International Journal of Advanced Research in
IT and Engineering ISSN: 2278-6244

ENCODING SCHEMES IN GENETIC ALGORITHM

Anit Kumar*

Abstract: *Genetic Algorithm (GA) are randomized searching and optimization techniques guided by the principles of evolution and natural genetic. They are efficient, adaptive and robust search processes. Genetic Algorithm handles a population of possible solutions represented by a chromosome and a chromosome is a sequence of genes. The main issue is how to represent the genes in a chromosome. Choosing the right scheme of encoding the genes is a crucial task. Encoding mainly depends on the type of problem. This paper studies different encoding schemes the problems in which they are used.*

Keywords: *Genetic Algorithm, encoding scheme, binary encoding, tree encoding, value encoding, Permutation encoding, octal encoding, hexadecimal encoding.*

*Research Scholar, Department of computer Science and Applications, Kurukshetra University, Kurukshetra

1. INTRODUCTION

Genetic algorithm [1] play a vital role with a long history as an emerging paradigm within Artificial Intelligence and Robotics. They are currently the most prominent and widely used models of evolution in artificial life systems. The two important issues in search are exploration and exploitation. Exploitation means to use the already available knowledge to find out the better solution and Exploration is to investigate new and unknown area in search space. In order to solve a real world problem, it is necessary to make balance between exploitation and exploration. Genetic Algorithm works on two types of spaces alternatively, coding space (genotype) and solution space (phenotype). The phenotype describes the outward appearance of an individual. A transformation exists between genotype and phenotype, also called mapping, which uses the genotypic information to construct the phenotype [4]. A chromosome refers to a string of certain length where all the genetic information of an individual is stored. Each chromosome consists of many alleles. Alleles are the smallest information units in a chromosome. In genetic algorithm, an encoding function is use to represent mapping of the object variables to a string code and mapping of string code to its object variable is achieve through decoding function as shown in figure 1.

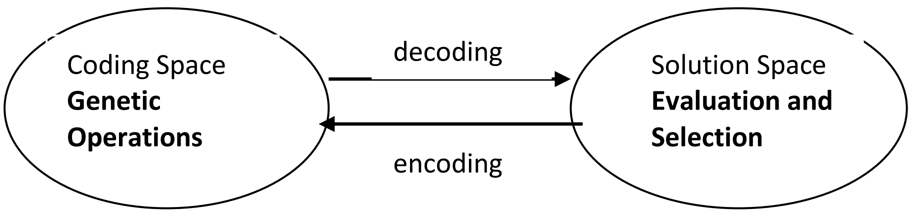


Figure .1: Encoding – Decoding method

2. PROPERTIES FOR SELECTION OF ENCODING SCHEME:

Genetic Algorithm is inspired from biological genetics model and most of its terminology has been borrowed from genetics. The genotype is the collection of genes possessed by an individual and genes are formed of DNA. In a living organism, the DNA combines with proteins to form the genes. The genes have a value called *alleles*. Each gene has a unique position on chromosome called *locus.* The encodings are chosen according to the following properties:

- Embodies the fundamental building blocks that are important for problem type [1].

- Is amenable to genetic operators that can propagate these building blocks from parents' genotype to offspring genotype [2].
- Allows a tractable mapping to the phenotype.

3. PRINCIPLES FOR SELECTION OF ENCODING

A schema is a template for describing a subset of chromosomes with similar sections. Schemas that are of low order, well defined and have above average fitness are termed as building blocks. Two basic principles [1] used for selection of encoding schemes are:

3.1 Principle of meaningful building blocks

According to Holland schema theorem, those building blocks which are above average in fitness, short defining length and low order will grow and others will degrade. Order of schema represents number of fixed position in schema and defining length cis the distance between the two fixed characters.

Example: For Schema 1*0*, defining length is 2 and order is 2.

3.2 Principle of minimal alphabets

The alphabets used for encoding should be as small as possible so as to increase the similarity pattern. It means that by reducing the cardinality of alphabets, one can increase the potential solutions. Example: (0,1) encoding is better than (A, B, C …….. Z) encoding.

4. TYPES OF ENCODING SCHEMES

Depending on the structure of encoding, they can be classified into 1-dimensional and 2-dimensional. Binary, Octal, Hexadecimal, permutation and value encodings are 1-dimensional and Tree encoding is 2-dimensional [3].

According to Goldberg, the fitness function for a specific encoding scheme depends on two factors - value and order. Encoding can be classified on fitness evaluation factor-

- Encoding schemes [5] where fitness depends on *order* only: f(o) e.g. Permutation encoding.
- Encoding schemes where fitness depends on *value and order*: f(v,o) e.g. Binary encoding.
- value only: f(v) e.g. Value encoding.

4.1 Binary Encoding: This is the most common form of encoding. In this encoding each chromosome is represented using a binary string. In binary encoding every chromosome is a string of bits, 0 or 1 [4]. Figure 2 shows the hexadecimal encoding.

Chromosome1	110101110010
Chromosome2	011010011101

Fig.2: Binary Encoding

In this encoding each bit show some characteristic of solution. On the other side each binary string represents a value. With smaller number of alleles, a number of chromosomes can be represented. Crossover operations possible in binary encoding are 1-point crossover, N-point crossover, Uniform crossover and Arithmetic crossover. The Mutation operator possible is Flip. In Flip mutation, bits changes from 0 to 1 and 1 to 0 based on generated mutation chromosome. This is generally use in Knapsack problem where binary encoding is used to show the presence of items say 1 to denote the presence of a item and 0 to the absence of a item.

4.2 Octal Encoding: In this encoding chromosome is represented using octal numbers (0-7). Figure 3 shows the Octal encoding.

Chromosome1	06254524
Chromosome2	63726425

Fig. 3: Octal Encoding

4.3 Hexadecimal Encoding: In this encoding chromosome is represented using Hexadecimal numbers (0-9, A-F).Figure 4 shows the hexadecimal encoding.

Chromosome1	97AE
Chromosome2	A2C6

Fig.4: Hexadecimal Encoding

4.4 Permutation Encoding: Permutation Encoding is used in ordering problems. In this, each chromosome represents position in a sequence e.g. in travelling salesman problem, the string of numbers represent the sequence of cities visited by salesman. Sometimes corrections have to be done after genetic operation is completed. Figure 5 shows the Permutation encoding.

Chromosome1	1 5 2 3 5 2 6 4 6 9 8
Chromosome2	8 6 3 6 3 9 6 3 1 5 8

Figure 5 : Permutation encoding

Permutation encoding is only useful for problems that have specific order. Some types of crossover and mutation corrections must be made to leave the chromosome consistent (i.e., have real sequence in it) for such kind of problems. Crossover operators performed on permutation encoding are Partially mapped crossover (PMX), Cycle crossover (OCX) and Order crossover (OX). Inversion [5] is the most commonly used mutation operator applied on ordered chromosomes. It changes the location of characters.

4.5 Value Encoding: In value encoding, each chromosome is represented as the string of some value. Value can be integer, real number, character or some object. In case of Integer values, the crossover operator applied are same as that applied on binary encoding. Values can be anything connected to problem, form numbers, real numbers or chars to some complicated objects. The figure 6 shows the Value encoding.

Chromosome1	1.23, 2.12, 3.14, 0.34, 4.62
Chromosome2	ABDJEIFJDHDDLDFLFEGT

Figure 6: Value Encoding

Value Encoding can be used in neural networks. This encoding is generally use in finding weights for neural network. Chromosome's value represents corresponding weights for inputs.

4.6 Tree Encoding: Tree encoding is mainly used for evolving programs or expressions for genetic programming. In tree encoding every chromosome is a tree of some objects, such as functions or commands in programming language. For example representation of an expression in binary tree forms as shown in the figure 7.

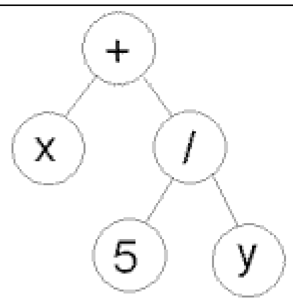
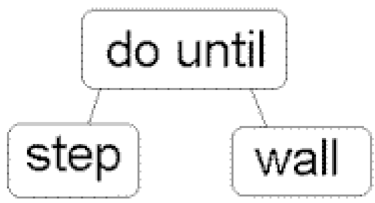
Chromosome 1	Chromosome 2
	
(+ x (/ 5 y))	(do until step wall)

Figure 7: Tree Encoding

Tree encoding is good for evolving programs. LISP is useful in this encoding as it helps in constructing tree for parsing and hence the crossover and mutation can be performed easily. Chromosomes are functions represented in a tree.

5. ANALYSIS OF DIFFERENT ENCODING SCHEMES

One can analyze that binary encoding is simplest and supports various crossover operators, but does not support inversion. As in the binary encoding, fitness depends on the value as well as order, so inversion operator results in disruption of building blocks and changes fitness value. Binary encoding supports the principle of minimal alphabets; as a result the number of exploitable schemas is maximized. On other hand, octal, hexadecimal, real encoding are used in limited number of applications. Permutation Encoding supports inversion, but does not support crossover that is supported by binary encoding. Permutation Encoding supports PMX, OCX, OX crossovers that are quite complex to implement. In Binary encoding scheme, fitness depends on value and order. In permutation encoding, it depends on order. In value encoding, it depends on value. There is no encoding scheme that is independent of these categories. So, there is a need to create a new encoding scheme that should be independent of value and order.

CONCLUSION AND FUTURE WORK:

After study of the existing encoding schemes, it has been established that encoding scheme used for representation of chromosome should follow two principles i.e. principle of meaningful building blocks and principle of minimal alphabets, that increases the survival rate of individuals. In some encoding schemes, the fitness depends on order of genes, while in others it depends on value. Different encoding schemes require different crossover and mutation operators. An operator applicable in one may or may not be applicable in others. So the applicability of various encoding schemes has lot to do with the problem specification and there is scope for further improvement in the encoding schemes through combination of features of the existing schemes and the operators they are applicable on.

REFERENCES

[1] D. Goldberg, Genetic algorithm in Search, Optimization and Machine Learning, Addison -Wesley, Reading, MA, 1989.

[2] B. Fox and M. McMahon, Genetic operators for sequencing problems, in Foundations of Genetic Algorithms, G. Rawlins, Ed. Morgan Kaufmann Publishers, San Mateo,CA, Ed. 1991, pp. 284-300.

[3] Mitsuo Gen, Genetic algorithms and engineering design, Wiley-IEEE, ISBN 0471127418, 1996.

[4] D. Fogel, Evolutionary Computation, IEEEPress,NewYork,Ed.1995,pp.631-637.

[5] Rakesh Kumar, Jyotishree, Novel Encoding Scheme in Genetic Algorithms for Better Fitness, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-6, August 2012.

[6] K. Deb, Handbook of Evolutionary computation, release 97/1. Oxford University Press, 1997.