The Effect of User Interaction Mechanisms in Multi-objective IGA

Alexandra Melike Brintrup
Cambridge University, Institute for Manufacturing
Cambridge CB2 1RX UK
+44 (0) 1223 765605
ab702 @ cam. ac. uk

Hideyuki Takagi
Kyushu University, Faculty of Design
4-9-1, Shiobaru, Minami-ku
Fukuoka, 815-8540 JAPAN
+81 92 553 4555
takagi @ design. kyushu-u. ac. jp

ABSTRACT
In this paper four mechanisms, fine and coarse grained fitness rating, linguistic evaluation and active user intervention are compared for use in the multi-objective IGA. The interaction mechanisms are tested on the ergonomic chair design problem. The active user intervention mechanism provided the best fitness convergence but resulted in the least diverse results. The fine grained evaluation provided a good blend of fitness convergence and diversity while the popular coarse grained discrete rating provided poor results. Linguistic evaluation resulted in poor qualitative fitness despite its fast speed of evaluation. The significant differences between interaction mechanisms show the need for further research.

Categories and Subject Descriptors
I.2.8 **[Computing Methodologies]:** Artificial Intelligence – *problem solving, control methods and search.*

General Terms
Algorithms, design

Keywords
IGA, multi-objective optimization, user interaction, design optimization, ergonomic chair design.

1. INTRODUCTION
Interactive Genetic Algorithms (IGA) are a type of GA, that optimize a target system based on human qualitative evaluation. Interaction between a user and the system can proceed in many ways depending on the task domain. In addition to assigning the qualitative fitness of an individual design, the user may intervene by choosing elite designs for survival, by modifying an individual and reinserting it into the population of designs, or by freezing parts of the design with the intention of reducing the search space dimensionality.

Most of previous IGA applications have been based on single objective optimisation frameworks, where the target system is evolved based only on the qualitative evaluation. See comprehensive review of such previous applications in reference [1].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
GECCO'07, July 7–11, 2007, London, England, United Kingdom.
Copyright 2007 ACM 978-1-59593-698-1/07/0007...$5.00.

On the other hand, application of IGA to quantitative optimization problems that have qualitative influences has many advantages. For instance, when IGA is applied to engineering design, interaction with a human evaluator facilitates the generation of solutions that incorporate human expertise without having to explicitly codify them into the optimization platform. In this respect optimization using a multi-objective approach gives the framework an ideal base for compromise decision making when qualitative and quantitative views co-exist. With this view, the multi-objective IGA was proposed by Brintrup *et al.* [2] and tested by applications on floor planning, ergonomic chair design and airfoil design [3]. In the multi-objective IGA qualitative influences are included in the optimisation process as an objective, and are simultaneously optimized with the quantitative objective.

The new multi-objective IGA approach brings on new needs. One of these needs is experimentation on the effect of different user interaction mechanisms when qualitative and quantitative objectives are co-present in the search space.

Contrary to single objective optimisation, multi-objective optimisation algorithms are assessed not only with the fitness convergence of the solutions they provide, but also with the spread of the solutions on the Pareto front. Furthermore the presence of qualitative objective requires user based assessment on the final solutions as the qualitative fitness results may include noise due to human inconsistency and absolute/relative evaluation. Given the recency of the multi-objective IGA, effects of the interaction mechanism on the above mentioned algorithm assessment criteria have not been examined.

Over the years many different evaluation methods were used in traditional single objective IGA. Kim and Cho used fine grained user evaluation in their IGA based fashion design system [4]. Ohsaki and Takagi reported that user fatigue is reduced when users can daringly evaluate individuals using coarse grained discrete evaluation [5]. Active intervention was used by Kishi and Takagi to freeze design features that a user perceives a need to be preserved [6]. These features are prevented from undergoing mutation and recombination. Another active user intervention mechanism is to modify the gene directly [7]. It is reported that active intervention accelerates convergence while reducing the search space. On the other hand user evaluation time increased.

On the other hand, of the limited number of multi-objective approaches that use IGA, discrete rating was the mainly used evaluation method. Brintrup et al, Kamalian et al and Machwe et al used coarse grained discrete rating [3, 8, 9]. With this paper we compare four types of user interaction in the multi-objective IGA by applying them to the ergonomic chair design problem. Section 2 presents a brief overview of the multi-objective IGA and the ergonomic chair design problem and describes the different interaction mechanisms used. Section 3 presents the experimental procedure, Section 4 lays out a discussion on preliminary results while Section 5 concludes on the findings.

2. ERGONOMIC CHAIR DESIGN

2.1 Multi-objective IGA

Our multi-objective IGA used in this paper is based on a modified version of a popular multi-objective optimization algorithm, the non-dominated sorting GA 2 devised by Deb *et al.* [10].

This algorithm enhances the non-domination-based sorting techniques by introducing the concepts of elitism and diversity. Elitism ensures the preservation of globally good solutions from generation to generation. Diversity ensures achieving a set of well-spread solutions in the objective space. Elitism in the NSGA 2 is achieved by combining parent and offspring populations before sorting them for non-domination. Diversity in the NSGA 2 is achieved by favoring well spread solutions in the search space. In the multi-objective IGA the fitness evaluation for the qualitative objective is obtained from the user interactively, whereas the quantitative fitness of a solution is assessed by the built-in fitness function. Readers may refer to [11] for a detailed description of the multi-objective IGA.

2.2 Ergonomic Chair Design Problem

Posture is simply the position of the body during an activity (including resting). Ergonomic chair design is concerned with accommodating a given posture comfortably while satisfying the user's visual aesthetic criteria. The designs are sometimes evaluated by measurements on how the chair fits to a given percentage of parts of the body in a given posture (i.e. reclining, working etc.), sometimes by live experiments in which a sitter's feeling of comfort is recorded, or often by a mixture of both. In the present study we define the ergonomic chair design problem as the problem of finding an optimum set of parameters that control the shape of the chair with respect to a given posture.

Two types of objectives are included in the problem: a qualitative objective based on the user's evaluation of how suitable the chair looks for a particular posture, and a quantitative objective that measures how closely the chair fits to the sitter's posture. The intention here is not to replace live comfort experimentation but to provide a complimentary tool that can illustrate and narrow down the available search space to the user. A detailed examination of the ergonomic chair design problem can be found in [3].

The parameters that constitute the gene are shown on Fig. 1 (a). The qualitative objective function is defined as:

$$f_1(c,l) = f_{user}(c,l) \tag{1}$$

where the user given fitness evaluation denotes the user's comfort (c) and liking (l) of the design.

The parameters used in the quantitative objective function are shown on Fig. 1 (b). The quantitative objective function is defined as:

$$f(x,y) = \left(\sum_{i=1}^{n} \sqrt{(ch\_spine_{ix} - p\_spine_{ix})^2 + (ch\_spine_{iy} - p\_spine_{iy})^2} \right) +$$

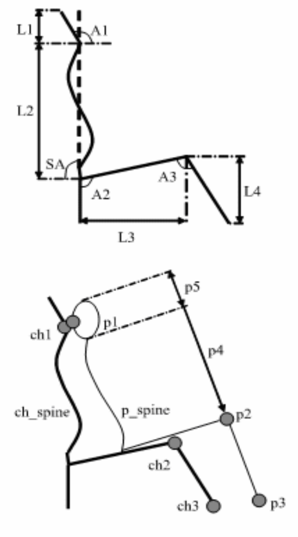

Figure 1. (a) chair design parameters (b) objective function parameters

$$\sqrt{(ch1_x - p1_x)^2 + (ch1_y - p1_y)^2} + \sqrt{(ch2_x - p2_x)^2 + (ch2_y - p2_y)^2} + \sqrt{(ch3_x - p3_x)^2 + (ch3_y - p3_y)^2} \tag{2}$$

where, *ch_spine* and *p_spine* are *n* B-spline points denoting a B-spline approximation of the chair's backrest spine and the person's spine, $ch1$ is the chair's headrest midpoint, $p1$ is the person's head midpoint projected to the back of head, $ch2$ is the chair's knee point, $p2$ is the person's knee point, $ch3$ is the chair's end of footrest point, $p3$ is the person's end of footrest point.

The following constraints were coded:

constraint[0] = 90 < chair knee angle < person knee angle (in degrees)

constraint[1] =90 < chair headrest angle < 180

constraint[2] =90 < chair angle of backrest to seat < 180

The seated human body's positional data is a predefined and fixed model, which is based on the 99 percentile white male body [13].

2.3 Interaction Mechanisms

Fig. 2 shows the graphical user interface where user interaction is pursued. At any given experiment only one of the interaction mechanisms is active. The following mechanisms are used:

The coarse grained discrete rating involves the user assigning a discrete numerical fitness score between 1 and 10 to the individual design produced by the multi-objective IGA. Any two designs may have the same score.

The fine grained discrete evaluation involves the user giving a finer fitness score between 1 and 100 to the individual design produced. The score is given by the use of a slider, and the numerical equivalent of the position of the slider is not shown to the user.

With linguistic evaluation the user gives the algorithm commands to "destroy", "keep", "promote", or "bonus" a particular design. The "destroy" command erases the design from the population. The "keep" command gives an average fitness value to the design

Table 1. Performances of interaction mechanisms

	Coarse grained discrete rating	Fine grained discrete user evaluation	Linguistic evaluation	Active user intervention
Average quantitative fitness	3300, risk %15	3092.8, risk %15	2760.82, risk %5	2580.48, risk %5
Average qualitative fitness	6.25, risk %15	5.56, risk %15	6.75, risk %15	3.83, risk %15
Diversity rank	2, risk %5	1, risk %5	4, risk %5	3, risk %5
User preference rank	4, risk %5	3, risk %5	2, risk %5	1, risk %5
Speed of evaluation rank	3, risk %15	2, risk %5	1, risk %15	4, risk %5

whereas the "promote" and "bonus" commands assign higher levels of fitness values respectively. Similar to previous fitness evaluation mechanisms designs may have the same evaluation.

During the active user intervention, the user directly modifies the design parameters that correspond to the gene through the active user intervention dialogue box shown on Fig. 2 (b). the user can preview changes or rollback to the original design. After the modification the user assigns the design a linguistic fitness value.

3. EXPERIMENTAL EVALUATION

The multi-objective IGA used a mutation rate of 0.01 and, and one-point simulated binary crossover with a rate of 0.9. Population size was 12. Human users were 2 females and 4 males of ages 24–33, whose expertise ranged over product design, design engineering and aerospace engineering. The users continued to run the programs for six generations. Each of the runs started with a different set of initial individuals.

Each user conducted one test for each interaction mechanism until significance could be concluded using the Wilcoxon signed rank test. This nonparametric pair observation test is used to compare the performance assessment obtained from each interaction mechanism with the other and rank the mechanisms in order of success. Table 1 shows the performance of each interaction mechanism including the risk factors obtained from the Wilcoxon signed rank test. A risk factor 5% means that the comparative round was in favor of the winner with a significance of 95%.

The solutions reached through different interaction mechanisms were assessed by three metrics: final average qualitative fitness, final average quantitative fitness convergence, and diversity. In addition, a preference ranking of users for interaction mechanism was taken and the speed of user evaluation with each mechanism was noted.

Many diversity metrics assess diversity by comparison of obtained solutions to those in the known Pareto front [12]. However, we cannot pinpoint the Pareto front as there exists no consistent relationship between the qualitative and quantitative objective functions. Therefore, Deb et al.'s diversity metric [10] was used to assess the spread of solutions which does not require a priori knowledge of the Pareto front.

4. DISCUSSION

The significant differences among the results obtained through the use of different interaction mechanisms show that the choice of interaction mechanism affects the performance of the algorithm.

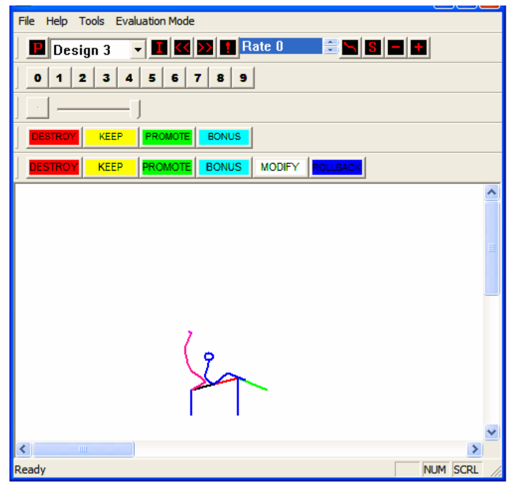

Figure 2. Graphical user interface

Active user intervention is the most preferred interaction mechanism as well as the best performing mechanism in terms of final average qualitative and quantitative fitness values it provided. Speed of evaluation and modest diversity are the drawbacks of this mechanism. Speed of evaluation was understandably the worst as users spent much time modifying the designs instead of simply evaluating them. The considerable slow down in evaluation hints that this method is more suitable for optimization problems that can be expected to optimize in a relatively small number of generations. Most designs were modified to bring them closer to the conceptual shape the user envisaged. This resulted in designs similar to one another and hence the diversity metric measured to be low performing.

Linguistic evaluation was the fastest evaluation method. Although it performed second best for final average quantitative fitness, the final average qualitative fitness was the worst. A poor performing qualitative fitness average is interesting since this method was the second mostly preferred after the active user interaction. It may be speculated that when many designs have the same discrete fitness value evaluation the algorithm might pursue the quantitative objective more profoundly than the qualitative objective.

Similarly the fine grained discrete fitness assignment was the third user preferred interaction mechanism but gave the second best qualitative fitness values and resulted in the most diverse set of designs. Although providing the second best speed of evaluation, it performed poorly on the final average quantitative fitness. It is possible that the wider numerical range resulted in a search space that held designs that were wide spread. Comparing the diversity provided by fine grained fitness evaluation and coarse grained user evaluation it can be said that the more fine grained is the fitness the more information the algorithm has on the spread of the solution and the crowding distance calculation is therefore more informative.

Interestingly the worst performing interaction method was the popular coarse grained rating where the users directly quantified the qualitative objective.

Overall users seem to prefer interaction mechanisms that do not require direct quantification by the users. Although the users prefer manual modification of designs most this is also the interaction mechanism that results in the most fatigue due to its slow speed.

It can be hypothesized that in a solution search where diversity is of importance and the quantitative objective requires a large number of iterations active user intervention is best used once per every set number of generations instead of continuous intervention. In this case fine grained discrete evaluation seems to be a good contender with a fine balance of multiple objective and diversity satisfaction.

It is important to note that user evaluated qualitative objective does not necessarily contradict or complement to the quantitative objective in the ergonomic chair design problem. Further experimentation is necessary for examining interaction mechanisms in which contradictory or complimentary objectives are in act.

5. CONCLUSION

This paper examined the effects interaction mechanisms for multi-objective interactive genetic algorithms where objectives are qualitative and quantitative. Four interaction mechanisms were assessed through diversity and fitness convergence the resulting designs as well as user preference and speed of evaluation.

Fine grained evaluation provided the most diverse results and second best qualitative fitness objective satisfaction. Linguistic evaluation was the fastest evaluation method and whereas active user intervention, provided the best qualitative and quantitative fitness average, however was the slowest evaluation method. The significant differences between interaction mechanisms hint the need for further research.

6. REFERENCES

[1] Takagi, H. Interactive Evolutionary Computation: Fusion of the capabilities of EC Computation and Human Evaluation, *Proceedings of the IEEE*, 89:9, pp. 1275-1296, 2001.

[2] Brintrup, A. M., Tiwari A., and Gao J., A Framework for Handling Qualitative and Quantitative Data in an Evolutionary Multiple Objective Design Space, *Int. J. of Computational Intelligence*, 1:4, pp.282-289, 2004.

[3] Brintrup, A. M., *Handling Qualitativeness in Design Optimisation using Interactive and Multi-objective Genetic Algorithms*, Ph.D. Thesis, Cranfield University, UK, 2007.

[4] Kim, H. S. and Cho, S. B., Application of Interactive Genetic Algorithm to Fashion Design, *Engineering Applications of Artificial Intelligence*, 13:6, pp. 635-644, 2000.

[5] Ohsaki, M., Takagi, H., and Ohya K., An Input Method Using Discrete Fitness Values for Interactive GA, J. *of Intelligent and Fuzzy Systems*, 6:1, pp.131-145, 1998

[6] Takagi, H. and Kishi, K*., On-line Knowledge Embedding for Interactive EC-based Montage System*, Int. Conf. on Knowledge-Based Intelligent Information Engineering Systems (KES'99): Adelaide, Australia, pp.280-283, 1999.

[7] Unemi, T. *SBART 2.4: an IEC Tool for Creating 2D Images, Movies, and Collage*, Leonardo, 35:2, pp.189-191, MIT Press, 2002.

[8] Kamalian, R., Takagi, H., and Agogino, A., *Optimized Design of MEMS by Evolutionary Multi-Objective Optimization with Interactive Evolutionary Computation*, Genetic and Evolutionary Computation Conference (GECCO2004): Seattle, pp. 1030-1041, 2004.

[9] Machwe, A., Parmee, I. C., and Miles, J., *Integrating Aesthetic Criteria with a User-centric Evolutionary System via a Component-based Design Representation*, Int. Conf. on Engineering Design (ICED2005), Melbourne, 2005.

[10] Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T., *A Fast Elitist Non dominated Sorting Genetic Algorithm for Multi Objective Optimization: NSGA-2*, Parallel Problem Solving from Nature (PPSN2000): Paris, pp. 858-862, 2000.

[11] Brintrup, A., Ramsden, J., and Tiwari, A., An Interactive Genetic Algorithm Based Framework for Handling Qualitative Criteria in Design Optimization, *J. of Computers in Industry*, 58:3, pp.279-291, 2007.

[12] Grosan, C., Oltean, M., and Dumitrescu, D. *Performance metrics for multiobjective optimization evolutionary algorithms*, *Conference on Applied and Industrial Mathematics* (CAIM2003): Oradea, 2003

[13] Tilley A. R., *The Measure of Man and Woman: Human Factors in Design*, Revised Edition, Wiley & Sons, 2002.