Real Jumping Gene Genetic Algorithm (RJGGA) for Evolutionary Multiobjective Optimization Problems

Submitted to
Department of Computer Science

in Partial Fulfillment of the Requirements for the Degree of Master of Philosophy

by

Kazi Shah Nawaz Ripon

October 2006
Abstract

Real-world applications often involve the satisfaction of multiple conflicting objectives simultaneously. The problems with two or more objectives are called “multiobjective” problems that usually have a set of trade-off solutions. Evolutionary multiobjective optimization is currently gaining significant attention from researchers in various fields due to its effectiveness and robustness in solving multiobjective problems. The Jumping Gene Genetic Algorithm (JGGA) is a relatively recent multiobjective evolutionary algorithm (MOEA), which is proven to be in the leading edge over many existing MOEAs. In its initial formulation, the search space solutions are binary-coded and it inherits the customary problems of conventional binary-coded genetic algorithm (BCGA). For this reason, it is very important to remodel the JGGA into a real-coded genetic algorithm (RCGA).

The primary focus of this thesis is to present the Real Jumping Gene Genetic Algorithm (RJGGA) as an improvement on the JGGA to remove the shortcomings associated with the original JGGA. In order to introduce the concept and verification of RJGGA, this thesis comprised the basic concept of jumping gene phenomenon and the necessary modification required for transforming JGGA into RJGGA. To validate the performance of the proposed approach, we systematically compare the RJGGA with various existing MOEAs using some carefully chosen benchmark test functions. Simulation results justify that the proposed approach is able to generate non-dominated solutions with a wider spread along the Pareto-optimal front and better address the issues regarding convergence and diversity in multiobjective optimization.

This thesis also presents the applications of RJGGA on two different practical multiobjective problems to justify its efficiency in solving such problems. First, we illustrate the application of the proposed method in case of multiobjective data clustering. Traditional clustering algorithms are limited to optimizing single clustering objective only and often fail to detect meaningful clusters because most practical data sets are characterized by a high-dimensional, inherently sparse, data space. Inspired by the
inherent multiobjective nature of data clustering, we devise an explicit multiobjective data clustering approach to deal with a range of complex data properties including overlapping clusters, elongated cluster shapes, unequal sized clusters, etc. For these reasons, we present a novel multiobjective evolutionary clustering approach using the Variable-Length Real Jumping Genes Genetic Algorithms (VRJGGA), which evolves clustering solutions using multiple clustering criteria without any prior knowledge of the actual number of clusters.

Finally, we investigate the application of RJGGA in solving the multiobjective Job-Shop Scheduling Problem (JSSP). During the last decades, various evolutionary approaches were investigated in solving the JSSPs. However, most of them are limited to a single objective, which is not suitable for real-world scheduling problems as they naturally involve multiple objectives. The proposed scheduling approach heuristically searches for the near-optimal schedules that optimize multiple criteria simultaneously. Experimental results for both of the above two applications confirm that the RJGGA consistently finds solutions with better diversity and convergence as compared to other existing evolutionary based approaches.
Table of Contents

Abstract ................................................................................................................................. ii

Acknowledgements ........................................................................................................... iv

Table of Contents................................................................................................................ v

List of Tables ...................................................................................................................... ix

List of Figures..................................................................................................................... x

Abbreviations..................................................................................................................... xiii

List of Publications .......................................................................................................... xiv

Chapter 1 Introduction ....................................................................................................... 1

1.1 Problems and Motivations ......................................................................................... 1

1.2 Aim ......................................................................................................................... 6

1.3 Objectives ............................................................................................................. 6

1.4 Organization of Thesis............................................................................................... 7

Chapter 2 Multiobjective Evolutionary Algorithm (MOEA)............................................. 8

2.1 Overview ................................................................................................................. 8

2.2 Multiobjective Optimization Problem (MOOP)....................................................... 8

2.3 Concept of Domination .......................................................................................... 10

2.4 Non-Dominated Set............................................................................................... 11

2.5 Pareto Front........................................................................................................... 11

2.6 Goals of MOEA..................................................................................................... 12

2.7 Why Evolutionary Algorithms in Multiobjective Optimization............................ 12

2.8 History of Multiobjective Evolutionary Algorithm (MOEA).................................. 13

Chapter 3 Real Jumping Gene Genetic Algorithm (RJGGA)........................................ 19

3.1 Introduction............................................................................................................. 19

3.2 Jumping Gene Genetic Algorithm (JGGA)........................................................... 19
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1 Jumping Gene Transposition</td>
<td>21</td>
</tr>
<tr>
<td>3.2.2 Effects of Jumping Operations on Multiobjective Functions</td>
<td>23</td>
</tr>
<tr>
<td>3.3 Real-Coded Genetic Algorithm</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Design Issues for Real Jumping Gene Genetic Algorithm</td>
<td>25</td>
</tr>
<tr>
<td>3.4.1 Crossover in RJGGA</td>
<td>27</td>
</tr>
<tr>
<td>3.4.2 Mutation in RJGGA</td>
<td>29</td>
</tr>
<tr>
<td>3.4.3 Jumping Gene Operators in RJGGA</td>
<td>30</td>
</tr>
<tr>
<td>3.5 Performance Metrics and Test Problems</td>
<td>32</td>
</tr>
<tr>
<td>3.5.1 Metrics Evaluating Convergence</td>
<td>34</td>
</tr>
<tr>
<td>3.5.1.1 Generational Distance (GD)</td>
<td>34</td>
</tr>
<tr>
<td>3.5.1.2 Set Coverage Metric (C)</td>
<td>34</td>
</tr>
<tr>
<td>3.5.2 Metrics Evaluating Diversity</td>
<td>35</td>
</tr>
<tr>
<td>3.5.2.1 Spacing (S)</td>
<td>35</td>
</tr>
<tr>
<td>3.5.2.2 Spread (Δ)</td>
<td>35</td>
</tr>
<tr>
<td>3.5.3 Metric Evaluating both Closeness and Diversity</td>
<td>36</td>
</tr>
<tr>
<td>3.5.4 Test Problems</td>
<td>37</td>
</tr>
<tr>
<td>3.6 Performance Assessments and Comparisons</td>
<td>38</td>
</tr>
<tr>
<td>3.7 Conclusions</td>
<td>53</td>
</tr>
<tr>
<td><strong>Chapter 4 Multiobjective Evolutionary Clustering</strong></td>
<td>54</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>4.2 Necessity of Multiobjective Clustering</td>
<td>54</td>
</tr>
<tr>
<td>4.3 Related Works</td>
<td>57</td>
</tr>
<tr>
<td>4.4 Variable-Length Real Jumping Gene Genetic Algorithm</td>
<td>59</td>
</tr>
<tr>
<td>4.4.1 Chromosome and Population Initialization</td>
<td>61</td>
</tr>
<tr>
<td>4.4.2 Cluster Regulation</td>
<td>61</td>
</tr>
<tr>
<td>4.4.3 Fitness Evaluation and Selection Scheme</td>
<td>62</td>
</tr>
</tbody>
</table>
5.8 Experimental Results ............................................................................................................. 99
  5.8.1 Benchmark Problem ...................................................................................................... 99
  5.8.2 Experimental Evaluation and Discussions ..................................................................... 100
  5.8.3 Single Objective Context ............................................................................................. 101
  5.8.4 Multiple Objective Contexts ....................................................................................... 103

5.9 Conclusions ....................................................................................................................... 108

Chapter 6 Conclusions and Future Works .............................................................................. 109

6.1 Conclusions ....................................................................................................................... 109

6.2 Future Works .................................................................................................................... 111

References ............................................................................................................................... 98

Appendix A ............................................................................................................................... 107
List of Tables

Table 3.1 Test problems used in this study........................................38
Table 3.2 Comparison between RJGGA and JGGA.................................45
Table 4.1 Characteristics of datasets...................................................67
Table 4.2 Comparative results of VRJGGA and NSGAII..........................75
Table 4.3 Comparison of VRJGGA and non-evolutionary approaches.........76
Table 5.1 A 6x6 job-shop scheduling benchmark problem......................81
Table 5.2 Benchmark problems for multiobjective JSSP..........................99
Table 5.3 Comparison with some GA based algorithms.........................102
Table 5.4 Comparison of the makespan obtained by various evolutionary
and heuristic methods.................................................................103
Table 5.5 Scheduling results of test problems......................................104
List of Figures

Fig. 2.1 Hypothetical trade-off solutions for a car buying decision-making problem .......... 9
Fig. 2.2 Pareto Front ......................................................................................................... 11
Fig. 3.1 Cut and paste transposition: (a) Same chromosome; (b) Different chromosome .... 21
Fig. 3.2 Copy and paste transposition: (a) Same chromosome; (b) Different chromosome ... 22
Fig. 3.3 Flowchart of JGGA ........................................................................................... 22
Fig. 3.4 Flowchart of transposition .................................................................................. 22
Fig. 3.5 Probability distribution of offspring; (a) Distant parents;
    (b) Closely spaced parents .......................................................................................... 28
Fig. 3.6 Flowchart of transposition in RJGGA ................................................................32
Fig. 3.7 Convergence metric \((GD)\): (a) Mean; (b) Variance .......................................... 40
Fig. 3.8 Set domination graphs based on the \(C\) metric .................................................. 41
Fig. 3.9 Diversity metric \((\Delta)\): (a) Mean; (b) Variance .................................................. 42
Fig. 3.10 Spacing \((S)\) metric for diversity ................................................................. 43
Fig. 3.11 Hyper volume ratio \((HR)\) metric for both convergence and diversity:
    (a) For all algorithms; (b) Only for close competitors ................................................. 44
Fig. 3.12 (a)--(e): Final non-dominated solutions of RJGGA and JGGA for test functions .... 46
Fig. 3.13 Set of non-dominated solutions searched by each MOEA for test function ZDT1:
    (a) For all comparing MOEAs; (b) Only closely contested MOEAs ............................... 48
Fig. 3.14 Set of non-dominated solutions searched by each MOEA for test function ZDT2:
    (a) For all comparing MOEAs; (b) Only closely contested MOEAs ............................... 49
Fig. 3.15 Set of non-dominated solutions searched by each MOEA for test function ZDT3:
    (a) For all comparing MOEAs; (b) Only closely contested MOEAs ............................... 50
Fig. 3.16 Set of non-dominated solutions searched by each MOEA for test function ZDT4
    (a) For all comparing MOEAs; (b) Only closely contested MOEAs ............................... 51
Fig. 3.17 Set of non-dominated solutions searched by each MOEA for test function ZDT6
    (a) For all comparing MOEAs; (b) Only closely contested MOEAs ............................... 52
Fig. 4.1 A simple three-cluster data set poising difficulties to many clustering methods ..... 55
Fig. 4.2 Flowchart of VRJGGA ..................................................................................60
Fig. 4.3 An example of chromosome representation ...........................................61
Fig. 4.4 Copy and paste operation within a single chromosome with 4 features .......65
Fig. 4.5 Copy and paste operation between different chromosomes with 3 features:
   (a) Before; (b) After .........................................................................................65
Fig. 4.6 Cut and paste operation between different chromosomes with 3 features:
   (a) Before; (b) After .........................................................................................66
Fig. 4.7 Performance on (a) Iris; (b) Wine data ..............................................69
Fig. 4.8 Performance on (a) Pima; (b) WBC data .........................................69
Fig. 4.9 Performance on (a) X2D2K; (b) X8D5K data ....................................70
Fig. 4.10 Performance on (a) Diamond9 data; (b) Clustering result obtained by VRJGGA on
   Diamond9 data ..............................................................................................70
Fig. 4.11 WDRP data: (a) Problem space with 8 labeled clusters; (b) performance obtained by
   VRJGGA ........................................................................................................72
Fig. 4.12 Diverse non dominated clustering solution obtained by VRJGGA: (a) solution with 8
   clusters; (b) solution with 10 clusters ..............................................................72
Fig. 4.13 Average performance trend on Iris data: (a) Overall deviation;
   (b) Generalized Dunn’s index ..........................................................................73
Fig. 4.14 Average performance trend on Wine data: (a) Overall deviation;
   (b) Generalized Dunn’s index ..........................................................................73
Fig. 4.15 Wine data: (a) Clustering with actual classes;
   (b) Clustering results with predicted grouping ................................................74
Fig. 4.16 WBC data: (a) Clustering with actual classes;
   (b) Clustering results with predicted grouping ................................................74
Fig. 5.1 Permutation with repetition approach for a 3x3 JSSP ..........................92
Fig. 5.2 Jumping operation within single chromosome ..................................94
Fig. 5.3 Selecting partial schedule ..................................................................95
Fig. 5.4 Exchange of partial schedules ..............................................................95
Fig. 5.5 Missed/ exceeded genes after exchange ..............................................95
Fig. 5.6 Repair work for O1: (a) Deletion; (b) Insertion ..................................96
Fig. 5.7 GOX crossover (the substring does not wrap around the end-points) ....97
Fig 5.8  GOX crossover (the substring wraps around the boundary) ........................................97
Fig 5.9  The job-pair exchange mutation ..............................................................................98
Fig 5.10 Complete evolutionary cycle for multiobjective JSSP ............................................98
Fig 5.11 Final Pareto-optimal front: (a) la21; (b) la24; (c) la38; (d) la40 ..............................106
Fig 5.12 Non-dominated solutions per generation (for a single run):
(a) la21; (b) la24; (c) la38; (d) mt6 ..................................................................................107
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCGA</td>
<td>Binary-Coded Genetic Algorithm</td>
</tr>
<tr>
<td>EA</td>
<td>Evolutionary Algorithm</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>JGGA</td>
<td>Jumping Gene Genetic Algorithm</td>
</tr>
<tr>
<td>JSSP</td>
<td>Job-Shop Scheduling Problem</td>
</tr>
<tr>
<td>MOEA</td>
<td>Multiobjective Evolutionary Algorithm</td>
</tr>
<tr>
<td>MOOP</td>
<td>Multiobjective Optimization problem</td>
</tr>
<tr>
<td>RCGA</td>
<td>Real-Coded Genetic Algorithm</td>
</tr>
<tr>
<td>RJGGA</td>
<td>Real Jumping Gene Genetic Algorithm</td>
</tr>
<tr>
<td>VRJGGA</td>
<td>Variable-Length Real Jumping Gene Genetic Algorithm</td>
</tr>
</tbody>
</table>
List of Publications

Journal Paper:


Contributed Book Chapter:


International Conference Papers:

