COMPARISON FOR SELECTION TECHNIQUES IN GENETIC ALGORITHM

MUHAMAD AZREE BIN MAT SAID

A thesis submitted in fulfillment of the requirements for the award of the degree of Bachelor of Computer Science (Software Engineering)

Faculty of Computer Systems & Software Engineering
University College of Engineering & Technology Malaysia

November, 2006
ABSTRACT

The purpose of this project is to make a comparison for three selection techniques in Genetic Algorithm. The Genetic Algorithm has been implemented in the previous module for Chess Tournament Management System. Based on the previous system, the selection method only used randomize. By this project module, only three selection methods were used for the comparison. They are Roulette Wheel, Steady-State and Rank selection. The result of this comparison has determined the appropriate selection for Genetic Algorithm implementation in Chess Tournament Management System. This will help Chess Tournament Management System to provide a better optimize schedule.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>DECLARATION OF ORIGINALITY AND EXCLUSIVENESS</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGEMENT</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ABSTRAK</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLE</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td></td>
<td>LIST OF ABBREVIATIONS</td>
<td>xii</td>
</tr>
<tr>
<td></td>
<td>LIST OF APPENDICES</td>
<td>xiii</td>
</tr>
</tbody>
</table>

1 INTRODUCTION

1.1 Introduction

1.1.1 Chess Tournament Scheduling System

1.1.2 Genetic Algorithm

1.1.3 Current Chess Tournament Scheduling System

1.2 Problem Statement

1.3 Objective

1.4 Scope

1.5 Thesis Organization

2 LITERATURE REVIEW

2.1 Methodology
2.1.1 Genetic Algorithm Approach 5
2.1.2 Initialization 6
2.1.3 Evaluation 7
2.1.4 Selection 7
  2.1.4.1 Roulette Wheel Selection 8
  2.1.4.2 Rank Selection 8
  2.1.4.3 Steady State Selection 9
2.1.5 Recombination 10
2.1.6 Crossover 10
  2.1.6.1 Single Point Crossover 10
  2.1.6.2 Multi Point Crossover 11
2.1.7 Mutation 12
2.1.8 Replacement 12

3 METHODOLOGY 13
3.1 Waterfall Model 13
3.2 Project Analysis 14
  3.2.1 Fitness Function 14
  3.2.2 Crossover 15
  3.2.3 Mutation 15
3.3 Project Design 16
  3.3.1 Selection Technique 19
    3.3.1.1 Roulette Wheel Selection 19
    3.3.1.2 Rank Selection 21
    3.3.1.3 Steady-State Selection 22
3.4 Interface Design 23
3.5 Developmental Issues 27
  3.5.1 Development 28
  3.5.2 Software and Hardware 28

4 RESULT AND DISCUSSION 30
4.1 Expected Result 30
4.2 Testing Result 30
4.3 Further Research 37
4.4 Constraint 37

5 CONCLUSION 38
5.1 Summary 38
5.2 Achieved Objectives 39
5.3 Lesson Learnt 39

REFERENCES 41
APPENDICES 42
Appendix A (Project Gantt Chart) 42
Appendix B (Testing Result) 44
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Listing of Item</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Hardware Requirement</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Software Requirement</td>
<td>29</td>
</tr>
<tr>
<td>4.1</td>
<td>Comparison Result for Three Selections in Case 1, Case 2 and Case 3</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison Result for Three Selections in Case 4, Case 5 and Case 6</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison Result for Three Selections in Case 7, Case 8 and Case 9</td>
<td>35</td>
</tr>
<tr>
<td>FIGURE NO</td>
<td>TITLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1</td>
<td>Genetic Algorithm high-level flow</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Population of Chromosomes</td>
<td>7</td>
</tr>
<tr>
<td>2.3</td>
<td>Roulette wheel selection</td>
<td>8</td>
</tr>
<tr>
<td>2.4</td>
<td>Situation before ranking (graph of fitness)</td>
<td>9</td>
</tr>
<tr>
<td>2.5</td>
<td>Situation after ranking (graph of order numbers)</td>
<td>9</td>
</tr>
<tr>
<td>2.6</td>
<td>Single Point Crossover</td>
<td>11</td>
</tr>
<tr>
<td>2.7</td>
<td>Multi Point Crossover</td>
<td>11</td>
</tr>
<tr>
<td>2.8</td>
<td>Mutation</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Waterfall Model</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>Fitness Function</td>
<td>14</td>
</tr>
<tr>
<td>3.3</td>
<td>Crossover of chromosome</td>
<td>15</td>
</tr>
<tr>
<td>3.4</td>
<td>Mutation on chromosome</td>
<td>16</td>
</tr>
<tr>
<td>3.5</td>
<td>Project Design</td>
<td>17</td>
</tr>
<tr>
<td>3.6</td>
<td>Roulette Wheel Flow</td>
<td>19</td>
</tr>
<tr>
<td>3.7</td>
<td>Probability Number</td>
<td>20</td>
</tr>
<tr>
<td>3.8</td>
<td>Sum Probability Number</td>
<td>20</td>
</tr>
<tr>
<td>3.9</td>
<td>Condition Set</td>
<td>20</td>
</tr>
<tr>
<td>3.10</td>
<td>Rank Selection Flow</td>
<td>21</td>
</tr>
<tr>
<td>3.11</td>
<td>Sum of Rank Number</td>
<td>22</td>
</tr>
<tr>
<td>3.12</td>
<td>Condition Set</td>
<td>22</td>
</tr>
<tr>
<td>3.13</td>
<td>Steady-State Selection Flow</td>
<td>23</td>
</tr>
<tr>
<td>3.14</td>
<td>Sort</td>
<td>23</td>
</tr>
<tr>
<td>3.15</td>
<td>Front Page Interface</td>
<td>24</td>
</tr>
<tr>
<td>3.16</td>
<td>Main Program interface</td>
<td>24</td>
</tr>
<tr>
<td>3.17</td>
<td>Setting Dialog Box</td>
<td>26</td>
</tr>
<tr>
<td>3.18</td>
<td>Generate Form</td>
<td>27</td>
</tr>
</tbody>
</table>
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>BCS</td>
<td>Bachelor Computer Science</td>
</tr>
<tr>
<td>FSKKP</td>
<td>Fakulti Sistem Komputer &amp; Kejuruteraan Perisian</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>KUKTEM</td>
<td>Kolej Universiti Kejuruteraan &amp; Teknologi Malaysia</td>
</tr>
<tr>
<td>RAM</td>
<td>Random Access Memory</td>
</tr>
<tr>
<td>to</td>
<td>Population of Chromosomes</td>
</tr>
<tr>
<td>VB.Net</td>
<td>Visual Basic .Net</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>OO</td>
<td>Object Oriented</td>
</tr>
</tbody>
</table>
## LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Project Gantt Chart</td>
<td>42</td>
</tr>
<tr>
<td>B</td>
<td>Testing Result</td>
<td>44</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

In this chapter, an introduction of the system follows by brief introduction of the genetic algorithm, problem statement, objectives and scope of the system.

1.1 Introduction

This chapter discusses a brief introduction about genetic algorithm and Chess Tournament Management System.

1.1.1 Chess Tournament Management System

A Chess Tournament Management System (CTMS) is developing to organize and manage the tournament. This system use Genetic Algorithm to optimize schedule with constraint. The resource constraint will be set by user such as venue constraint and arbiter constraint. After the system running the result will be show the match information such as the venue, arbiter, time, date and players. The main purpose of this system is to help admin to manage a tournament with automatically generate a tournament schedule.
1.1.2 Genetic Algorithm

Genetic Algorithms (GAs) first proposed and analyzed by John Holland in 1975, are optimization techniques based on natural evolution and adaptation. The basic flow algorithm is started with a set of population. Each possible solution within the population is also called chromosome. Each chromosome is assigned with fitness value. Solutions from population are taken and used to construct a new population. This process repeats again based on generation by hope the new population will be fitter than the old population [2].

The genetic algorithms have a cycle of process that hopes this process can get a fitness value of chromosome in population. The process is beginning with initialization by randomly creating chromosomes in population. Follow by evaluate a fitness value each chromosome. After evaluation, each chromosome will be selected by using several selection methods. After selected, two parents will be crossover to form a new offspring or new children. Crossover will be single-point crossover and multi-point crossover. Follow by mutation that mutates a chromosome to a new offspring. The new chromosomes will replace to population in the next generation. This process is looping to get best population to the next generation. This project has used three method of selection are listed below:

i. Roulette Wheel Selection
The idea of Roulette Wheel Selection is the chromosomes with higher fitness value have more chances to be selected than chromosomes with lower fitness value. The chromosomes with lower fitness value have a change to be selected but the change to be selected is too low.

ii. Rank Selection
The key of rank selection is ranks the population first and then every chromosome receives fitness value determined by this ranking. After ranking, each chromosome does not differ so much for a fitness value. All the chromosomes have a same chance to be selected.
iii. Steady State Selection
The main idea of this type of selecting is each chromosome with higher fitness is selected for create a new offspring. This new chromosome will replace the some bad chromosomes with lower fitness. The others population will survive to next generation.

1.1.3 Current Chess Tournament Scheduling System

For the last year, Chess Tournament Management System had developed by student and has functionality to record the profile of each player, their rating and games played. This system also can be capable to record all games player throughout the tournament. This system also can perform pairing for different type of tournament system available. [1]

1.2 Problem Statement

A Chess Tournament Management System (CTMS) has managed and organizes the tournament. That system was use by administrator. CTMS provide function for record rating, game played and profile every player. On this system apply genetic algorithm technique to generate optimize schedule for chess tournament, but in the selection technique their only using randomly selection. In genetic algorithm, the selection technique have many type of selection and every different technique have own method to select the chromosome. Random selection is one selection technique and does not know this technique is appropriate in selection technique to apply in CTMS.
1.3 Objective

The objectives of the project are:

i. To make a comparison for selection techniques in Genetic Algorithm

ii. To develop a prototype for the comparison using the Chess Tournament Management System.

1.4 Scope

i. Three techniques are use for comparison:
   a. Roulette Wheel Selection
   b. Rank Selection
   c. Steady-State Selection

ii. The techniques are applied for Chess Tournament Management System (Scheduling Module).

iii. The tool use in the development of this system is the Visual Studio .Net and Microsoft SQL Server 2003 (MSSQL) for storage a data.

1.5 Thesis Organization

This thesis divided into five chapters and every chapter has a different explanation. Chapter 1 is the introduction of the thesis. Introduction to the project is presented along with the project's problem statement, objectives of the project and the scopes of the project. Chapter 2 is the literature review. Research and literature review related to the project is presented. Some methods or methodologies that are applied in this study reviewed. Chapter 3 will present the research method and procedures like project analysis, design and methodology are presented. Chapter 4 will be discuss the testing result of the system and discussion on the result. Chapter 5 is concludes the studies that are carried out in the thesis. It consists of summary and lesson learnt from the developed project system.
CHAPTER 2

LITERATURE REVIEW

In this chapter elaborate about the technique, algorithm, and tools that are use to develop the system.

2.1 Methodology

2.1.1 Genetic Algorithm Approach

This chapter explains a detail about Genetic Algorithm. Base on Darwin’s theory about Genetic Algorithm, the algorithm is started with a set of population represent about a set of possible solution. Each possible solution is also called chromosome and this chromosome is assigned with a fitness value base on the fitness function. New populations are constructing by get a solution from taken a one population. This is selected by a hope that the new population will be fitter than the old one. Solutions, which are selected to construct new solutions (offspring), are selected according to their fitness - the fitter they are, the more chances they can have for reproduction. The offspring replace the old population and a generation is complete. This process is repeated until certain criteria are met. [2]. Figure 2.1 show how the Genetic Algorithm process.
2.1.2 Initialization

Initialization is a starting point for the algorithm with creating the initial population. At the beginning, the population is done by randomly creating the chromosomes. But if processes already have done a looping, the fitness chromosome will replace the old chromosomes. The initialization process can be illustration into Figure 2.2.
2.1.3 Evaluation

After the initialization, the next process is the chromosomes evaluation to determine how good the solution is. This step involves decoding the chromosome into the variable space of the problem and then checking the result of the problem using parameters.

2.1.4 Selection

Process of selection is to pick out chromosome from population. Chromosomes are selected based on their fitness that means how well they solve the problem at hand. Too many methods are present in selection section. The basically way is select by randomly, but is no guarantee the future generation will increase in fitness. Several method of selection is use in genetic algorithm is Roulette Wheel, Rank selection, Steady-State and Tournament selection.
2.1.4.1 Roulette Wheel Selection

Base on the roulette-wheel selection, parent is selected base on their fitness. The better chromosome is more change to be selected. Figure 2.3 show the roulette-wheel chart, chromosome 1 is more large size better than other chromosomes. Chromosome 1 has a bigger change to be selected. Other chromosome can be selected but probability to be selected is little change than chromosome 1.

![Roulette wheel selection](image)

Figure 2.3 Roulette wheel selection [3]

2.1.4.2 Rank Selection

Roulette Wheel become not appropriate to use when one of the chromosomes is 90% fitness compare with other chromosomes, because other chromosomes will have very few chances to be selected. In this situation, Rank Selection is appropriate technique to be used [3]. Rank Selection ranks the population first and then every chromosome receives fitness value determined by this ranking. The worst will have the fitness 1, the second worst 2 etc. and the best will have fitness \( N \) (number of chromosomes in population). Figure 2.4 show before ranking and Figure 2.5 show how the situation changes after changing fitness to the numbers determined by the ranking.
Now all the chromosomes have a chance to be selected. However this method can lead to slower convergence, because the best chromosomes do not differ so much from other ones.

2.1.4.3 Steady State Selection

In a steady-state genetic algorithm one member of the population is changed at a time. To perform selection a member of the population is chosen according to its fitness. It is copied and the copy mutated. A second member of the population is selected which is replaced by the mutated string. In crossover two members of the population are chosen, a single child created which replaces a member of the population. Any selection method can be used to select the individual for mutation or the parents. There are a number of different replacement strategies. [3]

i. Replace worst
ii. Replace a randomly chosen member
iii. Select replacement using the negative fitness
2.1.5 Recombination

During the recombination, pair of chromosomes are recombined, modified the chromosomes and then place back into the population as the next generation. Basically in this section the set of chromosome are refer as a parent to reform a new children. With the recombination they have some possible operators including crossover and mutation. The result after recombination is a new population of chromosomes.

2.1.6 Crossover

In this phase, the idea is both parent are selected base on the fitness function and then combine together to form a new child or offspring. There are two of methods to use to perform a crossover.

2.1.6.1 Single Point Crossover

The idea for single point crossover is before to form a two new child or offspring (new chromosomes). Two parents from the fitness chromosomes will be crossover base on parents 1 split some part of body and replace into parent 2. Same things for parent 2, they must split same part of body and replace into parent 1. The two children will form base on single crossover from two parents in the same part. Figure 2.6 show the example for single point crossover.
2.1.6.2 Multi Point Crossover

Same things for the multi point crossover but they have a few different for crossover. The concept in multi point crossover is, to form a new offspring or children they must crossover two parts in parent A and parent B. Figure 2.7 can explain how the multi point crossover process. From parent A, part A1 and A2 will change position with part B1 and B2 in parent B. The point is both of this parents change two part of body in each others.
2.1.7 Mutation

The mutation can perform after chromosomes have done crossover process. Basically mutations randomly change their gene to form new child. Refer as a binary code, mutation randomly pick and change small number of bit in chromosomes.

![Mutation Diagram](image)

**Figure 2.8 Mutation [2]**

2.1.8 Replacement

After all the process and perform a new chromosomes, this chromosomes will replace old chromosomes into a population. This process will repeat again until a number of generations.
CHAPTER 3

METHODOLOGY

In this chapter discusses about analysis of the system, design and development. That is how to implement the technique into the system.

3.1 Waterfall Model

The software process model are implied in the project development is waterfall model.

Figure 3.1 Waterfall Model [4]
Based on the Figure 3.1 Waterfall Model [9], the phases of waterfall model are including Requirement Analysis Definition, System and Software Design, Implementation and Unit Testing, Integration and System Testing, and Maintenance.

3.2 Project Analysis

Before develop this system they are required to analysis how the current Chess Tournament Management System manage the chess tournament by using genetic algorithm approach. Base on the previous system has developed. Start with population represents a group of chromosome object. After the initialization a fitness value will be assign and the populations are going to crossover, mutation and selection process. The detail about Literature Review was discussed in chapter 2.

3.2.1 Fitness Function

For the previous system the fitness function uses to determine the candidate schedule that will be select to form a new generation of solution. The fitness value of candidate schedule base on the time span for the schedule and number of resources consumes. The fitness value base on the shorter time is better than longer time. Then a few resource are use are more fitness value than a using more resources such as venue and arbiter. Formula to calculate a fitness value is show as Figure 3.2.

\[
F = \frac{T - t}{\sum (R_i \times m_i)}
\]

- \(F\) = fitness value
- \(T\) = Expected time span
- \(t\) = Solution time span
- \(R_i\) = Resource \(i\)
- \(R_m\) = Resource \(i\) multiplier

**Figure 3.2 Fitness Function [1]**