A STUDY ON GENETIC ALGORITHM

THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT
FOR THE DEGREE OF BACHELOR OF SCIENCE
UNIVERSITY TECHNOLOGY MARA

By

NURIZWANI BT MOHD NOOR
2002616522
BSc (Hons) INTELLIGENT SYSTEM
FACULTY OF INFORMATION TECHNOLOGY
AND QUANTITATIVE SCIENCE

OCTOBER 2004
A STUDY ON GENETIC ALGORITHM

NURIZWANI BT MOHD NOOR
2002616522

Supervised By

PUAN ZAIDAH IBRAHIM

COPYRIGHT © UiTM
APPROVAL

Name of Candidate: Nurizwani Bt Mohd Noor

Title of Thesis: A Study on Genetic Algorithm

Approved By:

Pn Zaidah Bt Ibrahim
Supervisor

Prof. Madya Dr Azlinah Hj Mohamed
Thesis Advisor

Date: 11th October 2004
DECLARATION

I hereby declare that this research report together with all of its contents is no other than those of my own work, except for some information taken and extracted from other sources that have been quoted respectively.

11th October 2004

NURIZWANI BT MOHD NOOR
(2002616522)
ACKNOWLEDGEMENT

Praise to Allah the Almighty for providing me the strength and knowledge throughout this period of study and for all that has bestowed on me. It is with His power the study is completed.

I would like to express a very special gratitude and thankfulness to my supervisor Puan Zaidah Bt Ibrahim who is giving me reliance and shrewd guidance that have profited me very much. She who is also gave lots of support in this knowledge quest.

My special appreciation to my thesis advisor, Prof. Madya Dr Azlinah Hj Mohamed, who had gave lots of knowledge for me through the completion of this project. This appreciation also to my Programme Head of Bachelor (Hons) Intelligent System, En Mohd Zaki Zakaria who has earned good management services throughout my degree program.

To all my friends who are simply too numerous to indicate, thanks for your time, understanding and buttress. May Allah will reciprocate you for all those things.

Last but not least, to my beloved family for their love, never-ending support and encouragements in making sure I do my best in completing this research study. Thank you very much and may Allah bless you all.
ABSTRACT

Genetic Algorithm is one of Artificial Intelligence's approaches that adapt the idea of simulating the nature's process of evolution and natural selection through computer implementation. Genetic Algorithms are one of a number of Artificial Intelligence tools used problem such as optimization, scheduling, data fitting, clustering, trend spotting, and path finding. This study is focused on the applications that used genetic algorithm such as university timetabling (course and exam), job scheduling, prediction and classification. It also focused on the identifying the best parameters used in genetic algorithms and identifying the strengths and weaknesses of the genetic algorithm. This research is focused on to study about genetic algorithms and identifies the applications that use genetic algorithm, and identify the strengths and weaknesses of genetic algorithm.
# TABLE OF CONTENT

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>i</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>ii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1.0 INTRODUCTION

1.1 INTRODUCTION 1

1.2 RESEARCH OBJECTIVE 3

1.3 RESEARCH SCOPE 3

1.4 RESEARCH SIGNIFICANCE 4

1.5 GENETIC OPTIMIZATION 4

1.5.1 Global Optimization 5

1.5.2 Constrained Optimization 5

1.5.3 Combinatorial Optimization 6

1.5.4 Multi-Objective Optimization 7
2.0 LITERATURE REVIEW

2.1 INTRODUCTION 9

2.2 DEFINITION OF GENETIC ALGORITHM 10

2.3 BASIC PRINCIPLES OF GENETIC ALGORITHM 17

2.3.1 Simple Genetic Algorithm and Basic Principle 18

2.4 HOW GENETIC ALGORITHMS WORK 20

2.4.1 Initialization 20

2.4.2 Selection 21

2.4.3 Crossover 23

2.4.4 Mutation 25

2.5 MEMETIC ALGORITHM 27

2.6 REVIEW OF DEVELOP SYSTEM USING GENETIC ALGORITHM 28

2.6.1 A Co-evolving Timeslot/Room Assignment Genetic Algorithm Technique for University Timetabling 28

2.6.2 A Multi-objective Genetic Algorithm for the Class/Teacher Timetabling Problem 29

2.6.3 A Simple Approach to Protein Structure Prediction Using Genetic Algorithms 30
4.0 ANALYSIS AND FINDINGS

4.1 INTRODUCTION

4.2 WHAT IS GENETIC ALGORITHM?

4.3 GENETIC ALGORITHM AND ITS PARAMETERS
   4.3.1 Encoding of a Chromosome
   4.3.2 Population Size
   4.3.3 Crossover and Mutation Probability
   4.3.4 Selection Schemes

4.4 APPLICATION OF GENETIC ALGORITHM

4.5 RECOMMENDATION THE PARAMETERS OF GENETIC ALGORITHM
   4.5.1 Types of Problems
   4.5.2 Population Representation
   4.5.3 Fitness Function
   4.5.4 Termination Criteria
   4.5.5 Selection
   4.5.6 Crossover
   4.5.7 Mutation Probability

4.6 STRENGTHS AND WEAKNESSES OF GENETIC ALGORITHM

4.7 CONCLUSION
5.0 CONCLUSION AND RECOMMENDATION

5.1 INTRODUCTION 66
5.2 CONCLUSION 67
5.3 RECOMMENDATION 68
5.4 SUMMARY 68

REFERENCES 70
## LIST OF TABLE

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>41</td>
</tr>
</tbody>
</table>

**Summarize Table for Application That Use Genetic Algorithm**
## LIST OF FIGURE

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Simple Genetic Algorithm</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>One-point Crossover</td>
<td>24</td>
</tr>
<tr>
<td>2.3</td>
<td>Two-point Crossover</td>
<td>24</td>
</tr>
<tr>
<td>2.4</td>
<td>Last Bit Mutation</td>
<td>27</td>
</tr>
<tr>
<td>3.1</td>
<td>Research Methodology</td>
<td>37</td>
</tr>
<tr>
<td>3.2</td>
<td>A Basic Genetic Algorithm</td>
<td>50</td>
</tr>
<tr>
<td>4.1</td>
<td>Representation of a Chromosome</td>
<td>57</td>
</tr>
<tr>
<td>4.2</td>
<td>Examples of Chromosome with Binary Encoding</td>
<td>57</td>
</tr>
<tr>
<td>4.3</td>
<td>Examples of Chromosome with Permutation Encoding</td>
<td>58</td>
</tr>
<tr>
<td>4.4</td>
<td>Examples of Chromosome with Value Encoding</td>
<td>58</td>
</tr>
<tr>
<td>4.5</td>
<td>Examples of Chromosome with Tree Encoding</td>
<td>59</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

Since the 1960's, there has been an increasing interest in imitating living beings to develop powerful algorithms for difficult optimization problems. Independently, in the 1960's to 1970's, three methodology of evolutionary algorithm has been developed, that are, genetic algorithm (GA), evolutionary strategies (ES) and evolutionary programming (EP), differentiated by the process of evolving its population.

GAs, as powerful and broadly applicable stochastic search and optimization techniques, are perhaps the most widely known types of evolutionary computation methods today. GA of evolutionary computation is inspired by Darwin's theory of evolution. A GA generates a population of possible solutions encoded as chromosomes, evaluates their fitness and creates a new population by applying genetic operators which are crossover and mutation. By repeating this process over many generations, the genetic algorithm has five basic components which are (Gen and Cheng, 2000):
i. a genetic representation of solutions to the problem;

ii. a way to create an initial population of solutions;

iii. an evaluation function rating solutions in terms of their fitness;

iv. genetic operators that alter the genetic composition of children during reproduction;

v. values for the parameters of genetic algorithms.

GAs uses fitness values of individual chromosomes to carry out reproduction. As reproduction takes place, the crossover operator exchanges parts of two single chromosomes and the mutation operator changes the gene value in some randomly chosen location of the chromosome. After a number of successive reproductions, the less fit chromosomes become extinct, while those best fit gradually come to dominate the population.

GAs works by discovering and recombining schemata which is good ‘building blocks’ of candidate solutions. The GAs does not need knowledge of the problem domain, but it requires the fitness function to evaluate the fitness of a solution.

Solving a problem using GAs involves defining constraints and optimum criteria, encoding the problem solutions as chromosomes, defining a fitness function to evaluate a chromosome’s performance and creating appropriate crossover and mutation operators.
In GAs, accumulated information is explored by the selection mechanism, while new regions of the search space are explored by means of genetic operators. In conventional genetic algorithms, the crossover operator is used as the principal operator and the performance of a genetic system is heavily dependent on it. The mutation operator which produces spontaneous random changes in various chromosomes is used as a background operator.

1.2 RESEARCH OBJECTIVE

The main objectives of this research are:

i. Study about genetic algorithms and identify the applications that use genetic algorithm;

ii. Identify the strength and weakness of genetic algorithm.

1.3 RESEARCH SCOPE

There are many types of application that used genetic algorithm to solve the problem for each application. In this paper, the scope of the research involves applications in university timetabling (course and exam), job scheduling, prediction and classification.
1.4 RESEARCH SIGNIFICANCE

Nowadays, genetic algorithm is a popular method to solve problem of optimization. There are many types of application used genetic algorithm to solve the problem such as timetabling, scheduling, automatic programming, protein structure prediction, immune system models, population genetic models and robot control. One of the most successful areas for genetic algorithm applications includes the problem of scheduling resources. Scheduling problems are complex and difficult to solve. They are usually approached with a combination of search techniques and heuristics.

1.5 GENETIC OPTIMIZATION

Optimization deals with problems of minimizing or maximizing a function with several variables usually subject to equality and/or inequality constraints. It plays a central role in operations research, management science, and engineering design. Many industrial engineering design problems are very complex and difficult to solve using conventional optimization techniques. In recent years, genetic algorithms have received considerable attention regarding their potential as a novel optimization technique. Because of their simplicity, ease of operation, minimal requirements, and parallel and global perspective, genetic algorithms have been applied successfully in a wide variety of problem domains (Gen and Cheng, 2000).
1.5.1 Global Optimization

Global optimization utilizes techniques that can distinguish between the global optimum and numerous local optima within a region of interest. Global optimization problems usually take the form of unconstrained optimization; that is, the problem is one of minimizing or maximizing a function in the absence of restriction. Conventional global optimization methods can roughly be categorized into two classes which are: (1) deterministic methods and (2) stochastic methods. Genetic algorithms have been fairly successful at solving problems of the type that are too ill-behaved, non-differentiable, and discontinuous for conventional hill-climbing and derivative-based techniques. Since the emergence of genetic algorithms in the early 1970s, global optimization has been one of the major targets, and a lot of effort has been devoted to developing powerful algorithms for global optimization problems. The usual way of applying genetic algorithms to solving global optimization problems is to encode each decision variable as a bit using either standard binary coding or Gray coding (Gen and Cheng, 2000).

1.5.2 Constrained Optimization

Non-linear programming (or constrained optimization) deals with the problem of optimizing an objective function in the presence of equality and/or inequality constraints. Non-linear programming is an extremely important tool used in almost every area of engineering, operations research, and mathematics because many
practical problems cannot be modeled successfully as a linear program. The central difficulty in applying genetic algorithms to these types of problems is how to handle constraints, because genetic operators used to manipulate the chromosomes often yield infeasible offspring. Several techniques have been proposed to handle constraints with genetic algorithms which are: rejecting methods, repairing methods, and penalty methods (Gen and Cheng, 2000).

1.5.3 Combinatorial Optimization

Combinatorial optimization problems are characterized by a finite number of feasible solutions. Although the optimal solution to such finite problem can be found by a simple enumeration, in practice it is frequently impossible, especially for practical problems of realistic size where the number of feasible solutions can be extremely high. Combinatorial optimizations contain a huge body of problems with different features and properties. Although these problems are quite different from each other, the essence of the problems can be characterized as one of the following types (Gen and Cheng, 2000):

- To determine a permutation of some items associated with the problem;
- To determine a combination of some items;
- To determine both permutation and combination of some items;
- Any one of the above subject to some constraints.
1.5.4 Multi-objective Optimization

Multiple-objective optimization problems have received increased interest from researchers with various backgrounds since early 1960. In a multi-objective optimization problem, multiple objective functions need to be optimized simultaneously. There is does not necessarily exist a solution that is best with respect to all objectives because of incommensurability and confliction among objectives. There has been an applying genetic algorithm to solving the multi-objective optimization problem, known as evolutionary multi-objective optimization or genetic multi-objective optimization (Gen and Cheng, 2000).

1.6 CONCLUSION

The evolutionary approach of artificial intelligence is based on the computation models of natural selection and genetic known as evolutionary computation. Evolutionary computation combines genetic algorithms, evolution strategies and genetic programming.

Genetic algorithms are a very powerful tool. However, coding the problem as a bit string may change the nature of the problem being investigated. There is always a danger that the coded representation represents a problem that is different from the one the researcher wants to solve.
The need to solve optimization problems arises in almost every field and in particular, is a dominant theme in the engineering world. Many optimization problems from the engineering world are very complex in nature and quite difficult to solve by conventional optimization techniques. Genetic algorithms have received considerable attention for complex problems and have been applied successfully in the area of industrial engineering (Gen and Cheng, 2000).
CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In nature, individuals that are best suited to the environment will win the competition for scanty resources. An individual's survival capacity is determined by various features that characterize it. The features in turn are determined by the individual's genetic content. The set of genes controlling features form the chromosomes, which are the keys for the individuals to survive in a competitive environment. Specifically, the joint action of nature selection and the recombination of genetic materials during reproduction is the driven force of evolution.

Since only the fittest individuals survive and reproduce, the genes of weaker individuals would die out gradually. If the environment does not change during the process, we can imagine that finally it will converge to a state that every individual has the fittest or the best genes.

Inspired by this natural evolution process, the use of analogies of natural behavior led to the development of the so-called "evolutionary algorithms" (EAs). They usually have four main elements which are an encoding structure that will be
replicated, operators to affect the individuals of a population, a fitness function that indicates how good an individual is and a selection mechanism. Genetic algorithm (GA) is one of the main paradigms within EAs. They operate on a population of individuals, each presenting a possible solution to a given problem. Each individual is assigned a fitness score based on the fitness function. A selection mechanism selects highly fit individuals to reproduce the offspring by “cross breeding” (crossover) and mutation techniques.

Genetic algorithms are not guaranteed to reach the global optimum, but they are generally good at finding an acceptable solution during an acceptable amount of time. They are mainly designed to solve optimization problems. However, when cooperating with other techniques it can also deal with problems with constraints. It is so robust that it can be applied to a wide range of problem areas. It also has good performance when solving some difficult problems which no existing specialized techniques can perform well on. Even if such specialized techniques exist, improvements could be made by hybridizing them with GAs.

2.2 DEFINITION OF GENETIC ALGORITHM

Genetic Algorithm (GA) model introduced by John Holland in 1975 triggered a wide interest in the application of these types of heuristic, which mimic the natural
evolutionary characteristic present in the biological species with the purpose of solving optimization problem efficiently (Carrasco & Pato, 2001).

GAs is a branch of artificial intelligence's stochastic search technique that is widely used in the field of optimization. The idea of GAs was first introduced by John Holland back in the 1970s and was later popularized by David Goldberg who was able to solve a difficult problem in the control of gas-pipeline transmission (Randy and Sue Ellen Haupt, 1998).

According to Biesbroek (1999), the study of GAs was originated from studies done on cellular automata, which took place when Goldberg was at the University of Michigan. Until the early 1980s, researches regarding GA were mainly theoretical, with only few real applications. Other famous experts such as De Jong and Hollstien who looked into similar areas in GA had also devoted their efforts in bringing GA to an upper level, in hope that this approach will be able to provide a different perspective in solving optimization problems. De Jong's work mainly attempted to capture the features of the adaptive mechanisms in the family of GAs while Hollstien's work focused on the effect of different selection schemes and mating strategies on the performance of GAs. The following years have witnessed the development of GA in the sciences, engineering and the business world where these algorithms had been successfully applied to problems such as optimization, scheduling, data fitting, clustering, trend spotting and path finding.