positive and negative correlation coefficients, then searches were done under four assembled objective functions.

Figure 4 is the visualization result of Pareto solutions searched for at the 1000th generation. The superiority of all solutions was evaluated based on the five objective functions, and inferior solutions are outlined. **Figs. 4(a)–(c)** show visualization results under four objective functions based on highly positive correlation coefficients and under five functions. **Figs. 4(b)–(d)** show those based on highly negative correlation coefficients and under five functions. **Figs. 4(a)–(b)** show all Pareto solutions and **Figs. 4(c)–(d)** show the inferior solutions in **Figs. 4(a)–(b)**. In **Fig. 4(c)**, the number of inferior solutions searched for under five objective functions is large and many inferior solutions are searched for under four objectives in **Fig. 4(d)**. The distribution of Pareto solutions under four objectives in **Fig. 4(b)** is small.

5. Conclusions

Assembling objective functions in multiobjective optimization problems were studied to determine an assembly methodology. NSGA-II was applied to a simplified nurse scheduling problem to determine the effects of assembling objective functions. First, for assembly into four functions based on meaning, we found that even if functions had similar meanings, they could become a tradeoff biasing evolution. We also studied assembly based on correlation coefficients, finding that they were useful in assembling objective functions. An experiment using a five objective 0/1 knapsack problem showed results similar to the simplified nurse scheduling problem. In future works, we plan to study other types of assembly and correlation coefficient threshold for assembly.

Acknowledgements

This work was supported in part by a Grant-in-Aid for Scientific Research (C) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, Grant number: 22500088.

References:

[1] J. H. Holland, "Adaptation in Natural and Artificial Systems," The Univ. Michigan Press, 1975.

[2] H. Ishibuchi, H. Ohyanagi, and Y. Nojima, "Evolution of Cooperative Behavior in a Spatial Iterated Prisoner's Dilemma Game with Different Representation Schemes of Game Strategies," Proc. of IEEE Int. Conf. on Fuzzy Systems, pp. 1568-1573, 2009.

[3] K. Deb, "Multi-Objective Optimization using Evolutionary Algorithms," Chichester, UK, Wiley, 2001.

[4] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan, "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization, NSGA-II," In KanGAL report 200001, Indian Institute of Technology, Kanpur, India, 2000.

[5] T. Wagner, N. Beume, and B. Naujoks, "Pareto-, Aggregation-, and Indicator-based Methods in Many-objective Optimization," Lecture Notes in Computer Science 4403, Evolutionary Multi-Criterion Optimization-EMO, pp. 742-756, Springer, Berlin, March, 2007.

[6] E. J. Hughes, "Evolutionary Many-Objective Optimization: Many Once or One Many?," Proc. IEEE Congress on Evolutionary Computation (CEC 2005), pp. 222-227, Sep. 2005.

[7] H. Aguirre and K. Tanaka, "Working Principles, Behavior, and Performance of MOEAs on MNK Landscapes," European J. of Operational Research, Vol.181, Issue 3, pp. 1670-1690, 2007.

[8] E. Cantú-Paz, "A Survey of Parallel Genetic Algorithms," Calculateurs Paralleles, Vol.10, No.2, 1998.

[9] L. Nang and K. Matsuo, "A survey on the parallel genetic algorithms," J. of the Society of Instrument and Control Engineers, Vol.33, No.6, pp. 500-509, 1994.

[10] J. Tang, M. H. Lim, Y. S. Ong, and M. J. Er, "Study of Migration Topology in Island Model Parallel Hybrid-GA for Large Scale Quadratic Assignment Problems," Proc. of the Eighth Int. Conf. on Control Automation and System, Vol.3, pp. 2286-2291, 2004.

[11] H. Ishiguro, T. Yoshikawa, and T. Furuhashi, "Visualization of Relationships between Genes and Evaluation Values in GA and Feedback into Genetic Operations," J. of Japan Society for Fuzzy Theory and Intelligent Informatics, Vol.21, No.3, pp. 327-337, 2009. (in Japanese)

[12] K. Deb, "A Fast and Elitist Multi-Objective Genetic Algorithm, NSGA-II," IEEE Trans. on Evolutionary Computation, Indian Institute of Technology, Vol.6, No.2, pp. 182-197, 2002.

[13] N. Ueno and T. Furuhashi, "A Nurse Scheduling Support System with Nurse-in-Chief's Know-How," Fuzzy, Artificial Intelligence, Neural Networks and Computational Intelligence (FAN Symposium '04), Vol.14, pp. 174-177, 2004. (in Japanese)

[14] H. Ishiguro, T. Yoshikawa, and T. Furuhashi, "Study on Visualization of Relationships between Genes and Evaluation Values in Evolution Process of MOGA for NSP," Proc. of Symposium on Evolutionary Computation, pp. 149-155, 2008. (in Japanese)

[15] J. W. Sammon, Jr, "A Nonlinear Mapping for Data Structure Analysis," IEEE Trans. on Computers, Vol.C-18, No.5, pp. 401-409, 1969.

Name:
Shun Otake

Affiliation:
Dept. of Computational Science and Engineering, Nagoya University

Address:
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

Brief Biographical History:

2009- Bachelor of Electrical Engineering and Computer Science, Nagoya University

Main Works:

• S. Otake, T. Yoshikawa, and T. Furuhashi, "Study on assembling of Objective Functions in Multi-objective Optimization Problems," Proc. the 25th Fuzzy System Symposium, 100210(CD-ROM), 2009.

Name:
Tomohiro Yoshikawa

Affiliation:
Dept. of Computational Science and Engineering, Nagoya University

Address:
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

Brief Biographical History:

1997- Ph.D., Dept. of Information Electronics, Nagoya University
1997-1998 Visiting Researcher, University of California at Berkeley
1998-2005 Assistant Professor, Dept. of Electrical and Electronic Engineering, Mie University
2005-2006 COE Designated Associate Professor, COE Project "Frontiers of Computational Science" at Nagoya University
2006- Associate Professor, Dept. of Computational Science and Engineering, Nagoya University

Main Works:

• T. Yoshikawa and T. Furuhashi, "Visualization Techniques for Mining of Solutions," Proc. of the 8th Int. Symposium on Advanced Intelligent Systems (ISIS'07), pp. 68-71, 2007.

Membership in Academic Societies:

• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
• Japan Society of Kansei Engineering (JSKE)
• The Institute of Electrical and Electronics Engineers (IEEE)

Name:
Takeshi Furuhashi

Affiliation:
Dept. of Computational Science and Engineering, Nagoya University

Address:
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

Brief Biographical History:

1988-1990 Assistant Professor of School of Engineering, Nagoya Univ.
1990-2001 Associate Professor of School of Engineering, Nagoya Univ.
2001-2004 Professor of School of Engineering, Mie Univ.
2004- Prof. of Dept. of Computational Science and Eng., Nagoya Univ.

Main Works:

• K. Yamamoto, T. Furuhashi, and T. Yoshikawa, "A Proposal of Visualization Method for Obtaining Interpretable Fuzzy Rules," CD-ROM Proc. of Int. Joint Conf. on Neural Networks, (IEEE-IJCNN2004), 2004.
• K. Yamamoto, T. Yoshikawa, and T. Furuhashi, "A Proposal of Fuzzy Modeling on Fusion Axes Considering the Data Structure," Proc. of Int. Conf. on Fuzzy Systems, (FUZZ-IEEE 2003), pp. 348-353, 2003.
• T. Furuhashi, "Fusion of Fuzzy/Neuro/Evolutionary Computing for Knowledge Acquisition," Proc. of the IEEE, Vol.89, No.9, pp. 1266-1274, 2001.
• N. E. Nawa and T. Furuhashi, "Fuzzy System Parameters Discovery by Bacterial Evolutionary Algorithm," IEEE Trans. on Fuzzy Systems, Vol.7, No.5, pp. 608-616, 1999.
• S. Horikawa, T. Furuhashi, and Y. Uchikawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algorithm," IEEE Trans. on Neural Networks, Vol.3, No.5, pp. 801-806, 1992.

Membership in Academic Societies:

• The Institute of Electrical and Electronics Engineers (IEEE)
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
• The Society of Instrument and Control Engineers (SICE)
• The Institute of Electronics, Information and Communication Engineers
• The Japanese Society for Artificial Intelligence (JSAI)
• The Institute of Electrical Engineers of Japan (IEEJ)