

Towards Understanding the Effects of Neutrality on the Sudoku Problem

Edgar Galvan-Lopez
Department of Computer S.
University of Essex, UK
egalva@essex.ac.uk

Julian Togelius
Department of Computer S.
University of Essex, UK
jtogel@essex.ac.uk

Simon Lucas
Department of Computer S.
University of Essex, UK
sml@essex.ac.uk

ABSTRACT

Over the last years, researchers have added neutrality in the evolutionary search in the hope that it can aid evolution. In this paper, we study the presence of neutrality that is already and to do so, we analysed the fitness landscape of the Sudoku problem. How and why neutrality affects evolutionary search is a reasonably well-studied but still not clearly understood topic. Here, we use neutral walks, neutrality trajectories and fitness distance correlation to attempt to throw new light on this topic.

Categories and Subject Descriptors: I.2.8 Artificial Intelligence [Problem Solving, Control Methods and Search]: Heuristic methods

General Terms: Algorithms

Keywords: Genetic Algorithms, Neutrality, Multimodal landscapes

1. NEUTRALITY AND SUDOKU

Recently, evolutionary computation researchers [1] have been increasingly interested in the effects of neutrality [3] in artificial evolution. Many fitness landscapes have large areas where fitness changes little or not at all between neighbouring genotypes. The questions under investigation relate to how such areas are distributed and how that in turn effects evolutionary search.

Sudoku is an immensely popular, originally Japanese, placement puzzle. We have previously studied how to use genetic algorithms to solve Sudoku puzzles[2]. The objective is to complete a Latin Square, to that each of the 9 rows, each of the 9 columns and each of the 9 sub-squares (boxes) contains exactly one of each integer between 1 and 9. The difficulty comes from each instance of the problem having a number of fixed positions, which have to be respected; e.g., if the initial board of a particular instance has the digit 3 in position 4.6, this position must have the same value in the solution.

Out previous work on evolving Sudoku solutions are based on a fitness function which simply counts whether each integer is present in each row, column and box, yielding a fitness between 0 and 243. Two different genotype spaces were compared, differing in whether rows are seen as permutations of the sequence 1-9 (the swap space) or rows are seen as sequences of digits between 1-9 with repetitions permitted. A number of mutation and crossover operators were

defined, all respecting the fixed positions; the swap space was found to be both smaller and smoother.

2. TOOLS FOR ANALYSING NEUTRALITY AND FITNESS LANDSCAPES

Neutral walks and mutation analysis are two ways of measuring how the neutral the neighbourhood around a particular genome is. A neutrality trajectory consists in following the best individual of an evolutionary run, and measuring its neutrality each time its fitness increases, using either neutral walks or mutation analysis. We found a very strong negative correlation between fitness and neutrality of a genome. Furthermore, there was no difference in neutrality at a given fitness between those evolutionary runs that eventually reached the global optimum and those that did not.

Fitness distance correlation (*fdc*) is a popular measure of problem hardness, it correlates the distance to the global optimum of a randomly generated genome with its fitness. Using this analysis, the Sudoku problem appears to be easier than it really is, something we argue is because it is deceptive only at higher fitness levels, not at the levels *fdc* samples. Further, we show that neutrality decreases *fdc*.

We also correlated the fitness and distance to global optimum of the best individuals at each fitness level of actual evolutionary runs. We showed that runs that eventually succeed differentiate from those that don't at a late stage.

3. CONCLUSIONS AND FUTURE WORK

We have seen that the amount of naturally occurring neutrality in this problem is strongly correlated to the fitness of an individual but not directly to the distance from the global optimum. This way, we can not tell whether neutrality increases evolvability. To test this, we will try to add artificial neutrality to the Sudoku problem in a forthcoming paper. Moreover, we are planning to use an extra problem with multimodal landscape features that can help us to complement our investigation.

4. REFERENCES

- [1] E. Galván-López, and R. Poli, An Empirical Investigation of How and Why Neutrality Affects Evolutionary Search. In *GECCO 2006*, pages 1149-1156, 2006, ACM Press.
- [2] A. Moraglio, J. Togelius and S. M. Lucas, Product Geometric Crossover for the Sudoku Puzzle. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 470-476, 2006, IEEE.
- [3] M. Kimura, Evolutionary rate at the molecular level. In *Nature*, volume 217, pages 624-626, 1968.