Reducing Premature Convergence Problem through Numbers Structuring in Genetic Algorithm

Abu Bakar Md Sultan†, Ramlan Mahmud†, Muhammad Nasir Sulaiman†

† Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Malaysia

Summary
Genetic Algorithm (GA) has been widely used in many types of optimization problem. Premature convergence was the main problem for GA performance cause by lower diversity of the population. Obviously maintaining higher diversity is important to obtained better result. Classic GA representing the solution using binary system that is 0 and 1. From the origin idea of binary system, we present odd and even number representation for GA's population. The odd and even number structure simulate the classic binary representation 0 and 1. The intention of proposed representation is to increase the diversity as well as preventing premature convergence. This short paper evaluates the performance of GA responding to new representation scheme.It was evaluated to solving timetabling problem and the result obtained is promising.

Key words:
Genetic Algorithm, Premature Convergence, Diversity, Timetabling, Optimization

1. Introduction

Genetic Algorithm is a stochastic search algorithm and a general-purpose optimization method based on Darwin Theory of evolution. GA operates on a population of solutions represented by some encoding. Each individual in the population is known as chromosomes that represent a set of solution. New solutions were obtained by combining different chromosomes to produce new better offspring or by altering existing member of the population (mutation). A series of evolution then takes place by first evaluating the fitness of each chromosome (individual) and select the fittest to survive to next generation.

A major problem in GA is that classic GA have tendency to converge to local optima. This premature convergence is caused by several algorithmic features, particularly selection pressure and too high gene flow between population member [1]. Population Diversity is undoubtedly a key issue in the performance of GA [2]. A common hypothesis is that higher diversity is important to avoid premature convergence and to escape local optima. Numerous methods have been applied to combat premature convergence. For instance [3] proposed a measurement to guide diversity search within the population. This technique emphasize on balancing between crossover and mutation under certain threshold values.

This paper presents the evaluation of unique numbering techniques for representing population in GA to prevent premature convergence and maintaining higher diversity in population. Section 2 described about problem description under study. Section 3 discussed about our proposed model. Section 4 present the readjusment procedure to maintain the representing sturuture . Section 5 describes result from simple experimental process over timetabling problem instances and we conclude in section 6.

2. Problem Description

The problem under study was a University Course Timetabling Problem (UCTP). A general timetabling problem consists of assigning set of lectures to rooms and timeslots and venue according a number of rules. In the optimization rules are usually divided in two groups: hard constraints and soft constraints. Hard constraints should under no circumstances be violated while the number of violation of soft constraint is as low as possible. Violation of hard constraints make the timetable infeasible while violations of soft constraint affect the quality of timetable thus it should be minimized. The work that was presented in this paper addresses UCTP Problem with the emphasis to determine the performance result on classic representation versus unique odd-even number reperesentation over GA population.

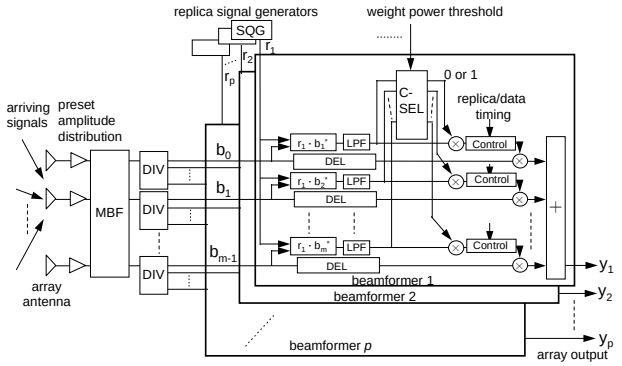

Fig. 1 Proposed beam former.

Manuscript received April 5, 2007
Manuscript revised April 25, 2007

2.2 Equations

If a displayed equation needs a number, place it flush with the right margin of the column (e.g., see Eq. 1).

$$y_i(N) = \sum_{n=0}^{m-1} w_n(N) b_n(N) = \sum_{n=0}^{m-1} \overline{b_n^*(N) r_i(N)} \cdot b_n(N) \quad (1)$$

3. Proposed Model

The odd and even number represents solution for course timetabling instances during the execution of the algorithm. Timetabling problem involved assigning event to timeslots and venues that satisfy all hard constraints and minimize as much soft constraint. These kinds of problems mostly exist in educational institution such as school and university. The concept behind the odd and even number to this problem is to differentiate between morning and afternoon session with the emphasis on increasing population diversity. The sequence of odd and even number are side by side along the chromosomes.

Figure 1 illustrates the chromosome structure in population. The o indicates odd numerical value inside the gene whereas e indicates even numeric values inside the gene.

o	e	o	e	…….

Figure 1**.** Chromosomes structure

To ensure the structures of chromosome are always in the position of holding odd and even number for respective genes, heuristic are employed for every iteration to make simple readjustments. For instances, if the gene supposed to hold odd number but after iteration process the value became even, then 1 will be added into to gene value. This method also applied to even number if changed.

Standard GA parameters were given for each series of experiment. Details description of GA parameters is given in Table 1. Series of experiments and comparative studies was being conducted to compare the performance of algorithm over two different representations that are OEGA (GA + odd-even representation) and basic GA representation methods. Classic GA used decimal number without specific concern on population structure.

Table 1
Genetic Algorithm Parameters

Items	Parameters
Size of population	40
Number of Generation	100
Crossover Rate	0.7
Mutation Rate	0.3

Both representations schemes were applied to the same GA algorithm. Figure 2 described the structure of GA used for this research.

```
Procedure GA
        Begin
            Generate random p(t)
            Evaluate p(t)
            Repeat
            Begin
                t  = t+1;
                crossover p(t);
                mutation  p(t);
                readjustment p(t);
                evaluate p(t);
            end
    end
```

Figure 2. Genetic Algorithm

The algorithms begin by generating initial random solution for UCTP problem. The rest of the processes are similar to almost basic GA system that required mutation and crossover. Readjustment procedures just only apply to odd-even representation only (OEGA). The result obtained from both experimental processes
are compared.

4. Readjusment Procedure

The important element inside the algorithm is the readjustment procedure. This procedure is necessary to avoid the structure of odd and even representation change

after each iteration. Even though the numbers keep on changing the readjustment procedure would readjust back any value that are not suppose loacate at respective gene type.

The readjustment procedure consist of simple repair instruction to readjustments back the value inside the gene if the number that suppose to odd appear even vice verse inside respective genes after each iteration. This procedure ensures the sequence of odd and even remains at respective gene after iterations. Figure 3 decribed the readjustment procedure.

```
Procedure Readjustment ()
Do
  Gene_location (o/e)
   If NumInGen not= Gene_location
   Add 1 to NumInGen
  While not end Chrom
```

Figure 3. Readjustment Procedure

5. Result

Table 2 shows the result from 20 tests run made for each method. Here we assumed that if the generation failed to converge before 1000 generation this means that it was trapped in premature convergence problems.

Table 2
Result from Test Run

	GA	OEGA
Number of Trials Run	20	20
Optimum Solution	3	12

The odd and even number has shown significant effect to the GA performance. The result obtained for OEGA shows significantly increase on number of optimum result.

6. Conclusion

This short paper presents unique population representation for genetic algorithm that is odd and even number representation. The major aim was to overcome the problem of premature convergence by increasing population diversity. The result obtained shows significantly increases in optimum solution for timetabling problem instances under study. This is the initial result from ongoing research. Even though the result is promising, extensive study is still being conducted to further conforming the result. Future efforts are also looking into the possibilities to adapt similar structure on different types of optimization result.

Acknowledgments

This research is sponsored by grants IRPA-04-02-04-02-EA001 from Ministry of Science, Technology and Innovation Malaysia (MOSTI).

References

[1] Qu, R. 2002. Case-based Reasoning for Course Timetabling Problem. Ph.D Thesis. Department of Computer Science, University of Nottingham UK.
[2] Popovic, D. 1997. Retaining Diversity of Search Point Distribution Through a Breeder Genetic Algorithm For Neural Network Learning. IEEE International Conference on Neural Network 1: 495-498. Houston Texas.
[3] Ursem, U K. 2002. Diversity-guided Evolutionary Algorithms. Lecture Notes in Computer Science. 2439:462-471.

Abu Bakar Md Sultan is a PhD student at faculty of Computer Science and information Technology, Universiti Putra Malaysia (UPM). He holds a Master Degree in Software Engineering from UPM. His research interest includes artificial intelligence, scheduling and software agents.

Ramlan Mahmod hold a PhD from University of Bradford, United Kingdom. Currently, he is a Associate Professor at Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, His research area are artificial intelligence.

Md Nasir Sulaiman obtained his PhD from Loughborough University in 1994. Currently he is a Associate Professor in Department of Computer Science, Faculty of Computer Science and Information Technology, University Putra Malaysia. His research interest is artificial intelligence, neural networks, data mining, pattern recognition and parallel computing

Mohd Rizam Abu Bakar obtained her PhD from the University of Bradford. Currently, he is a lecturer at the Faculty of Science, Universiti Putra Malaysia.