A STUDY OF PERFORMANCE OF STOCHASTIC UNIVERSAL SAMPLING VERSUS PROPORTIONAL SELECTION ON GENETIC ALGORITHMS

Minetti G., Salto, C., Alfonso H.
Proyecto UNLPAM-09/F009[1]
Departamento de Informática - Facultad de Ingeniería
Universidad Nacional de La Pampa
Calle 110 esq. 9 – Local 14
(6360) General Pico – La Pampa – Rep. Argentina
e-mail: { minettig,saltoc,alfonsoh}@ing.unlpam.edu.ar;
Phone: (02302)422780/422372, Ext. 6412

Gallard R.
Proyecto UNSL-338403[2]
Departamento de Informática
Universidad Nacional de San Luis
Ejército de los Andes 950 - Local 106
5700 - San Luis
Argentina
E-mail: rgallard@unsl.edu.ar
Phone: +54 2652 420823
Fax : +54 2652 430224

Abstract

Selection mechanisms favour reproduction of better individuals imposing a direction on the search process. According to this it is expected that the effective number of offspring of an individual in the next generation would always agree with the algorithmic sampling frequencies. This does not happens due to sampling errors. Stochastic universal sampling is a method that tries to remedy this problem.
This presentation discusses performance results on evolutionary algorithms optimizing a set of highly multimodal functions and a hard unimodal function, under Proportional selection and stochastic universal sampling. Contrasting results are shown.

1.Introduction

[1] The Research Group is supported by the Universidad Nacional de La Pampa.
[2] The Research Group is supported by the Universidad Nacional de San Luis and the ANPCYT (National Agency to Promote Science and Technology).

Selection mechanisms favour reproduction of better individuals imposing a direction on the search process. Rather than creating new individuals; it selects comparatively good individuals from a population for mating. The idea is that competing with other individuals, those with higher fitness have a higher probability to be selected for mating. In that manner, because the fitness of an individual gives a measure of its "goodness", selection introduces the influence of the fitness function to the evolutionary process [5], [6], [7], [8], [10]. Moreover, selection is the only operator of genetic algorithm where the fitness of an individual affects the evolution process. In such a process two important, strongly related, issues exist: selective pressure and population diversity [4].

In proportional selection (PS), individuals are chosen for mating from a population according to its fitness. This is the simplest selection scheme also known as *roulette-wheel selection* or *stochastic sampling with replacement.*
Here, individuals are mapped to contiguous segments in the real interval [0,1] in such a way that a segment corresponding to an individual has a size equal to the individual fitness. Then a random number in such interval is generated and the individual whose segment encompasses the random number is selected.
The *selection probability* is an important parameter of a selection mechanism and normally determines the number of expected copies of an individual after selection .

These expected values not always agree with the algorithmic sampling frequencies. Different algorithms provide large or minor differences between them. Baker [1], [2], introduced the Stochastic Universal Sampling (SUS). The idea is to make a single draw from a uniform distribution and use it for determining the exact number of copies from each parent and. He also defines the *bias* B as the absolute difference between an individual's actual sampling probability and its expected value η_i. Also he defined *spread* as the range of possible values for the number of copies an individual receives by a selection mechanism. Consequently the minimum spread allows a bias B = 0 and for it, the following assertion holds,

$$n_i \in \{\lfloor \eta_i \rfloor, \lceil \eta_i \rceil\}$$

According to the Baker's report SUS provides a bias B = 0 and minimum spread.

Grefenstette [9] introduced another kind of *bias* b (called *bias measure*) to define the *population diversity* as follows:

$$b(P(t)) = \frac{1}{l \cdot \mu} \sum_{j=1}^{l} max \left(\sum_{\substack{i=1 \\ a^t_{i,j}=0}}^{\mu} (1 - a^t_{i,j}), \sum_{\substack{i=1 \\ a^t_{i,j}=1}}^{\mu} a^t_{i,j} \right)$$

where l is the chromosome length and a^t_{ij} denotes the allele value.
The bias b ($0.5 \leq b \leq 1.0$) indicates the average percentage of the most outstanding value in each position of the individuals. Smaller values of b indicate higher genotypic diversity and vice versa. Bäck and Hoffmeister [3] used this concept to establish genotypic diversity. The bias b can be used to formulate an adequate termination criterion. For example, the search process can be stopped when b reaches a value near to 1 because at this time the genotypic diversity is very low indicating population convergence.

2. Experiments description

For experimentation a test suite of four functions with varied difficulty was selected:

f2: Michalewickz's highly multimodal function

$f(x_1,x_2) = 21.5 + x_1 \cdot sin(4\pi \cdot x_1) + x_2 \cdot sin(20\pi \cdot x_2)$, *for* :

$-3.0 \le x_1 \le 12.1, \quad 4.1 \le x_2 \le 5.8$

estimated maximum value : 38.850292

f3: Easom's Function

$$f(x_1,x_2) = -\cos(x_1)\cdot\cos(x_2)\cdot e^{\left[-((x_1-\pi)^2+(x_2-\pi)^2)\right]},$$

donde $-100 \le x_1, x_2 \le 100$

minimo global

$f(x_1,x_2) = -1, \quad en \ (x_1,x_2) = (\pi,\pi)$

f5: Branins's Rcos Function

$$f_4(x_1,x_2) = \left(x_2 - \frac{5.1}{(4\cdot\pi^2)}\cdot x_1^2 + \frac{5}{\pi}\cdot x_1 - 6\right)^2 + 10\cdot\left(1-\frac{1}{(8\cdot\pi)}\right)\cdot\cos(x_1) + 10,$$

$x_1 = -5:10, \ x_2 = 0:15;$

minimum global value: 0.397887

f6: Griewangk's Function F8

$$f_1(x_i) = 1 + \sum_{i=1}^{5}\frac{x_i^2}{4000} - \prod_{i=1}^{5}\left(\cos\left(\frac{x_i}{\sqrt{i}}\right)\right),$$

$x_i = -600:600, \ i = 1:5;$

minimum global value : 0.0

f7: Schaffer's Function

$$f(x_1,x_2) = \frac{1}{2} + \frac{\left(\left(sin\sqrt{x_1^2+x_2^2}\right)^2 - \frac{1}{2}\right)}{\left(1+0.001\left(x_1^2+x_2^2\right)\right)^2}$$

x_1=-100 : 100 x_2=-100 : 100

minimun global value : 0

For each experiment 20 runs with randomised initial population of size fixed to 80 individuals were performed on each function, using binary coded representation, elitism, one point crossover and bit flip mutation. The number maximum of generations was set to 1000 and probabilities for crossover and mutation were fixed to 0.65 y 0.001 for *f2*, and to 0.5 and 0.005 for *f5*, *f6* y *f7* .

To establish the behaviour of PS and SUS, the following performance variables were considered:

B: The Baker's bias , to determine the difference between an individual's actual sampling probability and its expected value η_i.

b: The Grefenstette's bias to determine genotypic diversity.

Ebest = (Abs(*opt_val* - best value)/*opt_val*)100
It is the percentile error of the best found individual when compared with the known, or estimated, optimum value *opt_val*. It gives us a measure of how far are we from that *opt_val*.

3. Conclusions

For any of the functions of the testing suite, the bias *B* is quite near to zero when Stochastic Universal Sampling is used. This indicates that there almost do not exists differences between the expected number of offspring for each individual and the effective sampling frequencies. Also a reasonable genetic diversity is preserved even at the final stages with values of *b* ranging between 0.6 to 0.8.
Regarding to Proportional Selection most of the cases show a value of *B* near to 0.8. This is an expected result due to the limited population size. In this case the genetic diversity is low with values of *b* ranging between 0.8 to 1.

4. Bibliography

[1] Baker J. E.:*Adaptive selection methods for genetic algorithms.* . Proc. Of the 1st International Conf. On Genetic Algorithms and Their Applications, pp 101-111. Lawrence Erlbaum Associates, 1985.

[2] Baker J. E.: *Reducing bias and inefficiency in the selection algorithm.* Proc. Of the 2nd International Conf. On Genetic Algorithms, pp. 14-21. J.J Grefenstette Ed, 1987.

[3] Bäck T., Hoffmeister F.: *Extended selection mechanisms in genetic algorithms.* Proc. of the 4th Int. Conf. on Genetic Algorithms, pp 92-97, Morgan Kaufmann, 1991.

[4] Bäck T.: *Selective pressure in evolutionary algorithms: a characterization of selection mechanisms.* Proc. 1st IEEE Conf. On Evolutionary Computation (Orlando, FL, 1994) (Piscataway, NJ: IEEE) pp 57-62.

[5] Blickle T. and Thiele, L.: *A comparison of selection schemes used in genetic algorithms.* Technical Report 11 TIK, Swiss Federal Institute of Technology, December 1995.

[6] Brindle A.: *Genetic algorithms for function optimization.* Doctoral Dissertation. University of Alberta, Edmonton, 1981.

[7] Esquivel S., Leiva A., Gallard R. –"Selection Mechanisms in Evolutionary Algorithms"- "- in Fundamenta Informaticae, IOS Press, vol 35, No 1-4, pp 17-33. Aug. 1998.

[8] Goldberg D. E. And Deb K.: *A comparison of selection schemes used in genetic algorithms.* Foundations of Genetic Algorithms 1 . Morgan Kaufmann, pp 69-93.

[9] Grefenstette J.J.: A user's guide to GENESIS. Navy Center for Applied Research in Artificial Intelligence, Washington D.C., 1987.

[10] Michalewicz, M.: *Genetic Algorithms + Data Structures = Evolution Programs.* Springer, third revised edition, 1996.