ON THE SHORT-TERM BEHAVIOUR
OF GENETIC ALGORITHMS

Dorin BOCU
Transilvania University of Brasov,
Romania

Abstract: The present paper is the first from a series of works dealing with the thematic scheme presented below:

      1. Formalization of the transient aspects of the genetic paradigm.

      2. Formalization of the asymptotic aspects of the genetic paradigm.

      3. Applications in search and optimization problems.

The immediat purpose of the paper is the critical evaluation of the approaches mentioned in the literature concerning the transient aspects of the genetic paradigm. Our goal is to eliminate ambiguities encountered in some works that use concepts like "population as multi-set".

At the same time, problems already formulated or new problems concerning genetic algorithms are treated in a uniform manner.

Keywords: Population, schema, oriented search space.

1. INTRODUCTION

First introduced by J. H. Holland [1], genetic algorithms were inspired by some of the selection and natural genetic mechanisms.

A possible reason for the researcher's option for a "paradigm" of the genetic algorithms may be briefly formulated as follows:

"Selection and natural genetics are the key to overall success obtained by the living world in its attempt to maintain its status-quo and adapt its existence in hostile ebvironments". "Overall" success is used because the survival of the living world is marked by failures; according to the uncompromising law of natural selection what is not fit inside a homogenous population and is not fit for the exigences of existence in eco-systems must be eliminated for a short term and undergo a long-term adapting process.

The concepts under discussion and their meanings are those already well-known in this field of interest: population, schema, selection operator, crossover operator, mutation operator, etc.

In this paper some of these concepts are presented using a formalism which combines simplicity with rigour.

In this context the essential aspects of genetic paradigms are discussed concerning implementation of the selection operator.

2. FORMALIZATION OF TRANSIENT ASPECTS OF GENETIC PARADIGMS

Definition 2.1. Given $A = \{0,1,...,\alpha-1\}$, $\alpha \geq 2$; $L \in N^*$; then: $\Omega = \{x_1, x_2,..., x_L$ $x_i \in A(\forall)1 \leq i \leq L\}$ is called <u>search space</u>; $x_1, x_2,..., x_L \in \Omega$ is called <u>point</u>; L is called the dimension of Ω. We use the following the notation: $Dim(\Omega) = L$.

<u>Remarks:</u>

$1°$ $\Omega = A \times A \times ... \times A$; therefore $|\Omega| = |A| \times |A| \times ... \times |A| = \alpha^L$

$2°$ considering $x_1, x_2,..., x_L$ an unsigned integer

$\Rightarrow \Omega = \{0,1,...,N-1\}$, $N = \alpha^L$; $i \in \Omega$ is called individual.

Definition 2.2. Let $M \geq 1$ $f: \{1,2,...M\} \rightarrow \Omega$ is called Ω – application. We denote $F = \{f: \{1,2,...,M\} \rightarrow \Omega\}$.

Definition 2.3. For $i \in \Omega$, $P(i) = \{j \in \{1,2,..., M\}, f(j) = i\}\}$ is called the weight of individual i, related to the Ω – application f.

Definition 2.4. Given $f, g \in F$; $f \sim g \xleftarrow{\text{by definition}} P_f(i) = P_g(i)$ $(\forall)i \in \Omega$.

Proposition 2.1. "~" is an equivalence relation.

P r o o f. $1°$ Reflexivity of the relation '~' i   vial.

$2°$ $\left. \begin{array}{c} f, g \in F \\ f \sim g \end{array} \right\} \Rightarrow P_f(i) = P_g(i)(\forall)i \in \Omega \Rightarrow P_g(i) = P_f(i)(\forall)i \in \Omega \Rightarrow g \sim f$

$\Rightarrow$ relation '~' is symetric.

$3°$ $f, g, h \in F$; $(f \sim g) \wedge (g \sim h)$

$\Rightarrow (P_f(i) = P_g(i)\ (\forall)i \in \Omega) \wedge (P_g(i) = P_h(i)\ (\forall)i \in \Omega)$

$\Rightarrow P_f(i) = P_h(i)\ (\forall)i \in \Omega) \Rightarrow f \sim h$.

Definition 2.5. We call polulation the N-touple $P = <n_0, n_1,..., n_{N-1}>$; $n_i = P_f(i)$; $i \in \Omega$, $f \in F$; where f is the representative of a class of equivalence of $F / \sim$.

Remark

It is obvious that the set of populations containing M individuals, that can be constructed with elements of Ω is of cardinal $|F/\sim|$.

Definition 2.6. Given the search space Ω, we call <u>schema</u> a word of lenght L over the alphabet $A \cup \{*\}$. $S\{\Omega\}$ is a notation for the set of schemas which can be constructed starting from Ω.

Definition 2.7. We say that the point $x_1, x_2, ..., x_L \in \Omega$ agrees with the schema $s = s_1 s_2 ... s_L \in \{A \cup \{*\}\}$ if $s_i = x_i \ (\forall) i \in \{1,2,...,L\}, s_i \neq *$.

In certain considerations a generalization of the schema notion is useful, namely the predicate notion. In this paper we do not refer to this aspect.

Definition 2.8

Let Ω, $Dim(\Omega) = L$; For $s \in S(\Omega)$; $s = s_1 s_2 ... s_L$ we define:

1° the <u>order</u> of the schema s by $O(s) = | A \cap \{s = s_1 s_2 ... s_L\}|$

2° the <u>defining length</u> of the schema s by $\delta(s) = LDP(s) - FDP(s)$, where:

$LDP(s)$ is a function which returns the last position different from '*' in s;

$FPD(s)$ is a function which returns the first different position from '*' in s.

Definition 2.9. Let Ω be as described in Definition 2.1, the population $P = <n_0, n_1, ..., n_{N-1}>$ and $s \in S(\Omega)$. We call representation of s in P the following:

$$R_p(s) = \sum_{i=0}^{N-1} n_i \varphi(E_{\alpha, L}(i), s), \text{ where:}$$

$E_{\alpha, L}(i)$ is the equivalent in base α on L positions of integer i,

$$\varphi(E_{\alpha, L}(i), s) = \begin{cases} 1 & \text{if } E_{\alpha, L}(i) \text{ concording with } s \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.10. We call the objective function associated to the search space Ω the function $f: \Omega \to R$; f-injective. The pair (Ω, f) is called <u>oriented search space</u>.

Remark.

If $x \in \Omega$ then $f(x)$ is called the utility or the adaption of x to a given population.

Definition 2.11. Given (Ω, f); $\Omega = \{0, 1, ..., N-1$; $N = \alpha^L$; $H \in S(\Omega)$; $P = < n_0, n_1, ..., n_{N-1} >$ we call the utility of schema H in population P the following:

$$f_p(H) = \frac{1}{R_p(H)} \sum_{i \in \Omega} \varphi(E_{\alpha, L}(i), H) \cdot n_i \cdot f(i).$$

Definition 2.12. We call <u>selection operator</u> the procedure, through which an element $i \in \Omega$ for which in $P_t = < n_0^t, n_1^t, ..., n_{N-1}^t >$ we have $n_i^t \neq 0$ is designated potential constituent of the population $P_{t+1} = < n_0^{t+1}, n_1^{t+1}, ..., n_{N-1}^{t+1} >$.

<u>Remarks</u>

$1°$ We say that an individual of some population P_t s designated a potential constituent of population P_{t+1} meaning that it is selected for the possible application of crossover and mutation operators (defined in the following), operators through which we obtain descendents, starting from two individuals, that permit the orientation of the search versus new points of Ω.

$2°$ To implement the idea of the survival of individuals with special aptitudes, the selection of an individual to participate in the generating process of the succesive states of a population is mostly proportional to the value of the objective function associated to the individuals as follows:

If $P_t = < n_0^t, n_1^t, ..., n_{N-1}^t >$ and Φ is a selection operator, then we have:

$K := \Phi(P_t, RND)$ where:

K is the returned value by the selection operator;
RND is a function which generates a random number according to a given probabilistic law (in a lot of examples RND returns a number from interval $[0,1]$ in a steady random manner).

$3°$ Because:

$$\min_{x \in D} \{f(x)\} = f(x^*) = -\max_{x \in D} \{f(x)\}$$

and because if

$$\left. \begin{array}{l} A = \{x^* \in D; \max_{x \in D} \{f(x)\} = f(x^*)\} \\ B = \{y^* \in D; \max_{y \in D} \{f(y) + c\} = f(y^*) + c\} \end{array} \right\} \Rightarrow A = B$$

we can always perform transformations after which the utilities of the individuals are strictly positive.

This is the reason why we consider that a standard genetic search problem is a problem of the type:

$$\begin{cases} \max\{f(i); i \in \Omega\} \\ 0 < f(i) < \infty; (\forall)i \in \Omega, f \text{ injective} \end{cases}$$

Under the above circumstances we have:

Theorem 2.1. Let $P = <n_0, n_1, ..., n_{N-1}>$ for which the utility vector is $F = <f_0, f_1, ..., f_{N-1}>$ where:

$$f_i = \begin{cases} f(i) > 0; n_i \neq 0 \\ \\ 0 \quad n_i = 0 \end{cases} \qquad i = \overline{0, N-}$$

Then, the algorithm:

$NrA := RND;$

$$Sel := \min\left\{ \frac{f_0 + f_1 + \cdots + f_{i-1}}{\bar{f}} \leq NrA < \frac{f_0 + f_1 + \cdots + f_i}{\bar{f}}; i \in \{1,2,...,N-1\}, n_{i-1} \neq 0 \right\}$$

ensures a proportional selection with the utility according to the folowing conditions:

RND is a function which returns a number in [0,1] randomly and uniformly distributed:

$$\bar{f} = f_0 + f_1 + \cdots + f_{N-1}$$

P r o o f. It must be shown that:

$$P(Sel = i) = \frac{f_i}{\bar{f}} \qquad (\forall)i; n_i \neq 0.$$

RND function generates uniformly random numbers in [0,1].

So, the probability density of the random variable associated to the generation procedure is:

$$p(x) = \begin{cases} 1; & x \in [0,1) \\ 0; & x \in R \setminus [0,1) \end{cases}$$

where ρ is the integrable function $\Rightarrow$ and the corresponding repartition function of ρ is:

$$Q(x) = \int_{-\infty}^{x} \rho(t)dt = \begin{cases} 0 & ;x \in (-\infty;0] \\ x & ;x \in (0,1) \\ 1 & ;x \in [1,\infty) \end{cases}$$

$$P(Sel = i) = P(\frac{f_0 + f_1 + \cdots + f_{i-1}}{\bar{f}} \leq NrA < \frac{f_0 + f_1 + \cdots + f_i}{\bar{f}}) =$$

$$= Q(\frac{f_0 + f_1 + \cdots + f_i}{\bar{f}}) - Q(\frac{f_0 + f_1 + \cdots + f_{i-1}}{\bar{f}}) = \frac{f_i}{\bar{f}} \quad (\forall)i \in \{1,2,...,N-1\}, n_i \neq 0$$

Definition 2.13. Let Ω; $Dim(\Omega) = L$; $x \in \Omega$, $x = x_1 x_2 \cdots x_L$; $x_i \in A; i = 1, L$; $y \in \Omega$; $y = y_1 y_2 \cdots y_L$; $y_i \in A; i = 1, L$;

$1°$ The operator characterized as follows is called single point crossover (SPC):

     a) It randomly selects a natural number $k \in \{1,2,...,L-1$;

     b) $\begin{array}{c} x_1 x_2 \cdots x_k x_L \\ y_1 y_2 \cdots y_k y_L \end{array} \xrightarrow{\text{SPM}} \begin{array}{c} x_1 x_2 \cdots x_k y_{k+1} \cdots y_L \\ y_1 y_2 \cdots y_k x_{k+1} \cdots x_L \end{array}$

$2°$ The operator characterized as follows is called the single point mutation (SPM) of element $x \in \Omega$:

     a) It randomly selects a natural number $k \in \{1,2,...,L$;

     b) $x_1 x_2 \cdots x_k \cdots x_L \xrightarrow{\text{SPM}} x_1 x_2 \cdots \overline{x_k} \cdots x_L; \; \overline{x_k} \in A \setminus \{x_k\}$
        $\overline{x_k}$ is randomly selected out of $A \setminus \{x_k\}$

Definition 2.14. Let Ω; $Dim(\Omega) = L$; $H \in S(\Omega)$; $H = x_1 x_2 \cdots x_L$; $x_i \in A \cup \{*\}$; $i = 1, L$. It is said that H is affected by the application of the SPC operator because $1 \leq k \leq L-1$ if $(\exists) \, x_i, x_j; i < j; j > k; (x_i \neq *)$ and $(x_j \neq *)$.

<u>Remark</u>

     It is easily proved that the probability that a schema H is affected by the application of the SPC operator is:

$$\overset{c}{P_{aa}} = \frac{\delta(H)}{L-1}.$$

Therefore the probability of survival when applying SPC operator with probability 1 is:

$$\rho_s = 1 - \frac{\delta(H)}{L-1}.$$

As the SPC operator itself is applied with probability p_c and the event which consists of the application of the SPC operator is independent of the event which consists of affecting a schema, it can be inferred that:

$$p_s^c \geq 1 - p_c \cdot \frac{\delta(H)}{L-1} \quad \text{where } p_s^c \text{ is the symbol for the probability of survival with}$$

the possible application of the SPC operator.

Definition 2.15. Let Ω; $Dim(\Omega) = L$; $H \in S(\Omega)$; $H = x_1 x_2 \cdots x_L$; $x_i \in A \cup \{*\}$; $i = 1, L$. It is said that H is affected by the application of the SPM operator on pisition k; $1 \leq k \leq L$ if $x_k \neq *$.

Remark

If the SPM operator is applied with probability p_m in an obvious way, the probability that H is affected is $p_m^{O(H)}$.

Therefore the probability of survival when applying operator SPM is:

$$p_s^m = (1 - p_m)^{O(H)}.$$

We can now present the central result of the mathematical theory of genetic algorithms, a result which not only gives information about the dynamics of representing a schema within a population for a short period, but also indicates the long-term tendencies related to the representation of a schema in the corresponding genetic population series.

Theorem 2.2. Let (Ω, f) bean oriented search space; $Dim(\Omega) = L$;. $\Omega = \{0, 1, ..., N-1\}$, $N = \alpha^L$

$$P_0 = < n_0, n_1, ..., n_{N-1} >; \quad H \in S(\Omega).$$

If it is considered that P_0 is changed under the influence of a canonical genetic algorithm, then:

$$E(R_{P_{i+1}}(H)) \geq R_{P_i}(H) \cdot \frac{f_{P_i}(H)}{f_{P_i}(\Omega)} \cdot (1 - \alpha(H, P_i));$$

$$\alpha(H, P_t) = p_c \cdot \frac{\delta(H)}{L-1} - O(H) \cdot p_m \text{ where } P_t, P_{t+1} \quad \text{designates} \quad \text{two} \quad \text{successive}$$

generations in the application of the canonical genetic algorithm.

P r o o f. Let $P_t = < \omega_0(t), \omega_1(t), \ldots, \omega_{N-1}(t) >; \omega_i(t) \geq 0; i = 0, N-1$. Obviously $R_{P_t}(\Omega) = \omega_0(t) + \omega_1(t) + \cdots + \omega_{N-1}(t)$.

If H is a scheme $\Rightarrow (\exists) P = < h_0(t), h_1(t), \ldots, h_{N-1}(t) >; h_i(t) \geq 0; i = 0, N-1$.

$$R_{P_t}(H) = h_0(t) + h_1(t) + \cdots + h_{N-1}(t).$$

$$h_i(t) = \varphi(E_{\alpha, L}(i), H) \cdot \omega_i(t); i = 0, N-1.$$

$$\Rightarrow \quad h_1(t) = \begin{cases} 0 & ; \omega_i(t) = 0 \\ \omega_i(t); (\omega_i(t) \neq 0) \wedge (\varphi(E_{\alpha, L}(i), H) = \cdot T \cdot) \\ 0 & ; (\omega_i(t) \neq 0) \wedge (\varphi(E_{\alpha, L}(i), H) = \cdot F \cdot) \end{cases}$$

By applying the selection operator to the P_t population in order to form the gene-pool we obtain:

$$E(R_{P_{t+1}}(H)) = R_{P_t}(H) \cdot \frac{f_{P_t}(H)}{f_{P_t}(\Omega)} \text{ where:}$$

$$f_{P_t}(H) = \frac{\sum\limits_{i=0}^{N-1} h_i(t) \cdot f(i)}{R_{P_t}(H)}; \quad f_{P_t}(\Omega) = \frac{\sum\limits_{i=0}^{N-1} \omega_i(t) \cdot f(i)}{R_{P_t}(\Omega)}$$

By successively applying the operators SPC and SPM to the individuals that form the gene-pool and taking account of the specific ways in which these operators can affect a schema, we have:

$$E(R_{P_{t+1}}(H)) \geq R_{P_t}(H) \cdot \frac{f_{P_t}(H)}{f_{P_t}(\Omega)} \cdot (1 - p_c \cdot \frac{\delta(H)}{L-1}) \cdot (1 - O(H) \cdot p_m).$$

As for low values of p_m we have:

$$p_s^m = (1 - p_m)^{O(H)} \approx 1 - O(H) \cdot p_m \text{ and as we also have:}$$

$$(1 - p_c \cdot \frac{\delta(H)}{L-1}) \cdot (1 - O(H) \cdot p_m) =$$

$$-O(H) \cdot p_m - p_c \cdot \frac{\delta(H)}{L-1} + p_c \cdot p_m \cdot O(H) \cdot \frac{\delta(H)}{L-1}$$

$$\geq 1 - p_c \cdot \frac{\delta(H)}{L-1} - p_m \cdot O(H)$$

$\Rightarrow$ the conclusion of theorem is immediate.

3. SUMMARY

The whole construction of the present paper leads to Theorem 2.1. and Theorem 2.2.

The role of selection in the implementation of genetic algorithms is interesting and is somehow complementary to the mutation and crossover operators.

While selection focuses on genetic search, mutation and crossover disclose new search areas starting from the potential offered by the current population.

The fundamental theorem of the scheme is the result which contains information related to the "phenomenology" of the short term passing from one population to another but also indicates possible problems which may occur regarding the long-term convergence of genetic algorithms.

REFERENCES

[1]   Bocu, D., "On the application of genetic algorithms in the optimization process", Presented at the *3rd Balkan Conference on Operational Research*, 16-19 October 1995, Thessaloniki, Greece.

[2]   Goldberg, D. E., *Genetic Algorithms in Search. Optimization Machine Learning*, Addison-Wesley, 1989.

[3]   Holland, J. H., *Adaption in Natural and Artifical Systems*, The University of Michigan Press, Ann Arbor, 1975.