ISSN:2320-0790

COMPUSOFT
An International Journal of Advanced Computer Technology

Mathematical Study OF Adaptive Genetic Algorithm (AGA) with Mutation and Crossover probabilities

Dipanjan kumar Dey
Assistant professor, Department of Computer Science & Engineering, PITM, Kolkata

Abstract: GAs is powerful search techniques that are used successfully to solve problems in many different disciplines.AGA are particularly easy to implement and promise substantial gains in performance. AGA has some parameters, such as population size, the crossover probability (P_c), the mutation probability (P_m) are varied while genetic algorithm is running. Genetic algorithm includes these parameters that should be adjusting so that the algorithm can provide positive results. The main aim of this paper is that how to design of adaptive crossover probability (P_c) and mutation probability (P_m).By varying P_c and P_m adaptively it response to the fitness values of the solution.

Depending the fitness value of the solution, in AGA the crossover probability (P_c) and the mutation probability(P_m) are varied. High- fitness solutions are 'protected', while solutions with sub average fatnesses are totally disrupted, that is by varying the crossover probability (P_c) and the mutation probability(P_m) adaptively in response to the fitness value of the solution: when the population tends to get struck at a local optimum, the crossover probability (P_c) and the mutation probability(P_m) are increased and when the population is scattered in the solution space, the crossover probability (P_c) and the mutation probability(P_m) are decreased.

Keywords: crossover probability (P_c), mutation probability(P_m)

INTRODUCTION

Crossover and mutation operators of genetic algorithms are used for constructing the adaptive abilities. The choice of (P_c) and (P_m) is known to critically affect the behavior and performance of GAs. The crossover rate controls the capability of GAs in exploiting a located hill to reach the local optima. The higher the crossover rate, the quicker exploitation proceeds. The crossover probability (P_c) controls the rate at which solutions are subjected to crossover. The higher the value of (P_c) the quicker are the new solutions introduced into the population. A (P_c) that is too large would disrupt individuals faster than they could be exploited. The mutation rate controls the speed of GAs in exploring a new area. Small (P_m) values are commonly adopted in GAs. Typical values of (P_c)are in the range 0.5~1.0, while typical values of (P_m) are in the range 0.001~0.05 are commonly employed in GA practice.

Adaptive probabilities of crossover and mutation operators

Two characteristics are held are essential in GAs for optimizing multimodal functions. The first characteristic is the capacity to converge to an optimum, local or global after locating the region containing the optimum. The second characteristic is that the capacity to explore the new regions of the solution space in search of the global optimum. The balance between these two characteristics of the GA is directed by the values of P_m and P_c and the type of the operators employed.

It is commonly understood that crossover plays a important role in conversing, by combining the solutions closed to an optimum. If we choose solutions with higher fitness values for crossover, we may expect GA to converge faster to the nearby optimum. On the other hand when population becomes too homogeneous it is not clever to favour solutions to that way, because there is a danger of getting stuck to a local optimum.

Mutation is the operator which is mainly responsible for preventing GA of becoming stuck. If a population converges to a local optimum, it is possible to drive it way with increased mutation probabilities. But to vary the choice of solutions to be crossed and the mutation rate, it is essential to be able to identify whether the GA is conversing to an optimum.

Design of adaptive crossover probability (P_c) and the mutation probability (P_m):

Let $\bar{f}$ be the average fitness value of the population, f_{max} be the maximum fitness value of the population, The value $f_{max} - \bar{f}$ is yardstick for detecting the convergence of genetic algorithm.

The value $f_{max}-\bar{f}$ is likely to be less for a population that has converged to an optimum solution than for a population scattered in the solution space.

From figure we notice that $f_{max}-\bar{f}$ decreases when the GA converges to a local optimum with a fitness value of 0.5.Note that the globally optimal solution has a fitness value of 1.0.

The values of P_c and P_m are varied depending on $f_{max}-\bar{f}$

Since P_c and P_m have to be increased when the GA converges to a local optimum, i.e. When $f_{max}-\bar{f}$ decreases, P_c and P_m varied inversely with $f_{max}-\bar{f}$. Thus P_c is inversely proportional to $f_{max}-\bar{f}$ and (P_m) is inversely proportional to $f_{max}-\bar{f}$, i.e. When P_c increased $f_{max}-\bar{f}$ decreased and vice versa. Similarly P_m increased $f_{max}-\bar{f}$ decreased and vice versa.

Thus we can write,

$P_c \propto \frac{1}{\text{fmax} - \bar{f}}$ and $P_m \propto \frac{1}{\text{fmax} - \bar{f}}$

Or, $P_c = \frac{K_1}{\text{fmax} - \bar{f}}$ ………(1)and $P_m = \frac{K_2}{\text{fmax} - \bar{f}}$……..(2),

Where K_1 and K_2 are constants of proportionality

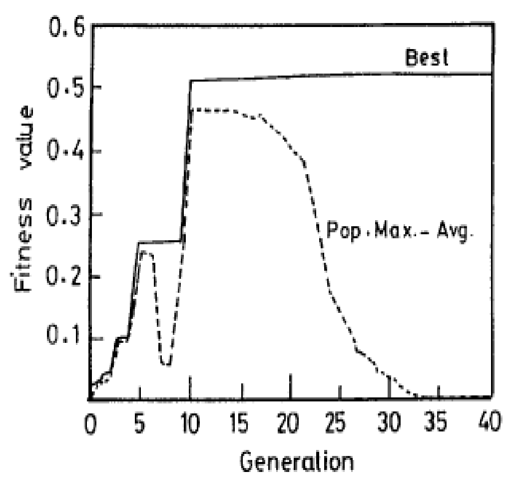

Fig-

Variation of $f_{max} - \bar{f}$ and f_{best} (best fitness).

Thus we see from equation (1) & (2) and from figure

1) P_c & P_m do not depend on the fitness value of any particular solution of the population.

2) P_c & P_m depends for all the solution of the population.

3) At the same levels of mutation and crossover are subjected to solution with high fitness values as well as solution with low fitness values.

4) P_c & P_m increase and may cause the disruption of the near optimal solutions. When a particular solution of the population converges to a globally optimal solution (or even locally optimal solution),the population may never converge to the global optimum.

Though we may prevent the GA from getting stuck at a local optimum, the performance of GA (in terms of the generations required for convergence) will certainly deteriorate.

To overcome the above-stated problem, we need to preserve 'good' solutions of the population. This can be achieved by having lower values of P_c & P_m for high fitness solutions and higher values of P_c & P_m for low fitness solutions. While the high fitness solutions aid in the convergence of the **GA,** the low fitness solutions prevent the **GA** from getting stuck at a local optimum. The value of P_m should depend not only on **fmax** -$\bar{f}$, but also on the fitness value f of the solution. Similarly, P_c should depend on the fitness values of both the parent solutions. The closer ***f*** is to **fmax,** the smaller P_m should be, i.e., P_m should vary directly as **fmax** – f . Similarly, P_c should vary directly as fmax - *f',* where f' is the larger of the fitness values of the solutions to be crossed.

The expressions for P_c now take the form

$$P_c \propto (\text{fmax} - f')$$

But we know, $P_c \propto \frac{1}{\text{fmax} - \bar{f}}$

Combine these two we get,

$$P_c \propto \left[\frac{\text{fmax} - f'}{\textbf{fmax} - \bar{f}}\right],$$

Or, $P_c = k_1\left[\frac{\text{fmax} - f'}{\textbf{fmax} - \bar{f}}\right],\ k_1 \leq 1.0$ ------------ (3)

Where K_1 is the constant of proportionality

And the expressions for P_m now take the form

$$P_m \propto (\mathbf{fmax} - f)$$

But we know, $P_m \propto \frac{1}{fmax - \overline{f}}$

Combine these two we get,

$$P_m \propto [\frac{\mathbf{fmax} - f}{\mathbf{fmax} - \overline{f}}],$$

Or, $P_m = k_2[\frac{\mathbf{fmax} - f}{\mathbf{fmax} - \overline{f}}]$, $k_2 \leq 1.0$ ------------- (4)

Where K_2 is the constant of proportionality

It is to be noted that P_c and P_m are zero for the solution with maximum fitness.

From equation (3) and (4) we have,

From equation (3), for a solution with $f' = \overline{f}$ then $P_c = k_1$

From equation (4), for a solution with $f = \overline{f}$ then $P_m = k_2$

If $f <$ f',that is solution with sub average fitness values, P_c and P_m might assume values larger than 1.0.Now to prevent the overshooting of P_c and P_m beyond 1.0,we have the following constraints

$$P_c = k_3, f' \leq \overline{f}$$

And $P_m = k_4$, $f \leq \overline{f}$, where k_3, $k_4 \leq 1.0$

To overcome the problem that P_c and P_m are zero for the solution with maximum fitness we introduce a default mutation rate (of0.005) for every solution in AGA.

For more convenience the expressions for P_c and P_m are given as

$$P_c = k_1[\frac{fmax - f'}{\mathbf{fmax} - \overline{f}}], f \geq \overline{f}$$

$$P_c = k_3, f' < \overline{f}$$

$$P_m = k_2[\frac{\mathbf{fmax} - f}{\mathbf{fmax} - \overline{f}}], f \geq \overline{f}$$

$$P_m = k_4, f < \overline{f}$$

Where $k_1, k_2, k_3, k_4 \leq 1.0$

The moderately large values of P_c promote the extensive recombination of schemata, while small values of P_m are necessary to prevent the disruption of the solutions. These guidelines, however, are useful and relevant when the values of P_c and P_m do not vary.

One of the goals of our approach is to prevent the **GA** from getting stuck at a local optimum. To achieve this goal, we employ solutions with sub average finesses to search the search space for the region containing the global optimum. Such solutions need to be completely disrupted, and for this purpose we use a value of 0.5 for $\mathbf{k_4}$. Since solutions with a fitness value of $\overline{f}$ should also be disrupted completely, we assign a value of 0.5 to $\mathbf{k_2}$ as well.

Based on similar reasoning, we assign k_1 and k_3 a value of 1 .O. This ensures that all solutions with a fitness value less than or equal to $\overline{f}$ compulsorily undergo crossover. The probability of crossover decreases as the fitness value (maximum of the fitness values of the parent solutions) tends to fmax and is 0.0 for solutions with a fitness value equal to fmax.

CONCLUSION

The probabilities of crossover (p_c) and mutation (p_m) greatly determine the degree of solution accuracy and the convergence speed that genetic algorithms can obtain. Instead of using fixed values of p_c and p_m, AGAs utilize the population information in each generation and adaptively adjust the p_c and p_m in order to maintain the population diversity as well as to sustain the convergence capacity. In AGA (adaptive genetic algorithm), the adjustment of p_c and p_m depends on the fitness values of the solutions. It has been well established in GA literature that moderately large values of P_c ($0.5 < P_c < 1.0$) and small value of P_m ($0.001 < P_m < 0.05$) are essential for the successful working of Gas.

REFERENCES

1. H. Aytug and G. J. Koehler, "Stopping criteria for finite length genetic algorithms, "INFORMS Journal on Computing, Vol. 8, 1996, pp. 183-191.
2. T. Bäck, "Self-adaptation in genetic algorithms," in Proceedings of the First European Conference on Artificial Life, 1992, pp. 263-271.
3. T. Bäck, "Optimal mutation rates in genetic search," in Proceedings of the Fifth International Conference on Genetic Algorithms, 1993, pp. 2-8.
4. K. Deb and S. Argrawal, "Understanding interactions among genetic algorithm parameters,"in Foundations of Genetic Algorithms 5, 1998, pp. 265-286.
5. K. A. D. Jong, "An analysis of the behavior of a class of genetic adaptive systems," PhD thesis, University of Michigan, 1975.

6. K. A. D. Jong, "Adaptive system design: A genetic approach,"IEEE Transactions on System, Man and Cybernetics, Vol. 10, 1980, pp. 566-574.
7. A. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms,"IEEE Transactions on Evolutionary Computation, Vol. 3, 1999, pp.124-141.
8. T. C. Fogarty, "Varying the probability of mutation in genetic algorithms," in Proceedings of the Third International Conference on Genetic Algorithms, 1989, pp.104-109.
9. D. E. Goldberg,Genetic Algorithms in Search, Optimization & Machine Learning,Addison Wesley, 1989.

Dipanjan Kumar Dey, graduated from Calcutta University. M.sc (Mathematics) and M.Tech (Computer Science &Engineering) from M.C.K.V Institute of Engineering (under West Bengal University & Technology, India). He is currently Assistant Professor of Mathematics & Computer Science in Prajnanananda Institute of Technology & Management, West Bengal India. He is also Faculty member of Institute of Chartered financial Analysis of India (ICFAI) and Academic Counselor, Assistant Coordinator of Indira Gandhi National Open University (IGNOU) study center 2804, Kolkata. He is a Science Journalist having Post Graduate certificate course on journalism and media practice from National Council for the Science and Technology Communication, GOVT. OF INDIA, New Delhi.
His research interests in Genetic Algorithms, Soft Computing, Fuzzy Set, Artificial intelligence.