Genetic Algorithm Using Migration and Modified GSX as Support in the Diversification of Populations.

Francisco Ornelas[1], Alejandro Padilla[1], Alberto Ochoa[2], Eunice Ponce de León[1], Felipe Padilla[1]

[1]Departamento de Ciencias de la Computación, Universidad Autónoma de Aguascalientes, México
[2]Inst. de Ing. Y Tec., Universidad Autónoma de Ciudad Juarez, México
fjornel@correo.uaa.mx

Abstract. This research presents a study of some of the operators that have been used in recent years to diversify populations during the evolutionary process of genetic algorithms. This allows us to determine the migration and proper selection and crossing, significantly enhance the quality of solutions obtained, because it prevents premature convergence, resulting from the loss of changes in the genetic material of individuals by inbreeding between them and significantly prevents the migration.

Keywords. Genetic Algorithms, Migration, GSX,

1 Introduction

Genetic Algorithms (GA) are search procedures based on natural selection and genetics, which have proven successful in various applications involving search, optimization and machine learning [1], [2] and its use has been on increasing in both academics and industry [6].

The (GA) working with populations of individuals (possible solutions) that compete to preserve their characteristics from generation to generation. The preservation process is given by some operators who are: evaluation, selection, crossover and mutation.

The evaluation is to qualify each individual according to how well solves the proposed problem. That is, if it comes to maximizing a function with a higher score qualify individuals who obtain a higher result, while minimizing it in the case of a higher score is given to individuals who receive fewer results.

The selection is a process that involves selecting a number of individuals (usually defined depending on the type of problem and the number of individuals) that best solve the problem to be solved (this type of selection is called elitist selection, but there other types such as roulette, tournament, etc.).

The cross is the process which combines the genetic material of two individuals selected (usually from among the best) and usually occurs one or more new individuals that are expected to better solve the problem. There are several different types of crosses that are also used based on the characteristics of the problem, among which crosses a point, multi-point crosses, GSX, among others.

Finally, in the simple GA have the mutation that is a process through which it converts part of the genetic material of one or more individuals that results in new solutions that can enable us to explore different areas of our search space, but not necessarily produce better results than those already obtained.

Unfortunately, this operator is recommended to be used in very small percentages [1], because it is a highly random and does not assess whether the solution to be modified is one of the best results achieved so far or the solution in which turn is better quality.

This paper intends to show some evaluation results to operators which are related to the diversification of populations and migration operator (which is implemented in some models GA) to propose some improvements to this process and increase the performance of the GA.

2 Assessment and Selection in the GA

The simplicity of the GA allows it to be easily adapted to almost any problem. The most difficult thing is to create an evaluation function for taking into account all factors involved to solve the problem and to represent them through a function.

The evaluation function is central to the functioning of the GA.

Once you define a good evaluation function is a selection of individuals who were part of the next generation and will have opportunity to cross for part of their genetic material is preserved.

To make the selection process there are several models proposed (we assess three that are elitism, roulette wheel and stochastic universal sampling), the best known is the elitist, that consists in selecting the best individuals of each generation.

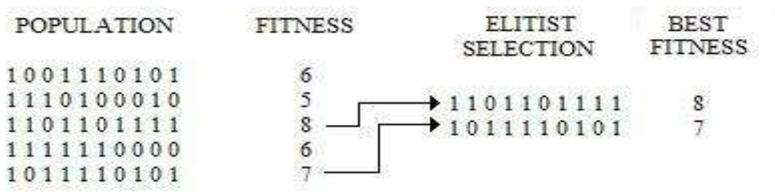

Image 1 Elitist selection.

The main problem we encountered in our research with this type of selection is the fact that some populations are so-called super individuals taking full control of the population and its quality is higher than the rest of the population (though not necessarily the best solution to find), which causes all individuals quickly begin to be part of the genetic material of this super individual that results in rapid convergence (typically a local optimum).

Another type of selection is the method of roulette that is allocated within an imaginary roulette each individual area of population is proportional to the quality with which solves the problem, so individuals get better results in the evaluation function will have a greater area within the spinner. Once this is set to 100% of roulette random values are obtained they determined that individuals are selected.

The selection by this method allows greater diversification than elitist method even though the fittest individuals have a greater chance of being selected.

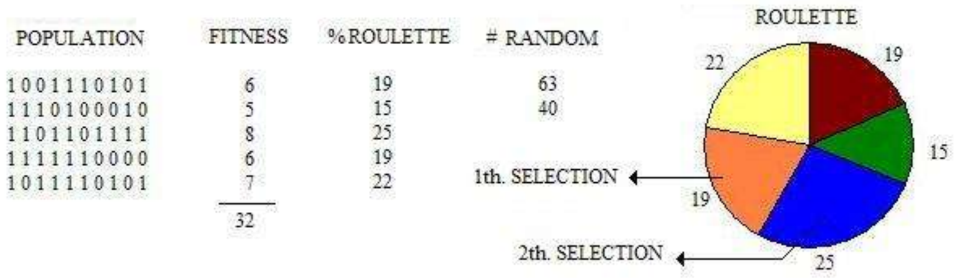

Image 2 Selection by the method of roulette

Finally the stochastic universal sampling selection [4] [5], is as in the previous selection in generating an imaginary roulette represent based on the area assigned to each individual's quality that it solves the problem, but main the main difference is that only the first individual is randomly selected and the rest are selected by dividing the wheel between the number of missing individuals which allows for greater uniformity in the selection and accepting both good and bad individuals.

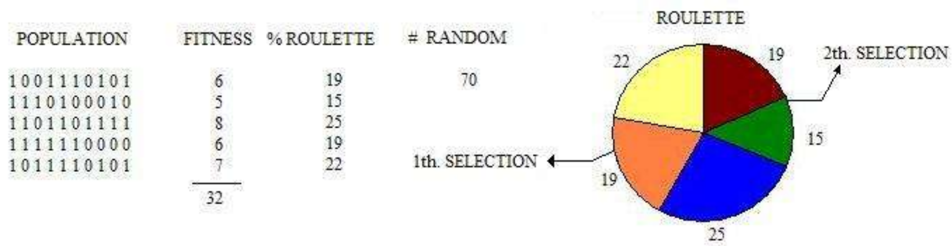

Image 3 Stochastic universal sampling selection.

Based on the study we did make use of an elitist approach to select not more than 10% of individuals to pass on to the next generation (and not lose good solutions that have thus far), and supplement it with other individuals obtained through a uniform universal sample selection (to provide greater diversification).

3 Cross Operator

As in the selection operator, the operator of crosses has different models that are used according to the characteristics of the problem and the type of representation with individuals (binary, alphabetic, alphanumeric, etc.).

To make the crosses are taken two individuals at random from the previously selected set of solutions in the previous process.

The most simple and well across is the crossing of a point. The new solution will consist of the elements containing one of the selected individuals (parents) from the start of it to a position (usually chosen at random) as a reference and take the other parent from this reference point until the end of same. This type of cross I may produce one or more new individuals taking the various parts of the parent solutions.

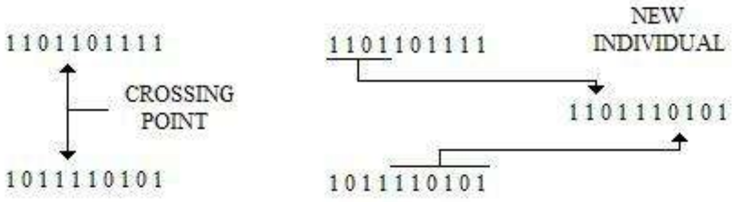

Image 4 Crosses from one point

Existen variantes de este tipo de cruza que consisten normalmente en incrementar los puntos de cruza e ir generando al nuevo individuo alternando las partes en que fueron divididos los padres.

The other type of crosses that were evaluated was the GSX (Greedy subtour Crossover) [6], which consists of randomly selecting a common element of both parents, and this element becomes part of the new individual. As a next step element is shifted to the left of the first parent and to the right of the second, and assess if current element is not already contained in the new individual, this process continues until some element and the content on the new individual or until no more elements are in one or both parents. Finally completed the new individual with the other elements that are not yet in the same either randomly or based on one of the parents and are inserted as they appear in it.

This type of cross subtours preserves that some problems are of vital importance such as the traveling salesman problem TSP.

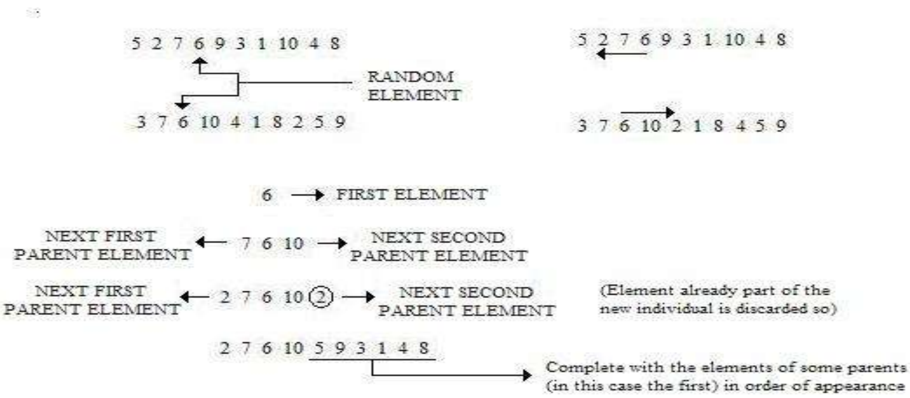

Image 5 GSX

In our research we believe that the GSX is a very important tool that has demonstrated excellent performance. However, a problem that we see is that sometimes the initial element is at the end of the father that causes almost immediately stop the construction of new individual and end up using a random process that occurs in most cases very different individual parents or used as a base to one parent, in which case the new individual would be almost identical to this one.

We propose a modification in the operation of the GSX and described below.

4 GSX Modified (GSXM)

As discussed above the GSX is a great tool for generating new individuals by crossing it allows parties retain complete genetic code from both parents. Unfortunately sometimes the way he works the GSX these pieces of code are very small and you end up with very different individuals parents (losing the desirable characteristics of the goods) or too equal to one parent (which also is desirable).

In nature, individuals are composed of equal parts of both parents' genetic code, which is why we try to modify the GSX to treat it as close as possible to this model. To this end we determined that the first thing we should change is that at least we had the chance to get half of each parent. For these reasons we select the most near the middle of one parent and evaluated against the other parent route also from closer to the middle element of it in the opposite direction as it will cross to begin forming the new individual, this until we find an equality between elements.

Once you have determined the item on the launch of the cross is the same procedure as in the GSX with two main variants. The first is that this process does not stop until it is reaches the end of one or both parents and if at this point there are still elements missing in the new individual, these are taken from one parent in the order they appear therein. The second is closely related to the first so that if an element is repeated this simply is not taken into account and continues into the following elements.

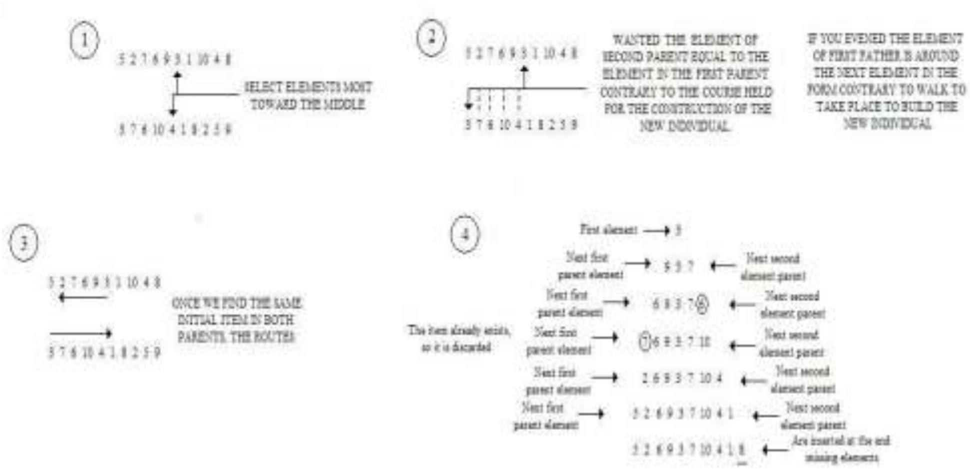

Image 6 GSXM

This change allows us to maintain subtour larger and larger number of subtour occasions that problems like the TSP is important.

5 Migration

Another operator we use in our research is migration.

Migration is a concept used in population genetics that is the area of genetics that deals with the study of the evolution of populations in nature and the factors involved in this process.

There is some work done with various proposals on the use of this operator in GA, one of them is the island model of dividing the initial population in subpopulations that evolve in parallel and each given number of generations send a copy of their best individuals of other stocks, that on certain criteria, some of which are similar to network topologies. Examples of such criteria are [4]: communication in star, which has a central population receives the best individuals from other populations and feedback with their best individuals, communication in mesh, which allows all people share their best individuals and communication ring, where people with the following individuals share with the people who have communication and receive feedback from the previous population in the connection.

The model we're using consists of two populations evolving in parallel. One is the main population that is receiving the best individuals of the population and in turn the secondary school population updates its population by introducing new individuals.

The migration process is held every nine generations, this so that individuals are completely discarded or strengthen their genetic material within the population.

We do not make the exchange of genetic material between the two populations based on our experiences because we have found that if a super individual not only takes over the population of which is generated, but to be the best evaluated moves to the other population and also takes control of it and even though the type of evaluation and selection we use largely avoids this situation if you can get to take control of it.

The new individuals enter in the second population can not stop the evolution of populations on search sites which include, but is followed by an exploration of it.

The migration rate we use is of a maximum of 10% to ensure that new individuals only for improvements on the genetic material without removing the original characteristics of the population.

6 Findings

In this paper we present some recommendations on the contractors carrying out the process of diversification in the evolutionary process in GA.

First, based on our research we found that an elitist selection causes a rapid loss of genetic diversity, because the better or best individuals take control in the evolutionary process and the operator of mutation is not sufficient to avoid this situation, what we use elitism in a very small amount and the complement to a process of selection and stochastic universal sampling has yielded good results.

It has also been found that the mutation operator is necessary but must be supplemented to improve genetic diversity. From the above we use the migration operator to comply with this process.

As part of research is developing a tool that meets all the characteristics mentioned above for verification if together give the same good results that modifications or improvements separately.

Bibliography

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

[2]Goldberg, D. The Design of Innovation: Lessons from and for Competent Genetic Algorithms. Kluwer Academic Publishers. 2002.

[3]Holland, J. Adaptation in Natural and Artificial Systems. The MIT Press. 1975

[4] Moujahid A., Inza I., Larrañaga P. Tema 2: Algoritmos Genéticos. Universidad del País Vasco. http://www.sc.ehu.es/isg/

[5] Gonzalez U., Solanas. A. Proteccion de la Privacidad mediante Microagregación Multivariante basada en Algoritmos Genéticos: Selección por Ruleta vs. Selección Uniforme

[6] Winward P. Fluctuating Crosstalk, GA Scalability, and the Disruption of Optima. University of Illinois at Urbana-Champaign. 2007