The Effects of Mutation and Directed Intervention Crossover when applied to Scheduling Chemotherapy

[Extended Abstract] *

Paul M. Godley, David E. Cairns, Julie Cowie, and Kevin M. Swingler
Dept. of Computer Science and Mathematics, University of Stirling
Stirling, Scotland
{pgo,dec,jco,kms}@cs.stir.ac.uk

John McCall
The School of Computing, Faculty of Design and Technology
The Robert Gordon University
Aberdeen, Scotland
jm@comp.rgu.ac.uk

ABSTRACT

This paper discusses the effects of mutation and directed intervention crossover approaches when applied to the derivation of cancer chemotherapy treatment schedules. Unlike traditional Uniform Crossover (UC), the directed intervention techniques actively choose the intervention level based on the fitness of the parents selected for crossover. This work describes how directed intervention crossover principles are more robust to mutation and lead to significant improvement over UC when applied to cancer chemotherapy treatment scheduling.

Categories and Subject Descriptors: I.2.8: Problem Solving, Control Methods, and Search: Scheduling

General Terms: Algorithms.

Keywords: Genetic Algorithms, Time Series, Chemotherapy, Crossover.

1. DIRECTED SEARCH

Genetic Algorithms (GAs) are frequently used to find effective solutions in a large solution space and may be applied to both static and time dependent problem domains [1]. Previous work into GA crossover approaches produced a directed intervention crossover method, Targeted Intervention with Stochastic Selection (TInSSel), suitable for application to time series problems [2]. This work introduces two alternative crossover approaches to TInSSel – Fitness Directed Search (FDS) and Directed Uniform Crossover (DUC). In contrast to UC; TInSSel, FDS and DUC operate on the principle of calculating the required number of interventions to use in offspring based on the relationships between parents fitness values and intervention levels.

To test these techniques, a cancer chemotherapy treatment problem was chosen as GAs have already been successfully used in chemotherapy design problems [4]. Varying mutation levels are reviewed to ascertain the crossover techniques robustness and ability to exploit diversity.

TInSSel is a directed intervention crossover approach which has previously been shown to outperform UC when applied to the scheduling of bio-control applications and is described in detail in [2]. Although TInSSel has proven that directed intervention can be effective, the question remains as to whether this is due to the sizing window approach, or indeed purely through providing a target intervention number for crossover to aim for. To investigate this, Directed Uniform Crossover (DUC) is presented as a simple technique which uses the fittest parent in the recombination pool and the number of intervention points utilized by this parent.

TInSSel uses the difference in intervention numbers between parents, but does not take into account the fitness difference. The Fitness Directed Search (FDS) approach utilises both these properties to calculate the number of interventions for offspring. The FDS approach places an emphasis on selecting intervention sizes that are close to the size of the fitter parent, while shifting in the direction that appears to offer the best improvement. The size of this shift is based on the intervention and fitness gradients between the two parents.

The FDS algorithm is presented with two parents for selection. F_1 is the normalised fitness associated with parent one and I_1 is the number of interventions utilised by parent one, with F_2 and I_2 the respective values for parent two. The normalised fitness score, F_{norm} is calculated by finding the maximum (F_{max}) and minimum (F_{min}) fitness scores contained in the current population and applying Equation 1, where F is the score being normalised.

The number of interventions used by the fitter parent is recorded as I_F and the normalised optimal fitness score is recorded as T, where *T=0* for a minimisation problem and *T=1* for a maximisation problem. The number of interventions to select in the offspring I_T, is calculated as shown in Equation 2.

I_T incorporates both the fitness and size difference between parents and therefore provides an intervention estimate that captures the relationship between the parents fitness and intervention values.

$$F_{norm} = \frac{F - F_{min}}{F_{max} - F_{min}} \quad (1)$$

$$I_T = I_F + (2T - 1)(I_1 - I_2)(F_1 - F_2) \quad (2)$$

*An expanded paper relating to this abstract is available at: http://www.cs.stir.ac.uk/∼dec/research/papers/fds

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-130-9/08/07.

2. RESULTS & CONCLUSION

Constructing an effective chemotherapy treatment schedule is a non-trivial task, with chemotherapy often considered one of the most complex cancer treatments [5]. This complexity makes this problem an ideal test to assess the abilities of FDS and DUC in searching a multi-constraint, extensive search space.The problem formulation is from [3]. To gauge the effectiveness of these directed crossover approaches, we compare their performance to the established UC approach.

Tournament selection was used to select parents with 2 potential parents in each tournament and a steady state population update. The population size was 100 and a crossover rate of 1.0 was used. 200 runs of each approach were conducted with 3 different mutation probabilities (0,0.001,0.01). Each run was recorded up to 10,000 Fitness Function Evaluations (FFE) to assess performance of the crossover approaches over time.

The results for both the 0.001 and 0.01 mutation experiments are shown in Figure 1. This shows the median performance of the relevant crossover approach at a given point in time (as measured by FFEs), where a high score indicates an effective treatment schedule. The graphs plot the median values from the runs with the error bars showing the first and third quartiles.

Where no mutation is present, all four crossover approaches converged prematurely and the results are not shown. The DUC approach produced the best results in this case since it just uses the best parent's intervention level as the guide for offspring. As UC is unguided, it has no means of adjusting its current intervention position and thus relies purely on mutation to find the best intervention level to use. The TInSSel and FDS approaches attempt to move the intervention level based upon the relative differences in the parent intervention and fitness scores. At the start of the evolutionary process, there is some variance in the population and progress is made, however as the variance disappears the population stagnates. With no variance to exploit, TInSSel and FDS also cease to improve.

Figure 1 shows the results for mutation levels of 0.001 and 0.01. It was observed that all approaches perform significantly better than the no mutation case. The main advantage of the directed approaches appears to be their relative efficiency at finding a good solution in a given time. UC is now able to use the variance introduced by this level of mutation to return better solutions when compared to its previous performance where no mutation was present. Even when FFEs equals 10,000, all the directed approaches were found to statistically outperform UC.

Where the mutation level is set to 0.01, all approaches achieve worse scores than the 0.001 mutation case. However, UC performs significantly worse than the directed approaches and would therefore be more sensitive to incorrect parameter tuning. As noise in a chromosome is likely to be relatively high, the DUC approach will be targeting potentially erroneous intervention levels. In contrast, the FDS approach is able to use the fitness and intervention gradient between the two parents to calculate a suitable level of interventions. From 2,000 FFEs onward, there was a measured statistically significant difference between the searching approaches of TInSSel and FDS versus DUC, which shows the need for variability in child sizing as discussed in Section 1. These results would seem to indicate that it is important to

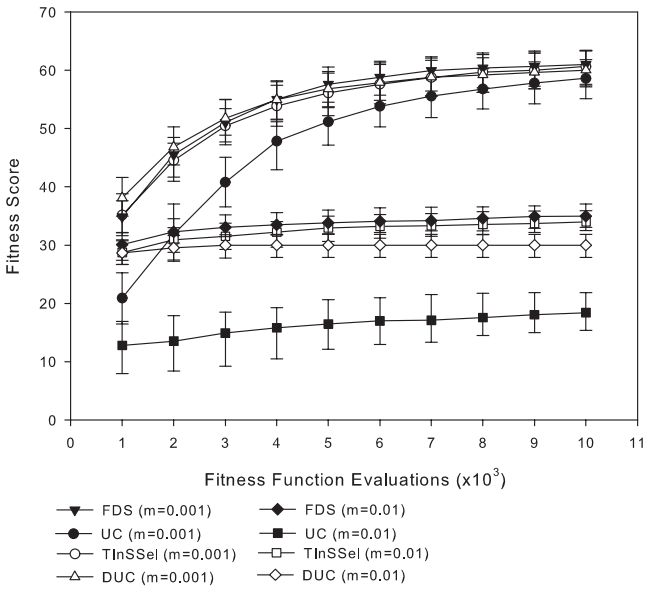

Figure 1: Fitness for mutation of 0.001 and 0.01

vary intervention levels around the currently observed best intervention level if the system is to remain robust to noise.

The experiments show a distinct advantage in using a directed intervention technique over traditional UC regardless of the fitness function evaluation point being observed. The directed approaches are quicker than UC at finding good scores and appear to be able to more effectively exploit the diversity introduced by mutation. In the case where one wishes to obtain optimal intervention schedules for a time-series domain problem, we therefore recommend using a technique such as FDS to improve search efficiency.

3. REFERENCES

[1] P. Collard, C. Escazut, and A. Gaspar. An evolutionary approach for time dependant optimization. In *Int. Conf. on Tools for Artificial Intelligence 96*, pages 2–9. IEEE Computer Society Press, 1996.

[2] P. M. Godley, D. E. Cairns, and J. Cowie. Directed intervention crossover applied to bio-control scheduling. In *IEEE CEC 2007: Proceedings of the IEEE Congress On Evolutionary Computation*. IEEE press, 2007.

[3] J. McCall, A. Petrovski, and S. Shakya. Evolutionary algorithms for cancer chemotherapy optimization. In G. B. Fogel, D. W. Corne, and Y. Pan, editors, *Computational Intelligence in Bioinformatics*, chapter 12, pages 265–296. Wiley IEEE Press, 2008.

[4] A. Petrovski, S. Shakya, and J. McCall. Optimising cancer chemotherapy using an estimation of distribution algorithm and genetic algorithms. In *GECCO '06: Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pages 413–418, New York, NY, USA, 2006. ACM.

[5] T. Wheldon. *Mathematical models in cancer research*. IOP Publishing Ltd., Bristol, 1988.