

# Markov Chain Model and Poli Based Schema Model for Bentley's Genetic Algorithm

Anna Paszynska

Institute of Computer Science, Jagiellonian University  
ul.Nawojki 11, Krakow, Poland  
apaszynska@hotmail.com

## Abstract

In this paper, I present the theoretical results obtained for Bentley's genetic algorithm, which is used in CAD system to generate 3D-solids designs. The Vose-like-Markov-chain model for the Bentley's GA is proposed. The transition matrices are found and the ergodicity of the Markov chain and the asymptotic correctness in the probabilistic sense are shown by using the model. The microscopic Exact Poli GP Schema Theory for Subtree-Swapping Crossovers are applied for the Bentley's GA to calculate the effective fitness and the total transmission probability for a fixed-size-and-shape schema under hierarchical crossover.

## Introduction

The GA based on hierarchical structures like tree based genetic programming [3, 4], graph-based GA and hierarchical chromosome based GA [1, 5] are often used as the representation in evolutionary algorithms. There are not known effective methods of investigation of asymptotic properties or convergence for that class of algorithms. The paper deals with the two approaches and introduces two models for the Bentley's GA (with hierarchical chromosome and hierarchical crossover) [1] – Vose-like Markov chain Model and Poli Schema Theory based one. Each approach models other aspects of genetic algorithms. The Markov chain model enables investigation on asymptotic properties and transition matrices. The second approach gives us possibility to calculate the total transmission probability for a fixed-size-and-shape schema under hierarchical crossover and the effective fitness.

## Markov chain model for Bentley's GA

In this section I will demonstrate a Markov chain Model for the Bentley's GA under the assumptions, that crossover points are numbers of primitives and that only mutation of alleles is used.

The phenotype space  $Fen$  is defined as a set of individuals consisting of  $n/3$  primitives of each of three classes: *base*, *seat*, *back*. Each primitive is described by  $p$  genes (sequences of  $q$  bits). The search space  $X$  is defined as  $X = (Z_2)^{npq}$ . After defining the coding and decoding functions, the mutation and crossover are defined as genetic operators on the space  $X$ . The crossover operator with two similarity points can be replaced with crossover with one similarity point by the assumption, that the primitives are arbitrary located within the chromosome. The probability distribution of the result of the crossing of the codes  $x, y$  is given by the following formula:

$$cross_{x,y}(\{z\}) = \sum_{t \in X} \frac{\eta_t + \eta_{\bar{t}}}{2} [(x \otimes t) \oplus (\bar{t} \otimes y) = z] \text{ where}$$

$$\eta_t = \begin{cases} p_c type_t & \text{for } t \neq \mathbf{0} \\ 1 - p_c - p_c type_0 & \text{for } t = \mathbf{0} \end{cases}$$

where  $p_c$  -crossover intensity,  $type$  -crossover type

$$type_t = \begin{cases} \frac{1}{n-1} \text{ gdy } (\exists l \in (0, npq) : t = 2^l - 1) \wedge \\ \wedge (\exists u \in (0, npq) : (1, t) = upq) \\ 0 \text{ in the other case} \end{cases}$$

The probability distribution of the result of the mutation of the code  $x$  is given by the following formula:

$$mut_x(\{z\}) = \sum_{s \in X} \xi_s [x \oplus s = z] \text{ where } \xi_s = p_m^{(1,s)} (1 - p_m)^{npq - (1,s)}$$

where  $p_m \in [0,1]$  -mutation rate

There has been shown, that the probability distribution  $m_{x,y}$  of mixing operation has two essential symmetries:

$$\forall x, y, z \in X \quad m_{x,y}(\{z\}) = m_{y,x}(\{z\});$$

$$\forall x, y, z \in X \quad m_{x,y}(\{z\}) = m_{x \oplus z, y \oplus z}(\{\mathbf{0}\})$$

After defining of the genetic operator  $G$  there has been shown that if the mutation and crossover parameters are constant, then the transition matrix for the genetic

$$\text{algorithm can be written as } Q_{p,p} = \mu^t \prod_{j=0}^{r-1} \frac{(G(p)_j)^{\mu p_j}}{\mu \tilde{p}_j}.$$

There has been also shown that (1) every population has a

nonzero probability of being the next generation and (2) the transition matrix  $Q_x$  is strongly positive definite.

From 2 it follows, that the Markov chain describing the dynamics of the genetic algorithm is ergodic. From the Ergodic Theorem it follows that

- for the arbitrary started probability measure  $\pi_\mu^0$  on  $X_\mu$

exists  $\pi_\mu : \lim_{t \rightarrow +\infty} \pi_\mu^t = \lim_{t \rightarrow +\infty} Q^t \pi_\mu^0 = \pi_\mu$

- the GA is asymptotically correct in the probabilistic sense.

## Exact Microscopic Schema Theory for Bentley's GA

Theoretical results presented below are obtained on the basis of Microscopic Schema Theory for GP with subtree-Swapping Crossover, introduced in [3]. Hierarchical crossover used in Bentley's GA is modeled as strongly typed crossover described by D.J.Montana in [2]. To describe hierarchical crossover by means of strongly typed crossover, I have to reorganize the hierarchical chromosome structure in the following way:

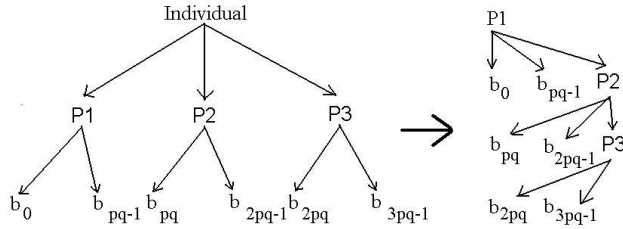


Figure 1. Reorganizing of the hierarchical chromosome

A node in the tree is coded in the Node Reference System as a pair  $(d,i)$ , where  $d$  is the layer number and  $i$  denotes node index in the layer. The following definitions defines a function other the node reference system:

- The function  $IB(d,i)$  returning 1 if the node with coordinates  $(d,i)$  is a bit and 0 in the other case
- $C(d,i,h)$  returning a class of the node with coordinates  $(d,i)$  in the tree with root  $h$
- Function  $CM$  returning 1 when nodes are in the same class and 0 in the other case.

For hierarchical crossover probability of choosing nodes with coordinates  $(d_1, i_1), (d_2, i_2)$  in  $h_1, h_2$  respectively equals:

$$P_{scgp}(d_1, i_1, d_2, i_2 | h_1, h_2) = \frac{CM(d_1, i_1, d_2, i_2, h_1, h_2)}{\sum_{D_1=0}^{\max\_depth(h_1)} \sum_{D_2=0}^{\max\_depth(h_2)} \sum_{I_1=0}^{pq} \sum_{I_2=0}^{pq} CM(D_1, I_1, D_2, I_2, h_1, h_2)}$$

A variable arity hyperschema is defined as in [3]. The only one difference is that the function set  $F$  and the terminal set  $T$  are defined as:  $F=\{base, seat, back\}$ ;  $T=\{0,1\}$ . The total transmission probability for a fixed-size-and-shape schema  $H$  under hierarchical crossover and no mutation is :

$$\alpha(H, t) = (1 - p_{x_0})p(H, t) + p_{x_0} \sum_{h_1} \sum_{h_2} p(h_1, t)p(h_2, t) *$$

$$\sum_{i \in H} \sum_j p(i, j | h_1, h_2) * [h_1 \in U(H, i)][h_2 \in L(H, i, j)]$$

where all the variables are defined as in [3].

The effective fitness for fixed size and shape schema  $F$  with hierarchical crossover and without mutation is:

$$f_{eff}(H, t) = \frac{f(H, t)}{p(H, t)} *$$

$$\left( (1 - p_{x_0})p(H, t) + p_{x_0} \sum_{h_1} \sum_{h_2} p(h_1, t)p(h_2, t) \sum_{i \in H} \sum_j p(i, j | h_1, h_2) \right) [h_1 \in U(H, i)][h_2 \in L(H, i, j)]$$

## Conclusions

In the paper, I have presented theoretical results obtained for Bentley's GA, which has been used in CAD system to generate 3D-solids designs. To model several aspects of the genetic algorithm, two approaches has been used. Firstly, the Vose-like-Markov-chain model for Bentley's GA has been introduced. This theoretical approach provides an excellent framework for studying the dynamics of genetic algorithms based on hierarchical structures. Presented model has been used for finding the transition matrices and for investigation of asymptotic properties by using the Markov transition functions. The ergodicity of the Markov chain describing the Bentley's GA has also been shown. Secondly, the microscopic Exact Poli GP Schema Theory for Subtree-Swapping Crossovers has been applied to Bentley's GA to calculate the effective fitness and the total transmission probability for the fixed-size-and-shape schema.

## References

1. Bentley, P.J. 1997. Genetic Evolutionary Design of Solid Objects using a Genetic Algorithm, Ph.D.diss, UCL London.
2. Montana, D.J. 1995. Strongly typed genetic programming. *Evolutionary Computation*. 3(2):199-230.
3. Poli, R. 2001. General Schema Theory for Genetic Programming with subtree-Swapping Crossover, Genetic Programming. In *Proceedings of EuroGP*, LNCS, (Milan):Springer-Verlag.
4. Poli, R.; Rove, J.E.; McPhee, N.F. 2001. Markov Chain Models for GP and variable-length Gas with Homologous Crossover , In *Proceedings of the Genetic and Evolutionary Computation Conference*. San Francisco, Calif.: Morgan Kaufmann.
5. Schaefer, R.; Kołodziej, J.; Paszynska, A. 2004. Hierarchical genetic computations in optimal design. *A Journal of Theoretical and Applied Mechanics*. 3(42):519-538.
6. Vose, M.D. 1999. *The simple genetic algorithm*, MIT Press.