

A COMPARISON OF FITNESS SCALING METHODS IN EVOLUTIONARY ALGORITHMS

Bertone E., Alfonso H.
Proyecto UNLPAM-09/F009¹
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: {
bertonee,alfonsoh}@ing.unlpam.edu.ar;
Phone: (02302)422780/422372, Ext. 6412

Gallard R.
Proyecto UNSL-338403²
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

Proportional selection (PS), as a selection mechanism for mating (reproduction with emphasis), selects individuals according to their fitness. Consequently the probability of an individual to obtain a number of offspring is directly proportional to its fitness value. This can lead to a loss of selective pressure in the final stages of the evolutionary process degrading the search.

This presentation discusses performance results on evolutionary algorithms optimizing two highly multimodal (Michalewicz's and Griewank's) functions and a hard unimodal (Easom's) function. Experiments were addressed to contrast the behaviour of a simple genetic algorithm against three scaling methods: linear, sigma truncation and recency-weighted-running-average. Diverse measures of performance were used to establish quality of results and convergence speed.

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).

From an static point of view, one pernicious consequence of this assignment of probabilities resides in the different behaviour showed by an evolutionary algorithm for functions that are equivalent from the optimization point of view such as $f(x) = ax^2$ and $g(x) = ax^2 + b$. For example, if for certain values of x , it results $b = -ax^2$ then the selection probabilities of many individuals would be extremely similar and the selective pressure would result too weak. Consequently optimization of $g(x)$ becomes a random search process.

From a dynamic point of view, as long as the evolutionary process progresses, the population often becomes dominated by super-individuals with a narrow range of objective values. In this condition all members of the population have similar fitness value leading to a loss of selective pressure towards the better individuals.

To avoid this undesirable behaviour the fitness function can be scaled (De Jong [2]) to the worst individual and instead of absolute individual's fitness, we manage with an individual's fitness relative to the worst individual.

But on the other hand, when scaling to the worst individual, the inverse effect (excessive selective pressure) can occur as long as a super-performer appears in the population. Copies of this super-individual will rapidly invade the whole population.

Scaling methods try to cope with problems, which are dependent on the characteristics of the function being optimized and earliest works are due to Bagley [1], Rosenberg [6] and Forrest [3]. Different categories of scaling were defined.

Goldberg presented; *linear*, *sigma truncation* and *power law* scaling [4]. Grefenstette [5] defines a scaling approach by using a fitness function as a time varying linear transformation considering the worst value seen in the last generations. Let us call this approach *time varying linear* scaling. But this method showed to be sensitive to 'lethals', poorly performing individuals arising from crossover and mutation. Better scaling was achieved by using a *recency weighted running average* of the worst observed objective values.

2. Scaling methods

Time varying linear scaling

Grefenstette defines his scaling approach using a fitness function as a linear time dependent transformation.

$$\phi(a_i(t)) = \alpha \cdot f(a_i(t)) - \beta(t)$$

where $\alpha = 1$ for maximization and -1 for minimization and $\beta(t)$ represents the worst value observed so far.

Sigma truncation

Here the scaled fitness is given by

$$\phi(a_i(t)) = \begin{cases} f(a_i(t)) - (f(t) - c\sigma_f(t)) & \text{if } f(a_i(t)) > (f(t) - c\sigma_f(t)) \\ 0 & \text{otherwise} \end{cases}$$

where c is an small integer (between 1 and 5) and σ_f is the standard deviation of the population. possible negative evaluations are set to zero.

Recency weighted running average.

Linear scaling is sensitive to 'lethals', bad performer individuals that occasionally appear.

The *recency weighted running average* method provides a smoother scaling and uses the following estimation for the scaled fitness

$$\beta(t) = \delta \cdot \beta(t-1) + (1-\delta) \cdot f_{\text{worst}}(t)$$

with $0 \leq \delta \leq 1$, where δ measures the adaptation speed and $f_{\text{worst}}(t)$ is the worst value observed until time t .

3. Experimental tests

The above mentioned scaling methods were contrasted against a simple genetic algorithm. For our experiments, 30 runs with randomised initial population of size fixed to 50 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 0.25 and 0.01 for $f1$ and 0.65 and 0.001 for $f2$.

The stop criterion was the stability of the mean population fitness.

$f1$: Michalewicz's highly multimodal

$$f(x_1, x_2) = 21.5 + x_1 \cdot \sin(4\pi \cdot x_1) + x_2 \cdot \sin(20\pi \cdot x_2), \text{ for } ;$$

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

estimated maximum value : 38.850292

$f2$: Griewangk's highly multimodal

$$f(x_1, \dots, x_5) = 1 + \sum_{i=1}^5 \frac{x_i^2}{4000} - \prod_{i=1}^5 \left(\cos \left(\frac{x_i}{\sqrt{i}} \right) \right),$$

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

minimum global value : 0.0

As an indication of the performance of the algorithms the following relevant variables were chosen:

$$\mathbf{Ebest} = ((opt_val - \text{best value}) / opt_val) 100$$

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 .

$$\mathbf{Epop} = ((opt_val - \text{pop mean fitness}) / opt_val) 100$$

It is the percentile error of the population mean fitness when compared with opt_val . It tells us how far the mean fitness is from that opt_val .

Time: It is the running time in seconds to arrive to the terminating generation

4. Conclusions

This work contrasted the behaviour of a genetic algorithm with and without scaling on a selected set of multimodal testing functions.

After analysis, the results indicate that all scaling methods produce individuals of better quality when they are compared with those provided by a simple genetic algorithm without scaling. Best mean Ebest values were achieved under sigma truncation.

Observing Epop values we noticed that when any of the scaling methods is applied on **highly multimodal optimization, the final population remains near the best found value.**

When optimizing *f1*, *linear scaling* and *recency weighted running average* increment running time in about 35% when compared with the simple genetic algorithm, while *sigma truncation* reduce this value in a 44%. When optimizing *f2* no sensible differences can be detected.

As a final conclusion we can say that even if scaling methods are not widely used, their application become beneficial when quality of solutions is considered an important issue in multimodal functions optimization.

5. Bibliography

- [1] Bagley J. D. The behaviour of adaptive systems which employ genetic and correlation algorithms. Doctoral dissertation, University of Michigan. Dissertation Abstracts International, 28(12), 5106B. university Microfilms No. 68-7556. 1967.
- [2] De Jong K. A.: *An analisis of the behavior of a class of genetic adaptive systems*. Doctoral dissertation, University of Michigan. Dissertation abstract International, 36(10), 5140B. (University Microfilms no 76-9381).
- [3] Forrest S. Documentation for Prisoners Dilema and Norms programs that use the genetic algorithm. Unpublished manuscript, University of Michigan, Ann Arbor, 1985.
- [4] Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA. (1989).
- [5] Grefenstette J.J.: Optimization of control parameters for genetic algorithms. IEEE Trans. Syst. Man Cybernetic. SMC-16, pp 122-128, 1986.
- [6] Rosemberg R. S. Simulation of genetic populations with biochemical properties. Doctoral dissertation, University of Michigan. Dissertation Abstracts International, 28(7), 2732B. university Microfilms No. 67-17863. 1967.

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

Minetti G., Salto, C., Alfonso H.
Proyecto UNLPAM-09/F009¹
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-338403²
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).