AGH

University of Science and Technology in Krakow

Faculty of Electrical Engineering, Automatics, Computer Science and Electronics

DEPARTMENT OF COMPUTER SCIENCE

MASTER OF SCIENCE THESIS

ADRIAN WOLNY

IMPROVING POPULATION-BASED ALGORITHMS USED IN GLOBAL OPTIMIZATION WITH FITNESS DETERIORATION TECHNIQUES

SUPERVISOR:
Robert Schaefer, Prof., PhD, DSc

Krakow 2011

I would like to thank ...

Contents

1. Introduction

> It is impossible for any optimization algorithm to outperform random walks on all possible problems.
>
> ... a conclusion from No Free Lunch Theorem

1.1. A statement of a problem

TODO: motivation of heuristics methods

A Global Optimization Algorithm is defined as optimization algorithm that employs measures that prevent convergence to local optima and increase the probability of finding a global optimum. However, there are classes of problems in which instead of finding the global optimum we are interested in finding many local optimums whose basins of attraction are properly wide and deep. One way to achieve this goal is to perform several runs of a evolutionary algorithm and alter the fitness function in every subsequent runs of the algorithm in a way that prevents exploration of basins which were found in previous runs of the algorithm. This work tries to find an effective fitness deterioration technique in high-dimensional domain spaces by interpolating fitness landscape in the area of basin of attraction for further fitness deterioration. Very often fitness function is computationally intensive and in such case it is unacceptable to perform classical interpolation of the fitness function. Making the assumption that clusters of population obtained after the single run of the algorithm are good estimators of basins of attraction, it would be better to exploit spatial characteristics of clusters and approximate basins of attraction by multidimensional Gaussain functions performing only a few fitness evaluation. Our goal is to create efficient method for fitness deterioration using the above schema and to analyze the relation between the deterioration accuracy and reproduction operators (mutation, crossover, etc.).

1.2. Definition of terms

1.3. Review of literature

TODO: sequential niching, prof. Obuchowicz works

2. Algorithm

2.1. A Hybrid Approach

TODO: algorithm description, pseudo-code description of the remaining chapters

2.2. Pseudocode

3. Clustering

TODO: cluster extension, set detection (specyfying set of individuals using small set of parameters, e.g center and radius), we make an assumption that clusters are good apprximation of basins of attraction, clustering metrics, clustering as an effective stop criterion (definition of the stop criterion, problems)

3.1. Cluster Extension

3.2. Clustering as a Stop Criterion

3.3. OPTICS

general description, optics ordering, extracting clusters, random samples, clustering, diagrams, reachability plots, why optics is good for further deterioration (describe in the next chapter)

4. Fitness Deterioration

TODO: description of the deterioration process, not very computationally intensive, easy to improve in subsequent runs, interpolation accuracy is not crucial, what is most important is to minimize the probability of finding the basins of attraction which was previously explored, use knowledge from clusteres;

4.1. Sequential niching

basic description of deterioration: when used, what approach, connection with clustering algorithms, advantages of OPTICS algorithm (improving deterioration by extracting clusters wiht different densities, cheapness)

4.2. Basic Scheme

detailed description of basic scheme deterioration

4.3. Adaptive Scheme

detailed description of adaptive scheme deterioration

5. Tested Algorithms

5.1. HGS

why HGS is well suited to our algorithm (suitable for clustering, fast convergence in leaves)

5.2. Tests

5.2.1. Benchmark functions

uni, bi and multimodal functions

5.2.2. Accuracy measures

how many optimas have been found, diagrams

5.2.3. Efficiency measures

6. Implementation

6.1. Architecture

6.2. Implementation in Java

clean structure, good test coverage, modular architercture, extensible,

6.2.1. Technologies

Spring, Maven, JUnit, Mockito, JAMA, TDD approach

6.2.2. Diagrams

class diagrams, sequence diagrams

7. Conclusions

7.1. Summary

7.2. Future Research

Bibliography

[1] *Global Optimization Algorithms - Theory and Application* by Thomas Weise

[2] *OPTICS: Ordering Points To Identify the Clustering Structure* by Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, Jrg Sander

[3] *Evolutionary search with soft selection* by A. Obuchowicz

[4] *Foundations of global genetic optimization* by Robert Schaefer, Henryk Telega

[5] *Introduction to Data Mining* by Tan, Steinbach, Kumar

[6] *Foundations of modern probability* by Kallenberg

[7] *No Free Lunch Theorems for Optimization* by David Wolpert, William Macready

[8] *Niching Methods for Genetic Algorithms* by Samir W. Mahfound