Towards an Optimal Restart Strategy
for Genetic Programming

Michael Solano
Oklahoma State University
700 N Greenwood, Tulsa, OK 74106
mike.solano@okstate.edu

Istvan Jonyer
Oklahoma State University
700 N Greenwood, Tulsa, OK 74106
jonyer@cs.okstate.edu

Categories and Subject Descriptors

I.2.8: Search, Heuristic Methods

General Terms

Performance, Experimentation.

The traditional approach to genetic programming attempts to find optimal individuals in a single run. This approach can suffer from *premature convergence* (Goldberg, 1989). Attempting to minimize this pitfall continues to be an area of active research.

Two methods that have shown promise are the implementation of a restart policy and the use of parallel multi-population island models. This work compares the performance of the island model and other restart policies. The comparison is conducted over a range of complexities and problem domains to determine if the difficulty of a problem or the domain has any bearing on which policy will be most successful.

A domain independent methodology for determining the complexity of a problem domain was proposed by Tomassini et al (2005) which defines the complexity of any problem in terms of a single value called the *fitness distance correlation*. Structural distance (Ekart and Nemeth, 2002) provides a convenient method to calculate the distance between any two program trees. Using the structural distance, it becomes possible to calculate the fitness distance correlation (*fdc*) for a particular problem domain. The importance of the *fdc* metric is that it provides a quantifiable relationship between the fitness of an individual and its distance from the ideal individual. This value can be used to gauge problem complexity (Tomassini et al, 2005).

The plot in figure 1 shows the general trends encountered using dynamic restart. Regardless of problem complexity, the best restart momentum was around four generations.
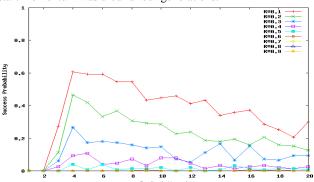

Figure 1: Dynamic Restart for Trap Functions with B=0.1

Figure 2 shows typical results for static restart, which seems to achieve good results around 50 generations, but better results with more generations. In figure 3, typical results for the island model experiments are shown. The best results are achieved using a small number of islands.

Analysis of the trap function domain shows that regardless of complexity, the dynamic restart method outperforms the other GP methods. The static restart method is the next best performer followed by the island model method. The worst performance, on average, was seen in the standard GP experiments.
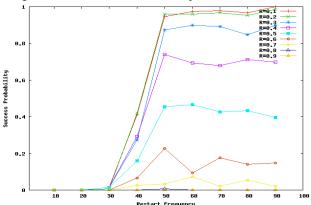

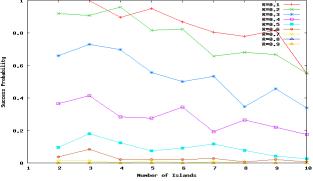**Figure 2:** Static Restart for Trap Functions with B=0.2

Figure 3: Trap Functions with B=0.2

[1] Ekart, A. and Nemeth, S. Z. 2002. Maintaining the diversity of genetic programs. In Genetic Programming, Proceedings of the 5th European Conference. Vol. 2278 of LNCS. pp. 162-171. Springer.

[2] Goldberg, David E. 1989. Genetic Algorithms in Search, Optimization, & Machine Learning. Boston, Massachusetts. Addison-Wesley.

[3] Tomassini, M., Vanneschi, L., Collard, P., Clergue, M. 2005. A Study of Fitness Distance Correlation as a Difficulty Measure in Genetic Programming. Evolutionary Computation. Vol. 13. Issue 2. pp. 13-239. Cambridge, Mass. The MIT Press.