Quality-Time Tradeoff in A Distributed Parameter-less Genetic Algorithm

George G. Mitchell
Dept of Computer Science
National University of Ireland, Maynooth
Co.Kildare,
Ireland.

georgem@cs.nuim.ie

Abstract

This paper introduces our Distributed Parameter-Less Genetic Algorithm and highlights the practical and economical motivations of such a system. With the Distributed Parameter-Less GA users do not have to perform trial and error experiments to find suitable parameter settings for the Genetic Algorithm. The users also achieve economically viable results in a far shorter space of time as a direct result of the *GATermination* operator. The *GATermination* operator is a Quality–Time Tradeoff operator, it has been designed and tested with NP-Hard combinatoprial optimisation problems in mind such as the TSPLIB benchmark problem set. Experiments show that a significant saving in computation time can be achieved when generating multi-optimisation solutions.

1 INTRODUCTION

Self-Organising Genetic Algorithms have been a focus of research recently; Harik's *Parameter-less Genetic Algorithm* (1999) triggered a number of similar research projects [Goldberg 1999, Lobo 2001, and Tongchim 2002]. This research together with the *Meta-Evolutionary Approach* proposed by Mernik (2000) provided a catalyst for the work presented in this paper.
The parameter-less Genetic Algorithm bases it's motivation on developing genetic algorithms that are easier to use, and as a result hopefully improving the appeal of genetically inspired search techniques to the greater computer community. The appeal of genetically inspired searching has to a large extent existed in a subset of the computer search community, typically with the developers of the Genetic Algorithms. The parameter-less Genetic Algorithm removes the need for users to be experts in the field. Specific knowledge of selection rates, crossover probability and optimal population size are controlled by the GA itself delivering a more "*black box*" approach to the use of Genetic Algorithm's.
Meta-Evolutionary Genetic Algorithms have been used to find the best combination of crossover operators for a given problem [Mernik 2000]. This research has been based on the premise that the use of many different crossover operators out performs single crossover Genetic Algorithms. The Meta-Evolutionary Genetic Algorithm is yet another attempt to simplify the use of Genetic Algorithms.
With this research, Genetic Algorithm usability and potential increase in the adoption of Genetic Algorithm's as a *mainstream* search technique has been the driving force, one important area has been neglected.

Little has been said about the need for these techniques to grapple with the commercial reality of generating reasonably good solutions in commercially acceptable lengths of computation time. Genetic Algorithms by their nature enable the searcher to examine and explore potentially better solution in a search space. But it is common that a number of lucky runs are computed during any one single test, each of these lucky runs can take a significant amount of computation time and each of these runs facilitates the fine-tuning of the genetic operator parameters – selection, crossover, mutation, population size and use of adjunct genetic operators.

We have developed a Distributed Parameter-less Genetic Algorithm DPLGA, in essence combining the best of Harik's and Merniks Genetic Algorithms together our GeneRepair adjunct genetic operator [Mitchell 2003] with a Quality-Time Tradeoff operator.
The papers starts by introducing our Distributed Parameter-less Genetic Algorithm, followed by the benchmark problem set. We then introduce our Quality-Time Tradeoff operator and conclude the paper with

recent experiment results and outline some extensions to the work.

2 TRAVELING SALESMAN PROBLEM

Hamiltonian search optimisation has for some time remained one of the core benchmarks for any optimisation algorithm. In its more familiar name TSP or the *Travelling Salesman Problem* consists of a minimal distance Hamiltonian cycles of a complete graph visiting all N nodes. The TSP is a classical example of an NP-hard problem, in cases where N can be very large some form of algorithm which generates sub-optimal and hopefully near-optimal solutions is desired.
Many approaches have been taken to *'solving'* the TSP such as 2-opt, 3-opt, Ant colony, Tabu search, multiple heuristic search enhanced GA and many more [Lawler 1985, Martin 1996].
With the selection of your desired technique the choice of design remains. With a genetic algorithm applied to the TSP some implementation details differ from the more general genetic algorithm. In particular we are concerned with the validity constraints of the TSP, we should visit all N node once and once only. Therefore it is common for genetic algorithms to have a repair operator of some form, either directly repairing the genetic strings (the TSP city tours) such as GeneRepair Mitchell (2003) or genetic operator that guarantees not to develop invalid genetic strings [Crawford, 1996].

3 DISTRIBUTED PARAMETER-LESS GA

Great strides have been made in the development of differing representations and operators, which can be applied to problems that are not solved properly with the application of traditional bit string representation, and operators. Crawford (1996) provides an exhaustive report on the many differing crossover and mutation operators that can be used to solve a wide range of problems. As just seen, one set of combinatorial optimisation problem pose significant problem validity constraints; the TSP, VRP and similar problems and it is the TSP that we will examine in this paper. With this ever increasing *toolbox* of operators and representations, the selection of the *'right'* operators/representation together with the selecting of the *'right'* parameter settings has become an optimisation problem in its own right.

Solving this optimisation problem in itself requires the use of evolutionary computation techniques. Attempts have been made to address this problem most notably by Mernik (2000) but also to lesser extents by Grefenstette (1986) and Freisleben (1993). All of these approaches have been used to determine the crossover probability, mutation rate, generation gap, scaling window and selection operator. Reports recently published by Goldberg (1999)show that the population size is a critical factor in the development of building blocks, the fundamentals of Genetic Algorithms.
Distributed Genetic algorithms are a novel approach to solving extremely large problems and differing architectures have been suggested [Belding 1995]. Approaches differ principally in either dividing the problem amongst the distributed clients (an island approach) or farming individual problem to each of the autonomous distributed clients. In this research we utilised the latter approach.

The Distributed Parameter-Less Genetic Algorithm consists of a *server genetic algorithm* and a number of *client genetic algorithms* (Figure 1) the server genetic algorithm sometimes referred to as a Meta-GA, essentially marshals the clients, distributes jobs, maintains a fault tolerant component which guarantees that each job is completed and also consists of a Genetic Algorithm embodying the key functions of crossover, mutation, selection and a fitness function. The marshalling of clients is facilitated by a JAVA RMI implementation similar to that implemented by Hoban (2002) although Hoban's application was not a search optimisation algorithm.
Testing of the systems proved successful and the fault tolerance of the system [Mitchell 2000,2001a &2001b] repeatedly satisfied system specification when fault were deliberately inserted in to the system.

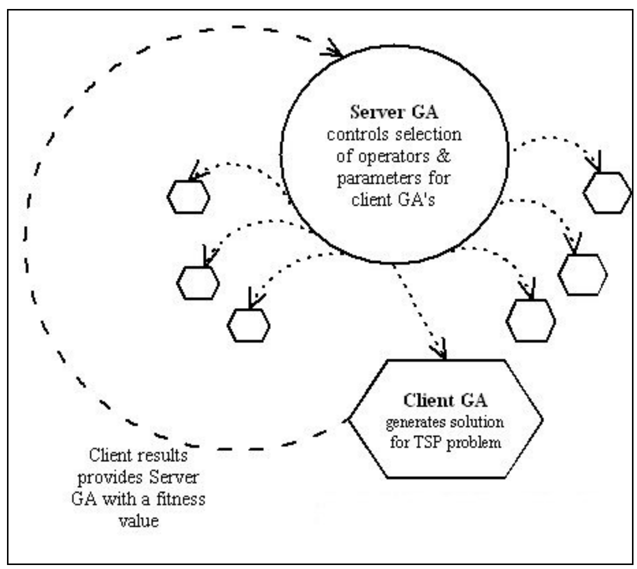

Figure 1 High-level view of the Dist. Parameter-less GA

Following testing of the Distribute Parameter-Less GA we found that a significant portion of operation time for the client GA's was spent on parameter settings which held little potential of improving. Whilst at first glance the decision to adopt a Multi-Objective Optimisation solution (pareto frointer) appeared to be most suitable, following examination it was decided to adopt an approach that would fulfil two distinct functions these were:

1. Fitness evaluation of the parameter setting - Server GA side

And

2. Termination of the search where a quality /time tradeoff was not being met – Client GA side

To accomplish this a new operator *GATermination* operator was developed (a quality-time tradeoff operator) for the Client GA. This consisted of a low pass filter of result within a sliding window appropriately sized for the give problem.

4 TRADEOFF OPERATOR

The use of Quality-Time tradeoff operators is not new, in fact Sosič (1994) developed a tradeoff function for a local optimisation algorithm, his Hill Climbing method was supplemented with his tradeoff function *Duty.* Duty minimized the *excess* (error) of a present solution compared to the benchmark optimal solution. An optimal present solution would be found to have an excess of zero. The generalised Duty measure gives equal weight to the quality of the solution and the computing time as follows:

$$D_k(t) = t^k * E(t)$$

Our *GATermination* operator operates as follows:

$$GAT(Q) = T + W(Q)$$

Weighting of the importance of the quality of the solution together with the time taken (*number of generations*) to find this solution generates a graph as illustrated in figure 4. By taking the minimum of the curve as the optimal solution of the GAT this provides the ability to halt the search of the Client GA if we believe further computation time would not be economically cost effective.

The selection of appropriate weighting is crucial to the correct operation of the Client GA and would yield an incorrect fitness value for the Server GA, thereby invalidating the results of the Distributed Parameter-less GA.

5 EXPERIMENT 1

We evaluated the effect of differing weightings for our Quality –Time operator, *GATermination*. We conducted several groups of tests on the TSP benchmark problems from the Heidelberg TSPLIB problem set (Reinelt, 1991). For these experiments we investigated the potential of the client GA result not only being used as a fitness value by the server GA, but also as a measure of the economic viability of solutions the client could generate from the given clients operator/representation parameters.

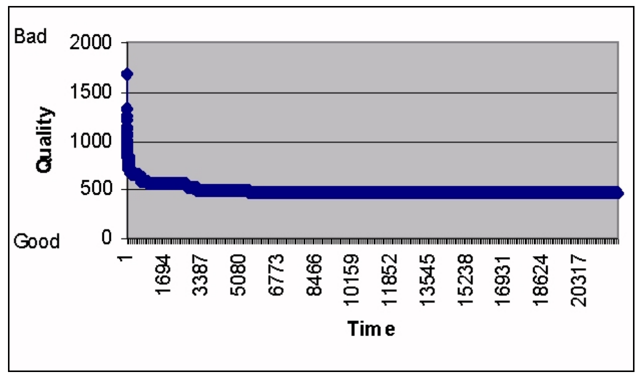

Figure 2: typical GA convergence on solution Q v's T

The optimal solution for the TSP eil51 problem is illustrated in figure 2; this test was performed with a mutation rate of 0.75, PMX crossover and tournament selection. From this it is clear that the majority of improvements occurred approximately in the initial 3,500 time units. Any termination prior to this would yield a solution that was significantly distant from the optimal solution calculated *apriori*.

Following this experiment we conducted experiments varying the weighting, figure 3 illustrates where a weighting of 10 has been used, a weighting of this would have yielded a solution within 10% of the optimal but outside of our permitted range. Figure 4 illustrates the optimal weighting found during experimentation, where it was determined that a weighting of twice the problem size was found to be the optimal weighting.

Figure 4 illustrates well the typical graph found during experimentation on the eil51, st70 and eil101 TSPLIB benchmark problem sets. Termination of the Client GA presented with the identical parameter settings as depicted in figure 4 would have occurred at time 5113. which very similar to the time take in figure 2 to find the optimal solution.

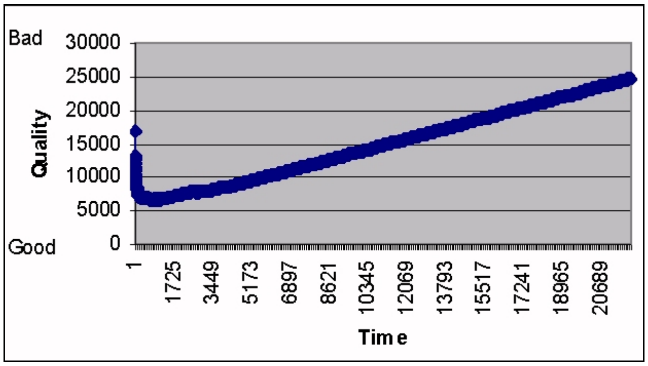

Figure 3: Premature termination of GA convergence on solution Q vs. T

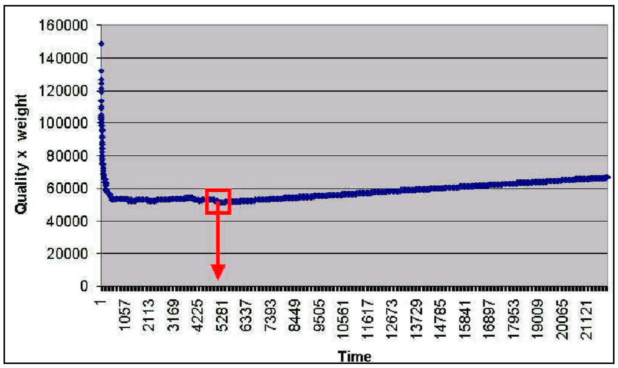

Figure 4: Termination of GA Client at time - 5113

6 EXTENSIONS

The experiment conducted so far are clearly work in progress and further examination of the applicability of the technique to larger problems such as fnl4461 a 4461 TSPLIB problem would be required before *GAT* can be accepted as a competent GA operator. The Distributed Parameter-less GA is currently in tests of problems in the range 100-1000 TSPLIB city problems, on completion of these experiments the DPLGA will be applied to the fnl4461 problem. A further extension to this project would involve the ability to adapt to any of the combinatorial problems previously mentioned such as the Vehicle Routing Problem, Printed Circuit Board problem and Data Packet Routing (Mitchell 2002). This should not prove to be difficult as the systems has been designed with the Client GA maintaining the problem specific information, all of these problems require very similar Genetic Algorithms and key differences exist on the problem validity constraints which we have located into a single operator GeneRepair [Mitchell 2000, 2003] Producing the system as a "seti at home" downloadable is on going, with this, the potential of increasing the computation power of the system is significant.

7 SUMMARY AND CONCLUSIONS

This paper reviewed the distributed parameter-less genetic algorithm and showed the practical and economical motivations of such a system. With the Distributed Parameter-Less GA users do not have to do trial and error experiments to find suitable parameter settings for the Genetic Algorithm, the users also will achieve economically viable results in a shorter space of time as a direct result of the GATermination operator. The GATermination operator is a quality –time tradeoff operator and has successfully been tested on the TSPLIB benchmark problem set where significant computation time can be saved.

REFERENCES

Harik, G. R., & Lobo, F. G. (1999). A parameter-less genetic algorithm. In Banzhaf, W., Daida, J.,Eiben, A. E., Garzon, M. H., Honavar, V., Jakiela,M., & Smith, R. E. (Eds.), GECCO-99: Proceed-ings of the Genetic and Evolutionary ComputationConference (pp. 258–267). San Francisco, CA:Morgan Kaufmann

D.E. Goldberg (1999): Using Time Efficiently: Genetic-Evolutionary Algorithms and the Continuation Problem. GECCO '99: 212-219

F.G. Lobo, and D.E. Goldberg, (2001), The parameter-less genetic algorithm in practice IlliGAL Report No 2001022, Illinois Genetic Algorithms Laboratory, USA.

Tongchim, S. and Chongstitvatana, P., "Parallel genetic algorithm with parameter adaptation", Information Processing Letters, Volume 82, Issue 1, pp. 47-54, 2002.

M. Mernik, M. Crepinsek, V. Zumer.(2000), A Meta Evolutionary Approach in Searching of the Best Combination of Crossover Cperators for the TSP. International Conference on Neural Networks, NN'2000, pp. 32-36.

G.G. Mitchell, D. O'Donoghue, D Barnes, M McCarville (2003), GeneRepair - A Repair Operator for Genetic Algorithms, Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '03, LBP, 88-93.

E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys.(1985) The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization. John Wiley Sons. Chichester, UK.

O.C. Martin and S.W. Otto,(1996), Combining Simulated Annealing with Local Search Heuristics. Annals of Operations Research 63, 57-75 .

T. C. Belding,(1995), The Distributed Genetic Algorithm Revisited. In Eshelman L. J. (ed.): Proceedings of the Sixth International Conference on Genetic Algorithms. Morgan Kaufmann, San Francisco, CA 114-121.

Crawford, K. D., R. Wainwright (1996), Research Question: How does one go about developing a new crossover operator with an a priori expectation of its merit? (A Survey of Crossover Operators for Genetic Algorithms), Technical Report UTULSA-MCS-96-2, The University of Tulsa, USA.

G.G Mitchell., W.M.Fitzgerald, J Doody. (2001) An Approach for Network Forwarding Systems Quality, IT&T 2001, pp 103 -111, ISSN 1649 - 1246, 2001.

G.G.Mitchell, , and S Brown,.(2001) A Model for Invisible Recovery Applied to Communications Networks, Applied Informatics 2001 , pp76 -81, ISBN 0-88986-320-2.

G.G.Mitchell, , and S Brown,.(2000)An Approach for Network Communications Systems Recovery, Applied Informatics 2000, pp 575-578, ISBN 0-88986-280-X.

J.J. Grefenstette, (1986) Optimization of Control Parameters for Genetic Algorithms. IEEE Transactions on Systems, Man & Cybernetics SMC-16, No. 1, pp. 122 –128.

B. Freisleben, M Hartfelder,(1993), In search of the best genetic algorithm for the traveling salesman problem. Proceedings of 9th International Conference on Control Systems and Computer Science, Bucharest, pp. 485 – 93.

R. Sosic, G.D. Wilby, (1994) Using the Quality-Time Tradeoff in Local Optimization. Proceedings of the IEEE Second ANZIIS Conference, Brisbane, Australia, pp. 253-257.

S. Hoban, J.Keating, B. Busschots (2002). Virtual Telescopes in Education, Journal of Digital Information 2(4):104

G. Reinelt. (1991), TSPLIB: A traveling salesman problem library. ORSA Journal on Computing, 3:376—384.

G.G. Mitchell, G.G., D. O'Donoghue, and A. Trenaman, (2000) A New Operator for Efficient Evolutionary Solutions to the Travelling Salesman Problem, Applied Informatics 2000, pp 771-774, ISBN 0-88986-280-X.

G.G. ,Mitchell Rafter A.M.., (2002) The Effect of a General Routing Algorithm Optimisation Function, IT&T 2002.