Salzburg Interest Group on Integrated Systems

siGis-002-1997 January 6, 1998

ptGAs – Genetic Algorithms Using Promoter/Terminator Sequences

Evolution of Number, Size, and Location of Parameters and Parts of the Representation

Unpublished Dissertation

Workshop on Exploring Non–coding Segments and Genetics–based Encodings

International Conference on Genetic Algorithms
July 19–23, 1997, Michigan State University

Helmut A. Mayer

helmut@cosy.sbg.ac.at

Department of Computer Science
University of Salzburg

siGis
Salzburg Interest Group on Integrated Systems

Correspondence to:
Universität Salzburg
Institut für Computerwissenschaften und Systemanalyse
Jakob–Haringer–Straße 2
A–5020 Salzburg
Austria

ptGAs – Genetic Algorithms Using Promoter/Terminator Sequences Evolution of Number, Size, and Location of Parameters and Parts of the Representation

Helmut A. Mayer
Department of Computer Science
University of Salzburg
A–5020 Salzburg, Austria
helmut@cosy.sbg.ac.at

Abstract

At the workshop we would like to present an overview of our work on chromosome structures for genetic algorithms based on biological principles. The use of promoter/terminator sequences (the initials *pt* and genetic algorithms (*GAs*) are blended to ptGA) defining start and end of a gene not only enables evolution of parameter values, but also allows for adaptation of number, size, and location of parameters on an artificial chromosome. The genes (parameters) on these ptGA chromosomes are separated by non–coding regions induced by the promoter/terminator sequences. Randomly generated chromosomes of fixed length carry different number of genes of varying size and locations. Evolution of these ptGA chromosomes drives the number of parameters and their values to (sub)optimal solutions. Moreover, the formation of tightly linked building blocks is enhanced by self–organization of gene locations. There are a variety of real world problems, where the (sub)optimal number of parameters is not known in advance, and the ptGA proposed in this work could be applied to those kind of problems. The talk would be organized as follows.

First, the basic information processing aspects of molecular genetics are summarized with specific emphasis on gene structure. We discuss the gene structures in basic groups of biological organisms and the way they are processed in order to build higher–level biological building blocks. We review theories on the origin of non–coding regions in the genetic code, and conclude with the presentation of the basic principles transferred from biology to machines, specifically to the proposed ptGA.

Second, we review the development of the use of non–coding regions on GA chromosomes and GA instances allowing for flexible representation of parameters. We conduct experiments with ptGA chromosomes employing non–coding regions at fixed locations and compare the results achieved by ptGA to similar approaches by means of a simple example taken from

the literature [Lev91]. Moreover, we give the results of a formal analysis of the disruption reducing effects of non–coding regions with respect to crossover.

Third, we briefly report on experiments with randomly generated ptGA chromosomes with promoter/terminator sequences. The dependencies of the distribution of the number of parameters on different pseudo random number generators are investigated. Also, the distributions of the length (size) of parameters is studied.

Fourth, we utilize the presented ideas for the algorithmic solution of test problems. Besides interpreting the artificial chromosome in a single pass from begin to end, we also investigate a method, where the chromosome is interpreted in multiple passes, i.e., *Framing*, by reusing the genetic code (see also [Wu96]). We start out with the evolution of the number of parameters without considering their sizes and values. We introduce a new, non–disruptive crossover operator emerging from the ptGA gene structure in a self–organizing manner and compare this mechanism to conventional crossover in GAs.

Fifth, we consider the specific choice of promoter/terminator sequences influencing the space of parameter values. As this possibly excludes the optimum from the searchable space, we proceed with additional evolution of promoter/terminator sequences. This induces a variety of species and corresponding subpopulations (niches). An extension to tournament selection is proposed so as to keep interactions of species at a low level. This results in a nearly independent, parallel evolution of subpopulations within the ptGA. Various population and tournament sizes for the new selection method are investigated by means of experiments with a test problem given in [Lev91].

Sixth, we change the representation of parameter values by splitting genes into separately encoded units which is based on the biological model of interpreting the genetic code. In a less general setting this idea has been published for use in an evolutionary fuzzy reasoning system [YFU97]. We devise another simple test problem in order to study various sizes of the encoded units and the interaction with promoter/terminator sequences.

Seventh, we put it all together and employ the ptGA for the successful treatment of five instances of an NP–complete combinatorial optimization problem (0/1 multiple knapsack problem). The results are compared to conventional GA performance described in literature.

Eighth, we conclude the talk by considering possible problem domains for future ptGA applications and by reflecting on some additional ptGA features to be adopted to the current algorithm.

References

[Lev91] James R. Levenick. Inserting Introns Improves Genetic Algorithm Sucess Rate: Taking a Cue from Biology. In Richard K. Belew and Lashon B. Booker, editors, *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 123–127. University of California, San Diego, Morgan Kaufmann, 1991.

[Wu96] Annie Siahung Wu. *Non–Coding DNA and Floating Building Blocks for the Genetic Algorithm*. PhD thesis, University of Michigan, 1996.

[YFU97] Tomohiro Yoshikawa, Takeshi Furuhashi, and Yoshiki Uchikawa. The Effects of Combination of DNA Coding Method with Pseudo–Bacterial GA. In *Proceedings of the 4th IEEE International Conference on Evolutionary Computation*, pages 285–290. IEEE Press, 1997.