A Survey of Mutation Techniques in Genetic Programming

Alan Piszcz
Department of Computer Science
University of Idaho
Moscow, ID 83844, USA
apiszcz@acm.org

Terence Soule
Computer Science Department
University of Idaho
Moscow, ID 83844, USA
tsoule@cs.uidaho.edu

ABSTRACT

The importance of mutation varies across evolutionary computation domains including: genetic programming, evolution strategies, and genetic algorithms. In the genetic programming community, researchers' view of mutation's effectiveness spans the range from an ineffective or marginal operator, to a neutral operator, to a highly effective operator that evolves solutions more effectively than genetic programming with crossover alone. Mutation implementation and associated parameters are often under reported in genetic programming research and typically lack context that justifies the technique and parameter selection. In part, reporting variance stems from the adaptation of mutation developed by the genetic algorithm community, and the creation of new mutation techniques in genetic programming. This survey describes the controversial operator in genetic programming applications, mutation selection operators, mutation techniques and offers an organization of mutation characteristics. We suggest methodologies to improve reporting of mutation parameters and related individual selection methods.

Categories and Subject Descriptors

I.2.2 [**Artificial Intelligence**]: Automatic Programming, program synthesis.

General Terms

Algorithms.

Keywords

Mutation, genetic programming.

1. INTRODUCTION

Genetic programming (GP) is a machine learning technique inspired by biological evolution to synthesize programs for a given computational task. The phases of the GP environment include: selection, crossover and mutation. Mutation alters an individual by modification of an evolving genetic program. In GP, mutation's role is to increase diversity that can improve the probability of overcoming local minima during the search process. The effectiveness of mutation can vary with factors like: the problem type, the researchers ability to attain optimal parameter values, the size of the population, the mutation technique and parameters.

Mutation receives considerable attention in other fields of evolutionary computation. In some communities, such as artificial life or evolution strategy, mutation is the primary method supporting variation and search. In this paper, we perform a survey of: 1) mutation's importance in research, 2) mutation selection, 3) common mutation techniques and 4) offer an organization of mutation as it relates to GP.

GECCO'06, July 8–12, 2006, Seattle, Washington, USA.
ACM 1-59593-186-4/06/0007.

2. BACKGROUND

Prior surveys provide overviews of GP with little discussion of mutation [2] [3] [1]. The sole focus of this survey is mutation; first we review related research that shows that mutation, when properly applied, can improve GP performance. Next, we review mutation implementation techniques for GP. Most mutation implementations have several common features: 1) Choosing the appropriate specification for controlling the mutation frequency of the population, individual or node 2) Deciding what portion of the population (all of the population, only the best N% of the individuals, only the worst N% of the individuals, etc.) will be mutated, which we refer to as *individual selection* 3) How mutation is implemented (subtree replacement, subtree swap, etc.) 4) Common mutation control parameters that influence the frequency of node or terminal selection, maximum depth of new subtree, and replacement subtree creation.

3. DISCUSSION

In many research papers incomplete reporting of mutation parameters limit the readers' ability to understand the role of mutation. Lack of reporting limits the understanding of mutation in many research papers. We propose a candidate tableau for mutation reporting in Table 1. Figure 1 is a candidate mutation reference organization for GP.

Proposed Tableau Entries: The tableau shown in Table 1 offers a proposed set of descriptive characteristics for mutation in GP. It is influenced by the results of this paper and aims to cover primary parameters and modes. The characteristics describing the environment parameters should be enumerated with the typical problem specification.

Organizing GP Mutation: Figure 1 summarizes the areas discussed in this paper and suggests some of the mutation choices in GP. Figure 1 shows a view of mutation, starting with populations, selection, implementation and control parameters.

	Selection				Mutation Adaptive Parameter Control Static Parameter Control		
					Implementation		
Population	Method	Control Parameters	Mutation Rate	Node Bias	Mutation Type	Mutation Sub Operation	Mutation Control Parameters
Multiple	All	Cutoff	Individual	Exon	Copy Subtree	Full	Cutoff
Single	Best	Tournament Size	Node	External	Delete Internal	Grow	Mutation Retry
	Fitness	Proportion	Population	Internal	Deflate	Half & Half	Ramp
	Fitness Overselect			Intron	Hoist		Rate
	Inverse Fitness				Inflate		Tree Depth
	Random				Insert Internal		
	Tournament				Point		
	Worst				Shrink		
					Swap Subtree		
					Uniform Subtree		

Figure 1: Reference GP Mutation Organization. A method for organizing GP mutation techniques and parameters that have appeared in the published literature.

Table 1: Proposed GP Mutation Tableau

Mutation characteristic	**Options**
Individual selection	All, best, fitness, fitness overselect, inverse fitness, random, tournament or worst.
Mutation rate	Individual, node or population.
Node bias	Exon, external, internal or intron.
Mutation Type	Copy subtree, delete internal, hoist, insert internal, point, shrink or swap subtree.
Mutation Sub-operation	Full, grow or half and half
Depth limit	Specifies replacement subtree maximum size
Ramp	Specifies the growth limits for the mutation sub-operations grow and half and half.
Mutation retry	The number attempts meet the maximum depth specification.

4. CONCLUSION

Understanding the advantages of mutation in GP continues to be an important and beneficial research area. Techniques for selecting individuals for mutation and methods of implementing mutation show a variety of capabilities and advantages that have yet to be associated with particular problem types. The facets of mutation in this paper include: mutation selection specifications, (PF, IF and NF) , individual selection techniques to select the target individual for mutation, mutation techniques, mutation control parameters, and we propose an organization and tableau for mutation reporting. As a community, improving our description of characteristics will benefit the understanding of mutation's influence (positive or negative) in GP research. This survey exposes the breadth of mutation research and proposes guidance to describe its reporting in the future. In summary we raised the following issues associated with mutation:

- *Mutation selection specification :* PF, IF and NF are three common methods; however, prior descriptions vary in quality and depth of coverage. We propose considering these for future descriptions of mutation implementation.
- *Individual selection techniques :* The selection techniques are an important aspect of mutation implementations, so the associated selection technique for mutation needs to be stated in research results. Future research could investigate the various selection techniques across a variety of problem classes and report their effectiveness.
- *Mutation techniques :* We show the variety of techniques and summarize their implementation.
- *Mutation control parameters :* Experiments need to report the mutation control parameter values to improve the ability to reproduce results.
- *Reporting :* We offer a tableau and reference organization for mutation research.

5. REFERENCES

[1] W. Banzhaf. Artificial intelligence: Genetic programming, July 04 2000. Contract no: 20851A2/2/102.

[2] J. R. Koza. Survey of genetic algorithms and genetic programming. In *Proceedings of 1995 WESCON Conference*, pages 589–594. IEEE, 1995.

[3] M. Willis, H. Hiden, P. Marenbach, B. McKay, and G. A. Montague. Genetic programming: An introduction and survey of applications. In A. Zalzala, editor, *Second International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications, GALESIA*, University of Strathclyde, Glasgow, UK, 1-4 Sept. 1997. Institution of Electrical Engineers.