International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 5, May 2012)

An Overview of methods maintaining Diversity in Genetic Algorithms

Deepti Gupta[1], Shabina Ghafir[2]
[1]Student (M.Tech 3 yr), Jamia Hamdard, New Delhi, India
[2]Assistant professor, Jamia Hamdard, New Delhi, India
[1]deepti.star.22@gmail.com
[2]impmailsforme@gmail.com

***Abstract* - Genetic algorithm is a search & optimization method based on the Darwin's principle of Survival of the fittest. It is an abstraction of complex natural genetics and natural selection process. Genetic algorithm is based on the principle of natural selection for reproduction and various evolutionary operations as crossover and mutation. Two controlling factors that need to be balanced in the process of selection are Genetic Diversity and Selective Pressure. Population Diversity can be controlled by a means of ways as Fitness sharing, Deterministic crowding and so many other. In this paper we are providing a brief knowledge about variety of methods maintaining population Diversity.**

***Keywords*-Genetic algorithms (GA), Diversity, Population convergence, Selection, Crossover, and Fitness function.**

I. INTRODUCTION

Genetic Algorithm is adaptive heuristic based on ideas of natural selection and genetics. Genetic algorithm is one of the most known categories of evolutionary algorithm. Genetic Algorithm is based on the mechanics of biological evolution initially developed by John Holland University of Michigan (1970's) and further carried by De Jong and Goldberg. Genetic Algorithm was designed to understand processes in natural systems it was developed to design artificial systems retaining the robustness and adaptation properties of natural systems. Although Randomized, GAs is by no means random, instead they exploit historical information to direct the search into the region of better performance within the search space [5].

A GA works with a number of solutions (collectively known as population) in each iteration which is chosen randomly. These solutions are usually coded in binary strings. Every solution or individual is assigned a fitness which is directly related to the objective function of the search and optimization problem. Thereafter, the population of individual is modified to a new population by applying three operators similar to natural genetic operators-reproduction, crossover, and mutation.

It works in a iterative manner by successively applying these three operators in each generation till a termination criterion is satisfied [3].GA is the method of solving problems by utilizing the processes of selection, crossover and mutation.

II. WORKING OF GA

Genetic Algorithm is started with a set of solutions (represented by chromosomes) called population. Solutions from one population are taken and used to form a new population. This is motivated by a hope, that the new population will be better than the old one [2]. Solutions which are selected to form new solutions (offspring) are selected according to their fitness - the more suitable they are the more chances they have to reproduce. This is repeated until some condition (for example number of populations or improvement of the best solution) is satisfied. A Simple GA working principle is shown in following Figure 1.The steps of simple Genetic Algorithm are described below here.

A. Basic Genetic Algorithm

Start: Generate random population of n chromosomes.

Fitness: Evaluate the fitness f(x) of each chromosome x in the population.

New population: Create a new population by repeating following steps until the new population is complete.

Selection: Select two parent chromosomes from a population according to their fitness as the fitness is better, the bigger chance to be selected by using various selection schemes.

Crossovers: With a crossover probability crossover the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents.

Mutation: With a mutation probability mutate new offspring at each locus (position in chromosome).

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 5, May 2012)

Accepting: Place new offspring in a new population.

Replace: Use new generated population for a further run of algorithm.

Test: If the end condition is satisfied, stop, and return the best solution in current population else move to Loop step.

Loop: Go to fitness step.

1. Formulate initial population
2. Randomly initialize population
3. Repeat
4. Evaluate objective function
5. Find fitness function
6. Apply genetic operators
7. Reproduction
8. Crossover
9. Mutation
10. until stopping criteria

Figure 1: The Working Principle of a Simple Genetic Algorithm

In contrast to local search methods, genetic algorithms are based on a set of independent operators such as selection, crossover and mutation controlled by a probabilistic strategy. Genetic algorithms are a sub category of evolutionary algorithm. Evolutionary algorithms are used to solve problems that do not already have a well defined efficient solution [4]. Genetic algorithms have been used to solve optimization problem.

III. FACTORS INFLUENCING GENETIC ALGORITHM

If population is not chosen intelligently it become difficult to find the correct solution of the problem whether in the case of initial population selection or the selection of population for the next generation. Some factors to take into account when the initial population is generated randomly are mentioned in the following figure 2.

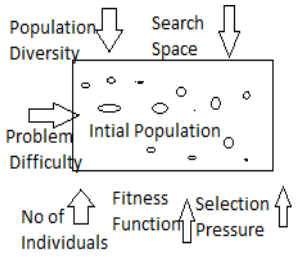

Figure 2: Factors affecting the initial population

Some factors that could influence the initial population or that should be taken into account when an initial population is generated randomly: the search space, the fitness function, the diversity, the problem difficulty, the selection pressure, and the number of individuals [6].We are discussing only two important factors here.

- *Diversity*
- *Selective pressure*

A. *Diversity*

The maintenance of a diverse solution population is required to ensure that the solution space is adequately searched, especially in the earlier stages of the optimization process. Population Diversity is considered as the primary reason for premature convergence. So a very homogeneous Population is found i.e. little Population Diversity is considered as the major reason for a Genetic Algorithm to premature converge [4].Premature convergence occurs when the population of a GA reaches such a suboptimal state that the genetic operators can no longer produce offspring that outperform their parents.

B. *Selective Pressure*

The tendency to select only the best members of the current generation to propagate to the next is required to direct the GA to an optimum.

Too much selective pressure can lower the genetic diversity so that the global optimum is overlooked & GA converges to a local optimum. Too little selective pressure prohibits GA to converge to an optimum in a reasonable time [28]. So a proper balance between Genetic Diversity and Selective Pressure is to be maintained for the GA to converge in a reasonable time to a global optimum.

IV. Methods For Maintaining Diversity

Population Diversity is qualitatively used for study the premature convergence. Degree of population diversity leads directly to premature convergence [7]. Techniques for diversifying a population typically reduce selection pressure, selection noise or operator disruption. Diversity-preserving mechanisms can help the optimization in two ways. A diverse population is able to deal with multimodal functions and can explore several hills in the fitness landscape simultaneously. Diversity-preserving methods can therefore support global exploration and help to locate several local and global optima. Methods for preserving diversity are mentioned as below.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 5, May 2012)

A. Nitching

A. De Jong introduced [8] the niching concept. A niche can be viewed as a subspace in the environment that can support different types of life. For each niche, the physical resources are finite and must be shared among the population of that niche. Nitching methods tend to achieve a natural emergence of niches and species in the search space. Nitching methods maintain population diversity and permit the GA to investigate many peaks in parallel. On the other hand, they prevent the GA from being trapped in local optima of the search space [25].

B. Crowding

De Jong subsequently presented the crowding concept to eliminate the most similar individual when a new one enters a subpopulation. Crowding is one of the methods that is based upon the restriction on selection methods. It is again of various types.

- As in standard crowding, only a fraction of the global population specified by a percentage G (generation gap) reproduces and dies each generation. In this crowding scheme, an offspring replaces the most similar individual (in terms of genotypic comparison) taken from a randomly drawn subpopulation of size CF (crowding factor) from the global population [8].
- Worst among most similar replacement policy [17] follows three steps. Firstly C_f crowding groups are created by randomly picking Cs crowding group size individuals (with replacement) per group from the population. Second one individual from each group that is most similar to offspring is identified. This gives C_f individuals that are candidate for replacement by virtue their similarity to offspring .The off spring will replace one of them. From this group of most similar candidates pick the one with the lowest fitness to die and be replaced by the offspring.
- Another type of crowding assumes that the parents would be one of the members of the population nearest to the new elements. In this way a family competition is held. These methods include deterministic crowding, keep best reproduction [19] and correlative family based selection [18].
- In deterministic crowding offspring compete directly with their respective parents. In every generation the population is partitioned into pairs of individuals. These pairs are then recombined and mutated. Every offspring then competes with one of its parents and may replace it if the offspring is not worse [7].

  Keep-best maintains the best parent and the best offspring in order to introduce good new genetic material into the population.
- Correlative family based selection chooses the best fitness individual as first survivor from each family then calculate distance of other family members with the highest fitness individual choosen and whosoever is minimum distant from the best individual is chosen as the survival for the next generation.

C. Restricted Mating

The restricted mating applies conditions such as restriction or encouragement, to select an individual and its mate partner. For example, the difference between pairs measured by the Hamming distance is used [9] for choosing the individuals for mating.

D. Sharing

Sharing method is the most frequently used technique for maintaining population diversity. It is inspired by natural ecosystem [10]. Each individual is forced to share its fitness value to its neighbours. The survival probability of an individual depends on its fitness value and its difference from others in the neighbourhood. Sharing must be used with the less biased selection methods. It tends to encourage search in unexplored area of the space. It comes with the limitation of more computation cost and priori knowledge of how far apart the optima are. Petrowski [26] suggested the clearing policy which encourages the winners to take all resources in a niche.

E. By Multiploidy

In nature many life form have poly-poid genotypes (multiploid) which consists of multiple sets of chromosomes with some mechanism for determining which gene is expressed [11]. Multiploid GA is able to recover from early genetic drift where good genes become lost in the initial selection process. It is useful in those cases where useful genetic material may otherwise be irretrievably lost.

F. Ranked space

Ranked space method is another strategy [12]. This embeds the diversity maintaining mechanism approach explicitly by the use of two ranks in the selection process called the quality rank and the diversity rank [29]. The combination of these two ranks is used to influence the selection probability. With this approach, the fitter individual is selected and at the same time the population diversity is maintained.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 5, May 2012)

G. DCGA

In the DCGA (Diversity control oriented GA) the structures for the next generation are selected from the merged population of parents and their offspring eliminating duplicates based on a selection probability, which is calculated using the hamming distance between the candidate structure and the structure with the best fitness value and is larger for structures with larger hamming distances. The idea is to exploit those worse solutions instead of discarding them by maintaining diversity of structures in the population [13].

H. Elitist

Elitist is one of the method [14] in which best two of these four (parent and offspring) go to the next generation. No separate selection and recombination phase but only a competition in each family which typically consist of two mating parents & their offspring. Best two of each family survive & are included in the next population. This method can very rapidly increase the performance of GA because it prevents loosing the best found solutions. Good solutions found are never lost unless even better solutions are created.

I. Injection

Injection strategy is another method [15]. Here fix point injection is used for certain number of generations. Inject new random number to the population to maintain the population diversity. The injection strategy should be carefully designed to avoid overlapping to the genes that have occupied the feasible slot. With it a sorting strategy is also introduced.

J. Removal of genotype or fitness duplicate

One simplest way to enforce diversity within the population is not to allow genotype duplicates. It prevents identical copies from entering the population as a natural way of ensuring diversity. One another restrictive mechanism is to avoid a fitness duplicate that is multiple individuals with the same fitness [7].

K. MOEA

A multi-objective evolutionary algorithm [16] keeping diversity of the population is the algorithmThis makes use of a metric based on entropy to measure the diversity of the population.

L. Replacement method

There are various replacement strategies try to maintain diversity. It can be categorized as Worst among most similar replacement, Family competition replacement scheme.

M. Using Tabu Multi parent Genetic Algorithm

The Tabu multi parent genetic algorithm (TMPGA) is presented. The mating of multiple parents in TMPGA is restricted by the strategy of tabu search [20]. The tabu list is used for preventing incest and maintaining the diversity of population.

N. CSGA

CSGA (complementary surrogate genetic algorithm) CSGA has the diversity-maintenance feature without an explicit mutation. The special feature of the CSGA is the inclusion of complementary surrogate set (CSS) into the population. The CSS is an individual or a set of individuals adding to the population for guaranteeing that each bit position of the whole population is diverse (not all '0' or all '1')[21] .

O. Fitness Uniform Selection Scheme

In the FUSS (fitness uniform selection scheme) lowest and highest fitness values in the population are (Let assume the representation) min f and max f respectively [22, 29]. The FUSS will select a fitness f uniformly in the interval [f min, f max]. Then the individual with fitness value nearest to f is selected. The FUSS maintains diversity better than a standard selection scheme since a distribution over the fitness value is used. Therefore, the higher and the lower fitness individuals are mixed in the selection.

Except all these above methods there are still more ways like Primal dual GA, Dual population GA, Multipopulation GAs, radius based methods [23],clustering [24], Adaptation of mutation rate, incest prevention [27], incest prevention in survival selection [13] , negative assortative mating[30]. All of these methods try to maintain diversity within the population so that global result may achieved in a proportionate balance time.

V. Conclusion

Genetic Algorithms are highly effective in searching a large, poorly defined search space even in the presence of difficulties such as high-dimensionality, multi-modality, discontinuity and noise.

Stochastic searching, intrinsically parallel nature with global perspective makes Genetic Algorithm of use. Maintaining diversity of individuals within a population is necessary for the long term success of any evolutionary system. So a balance between population diversity and selective pressure is to be maintained for finding an optimal solution in reasonable time. Genetic diversity helps a population adapt quickly to changes in the environment and it allows the population to continue searching for productive niches, thus avoiding becoming trapped at local optima. Thus improving diversity in GAs makes GA more useful efficient way to solve problems. So GAs is applied to a wide variety of searching and optimization problems in various fields from science, engineering and technology.

REFERENCES

[1] Kalyanmoy Deb. "Genetic Algorithm in search and optimization: The technique and application".

[2] Mir Asif iquebal" Genetic Algorithms and their application, "An overview ".

[3] Goldberg D.E. (1989) "Genetic Algorithm in search, optimization and machine learning".

[4] Darrell Whitely, "An overview of evolutionary algorithms: Practical issues and common pitfalls".

[5] R.Sivraj, Dr. Ravichandran, "A Review of selection methods in Genetic Algorithm ".

[6] Pedro A. Diaz-Gomez and Dean F. Hougen, "Initial Population for Genetic Algorithms: A Metric Approach".

[7] Tobias Friedrich, Pietro S. Oliveto, Dirk Sudholt, carsten witt, "Analysis of Diversity-Preserving Mechanisms for Global Exploration".

[8] K. A. DeJong, "An analysis of the behavior of a class of genetic adaptative systems," Ph.D. dissertation, Univ. of Michigan, Ann Arbor, 1975.

[9] L. J. Eshelman & J. D. Schaffer, (1991) "Preventing premature convergence in genetic algorithms by preventing incest", In Proceedings of the Fourth International Conference on Genetic Algorithms, pp. 115-122 .

[10] D. E. Goldberg & J. Richardson, (1987) "Genetic algorithms with sharing for multimodal function optimization", In Proceedings of the Second International Conference on Genetic algorithms and their application, pp. 41-49.

[11] Emma Collingwood, David Corne Peter Ross "Useful Diversity via Multiploidy".

[12] P. H. Winston, (1992) Artificial Intelligence, Addison- Wesley.

[13] Hisashi Shimodaira, "DCGA: A Diversity Control Oriented Genetic Algorithm".

[14] Goldberg, D.E. and K. Deb "A Comparative analysis of selection schemes used in GA ".

[15] Abu Bakar,Md. Sultan, Ramlan Mahmod,MD Nasir Sulaiman & Md. Rizam abu bakar"Maintaing diversity for Genetic algorithm: A case of timetabling problem".

[16] Xiaoning Shen, Min Zhang, Tao Li, "A Multi-objective Optimization Evolutionary Algorithm Addressing Diversity Maintenance"

[17] Manuel Lozano,Francisco Herrera,Jose ramon cano,"Replacement strategies to preserve useful diversity in SSGA".

[18] K. Matsui,"New selection methods to improve population diversity in GA".

[19] K.Wilese, S.D.Goodwin, Convergence characteristics of keep best Reproduction, Selected areas in cryptography".

[20] C. K. Ting & H. K. Buning, (2003) "A mating strategy for multi-Parent genetic algorithms by integrating tabu search", in Proceedings of the 2003 Congress on Evolutionary Computation, pp. 1259-1266.

[21] I.K. Evans, (1997)"Enhancing recombination with the Complementary surrogate genetic algorithm", In Proceedings of the IEEE International Conference on Evolutionary computation, pp. 97-102.

[22] M. Hutter, (2002) "Fitness uniform selection to preserve genetic Diversity", In Proceedings of the 2002 Congress on Evolutionry Computation, pp, 783-788.

[23] O. M. Shir and T. B¨ack, "Niching in Evolution Strategies". In 2005 Genetic and Evolutionary Computation Conference (GECCO '2005), volume 1, pages 915–916, New York, USA, June 2005. ACM Press.

[24] F. Streichert, G. Stein, H. Ulmer, and A. Zell. A Clustering Based Niching EA for Multimodal Search Spaces. In Artificial Evolution, 6th international Conference, Evolution Artificielle, EA 2003, Revised Selected Papers,pages 293–304, Marseille, France, Oct. 2004. Springer-Verlag. Lecture Notes in Computer Science Vol. 2936.

[25] Bruno Sareni and Laurent Kr¨ahenb¨uhl "Fitness Sharing and Niching Methods Revisited".

[26] A.Petrowski, "A New selection operator dedicated to speciation", In Proceeding of the seventh international conference on Genetic Algorithm,1997,pp. 144-151.

[27] L.J.Eshelman, J.D.Schaffer, "Preventing premature convergence in genetic algorithms by preventing incest", In Proceedings of fourth International conference on genetic algorithm 1991,pp. 115-122.

[28] SAS/IML User's guide 9.3 vol 1 chapter 20 page 501-505.

[29] Chaiwat Jassadapakorn and Prabhas Chongstitvatana, "Self adaptation mechanism to control the diversity of the population in Genetic Algorithm".

[30] C.Fernandes , A Rosa , "A Study on non-random mating and varying population size in genetic algorithm using a royal road function" ,in Proceedings of IEEE congress on evolutionary computation, Seoul, South Korea,2001.