Correspondence

Statistical Analysis of the Main Parameters Involved in the Design of a Genetic Algorithm

Ignacio Rojas, Jesús González, Héctor Pomares, J. J. Merelo, P. A. Castillo, and G. Romero

***Abstract*—Most genetic algorithm (GA) users adjust the main parameters of the design of a GA (crossover and mutation probability, population size, number of generations, crossover, mutation, and selection operators) manually. Nevertheless, when GA applications are being developed it is very important to know which parameters have the greatest influence on the behavior and performance of a GA. The purpose of this study was to analyze the dynamics of GAs when confronted with modifications to the principal parameters that define them, taking into account the two main characteristics of GAs; their capacity for exploration and exploitation. Therefore, the dynamics of GAs have been analyzed from two viewpoints. The first is to study the best solution found by the system, i.e., to observe its capacity to obtain a local or global optimum. The second viewpoint is the diversity within the population of GAs; to examine this, the average fitness was calculated. The relevancy and relative importance of the parameters involved in GA design are investigated by using a powerful statistical tool, the ANalysis Of the VAriance (ANOVA).**

***Index Terms*—ANalysis Of the VAriance (ANOVA), genetic algorithm (GA), statistical analysis.**

I. Introduction

The motivation of the present statistical study lies in the great variety of alternatives that a designer has to take into account when designing a genetic algorithm (GA). This decision is usually taken in terms of the most common values or experimental formulas given in the literature, or by means of trial and error [5], [8]. Nevertheless, it is very important to know which parameters involved in the design of a GA have the greatest influence on its behavior and performance [3], [11]. When analyzing the influence of each of these parameters, the designer should pay most attention to the one presenting the values that are statistically most significant. Thus, it should be possible to avoid the necessity for a detailed analysis of different configurations that might, in fact, lead to the design of various GAs with very similar behavior patterns. Although GAs have been applied to a wide range of difficult problems in numerous areas of science and engineering, there does not exist much theoretical or experimental analysis of the influence of the operators and the parameters (including interactions) involved in their design [4], [6]. Consequently, the goal of this correspondence is to obtain a better insight into which are the most relevant factors in GA design, in order to establish the elemental operations whose alternatives should be carefully studied when a real application is developed.

To do this, an appropriate statistical tool has been used: the multifactorial analysis of the variance [7], which consists of a set of statistical techniques that allow the analysis and comparison of experiments, by describing the interactions and interrelations between either the quantitative or qualitative variables (called *factors* in this context) of the system. In order to perform this analysis, the most relevant variables in the design of a GA have been taken as the main factors. Thus, population size, selection, crossover and mutation operators, the number of generations, and the crossover and mutation probabilities have been considered.

Manuscript received September 6, 1999; revised June 23, 2000, May 3, 2001, and February 26, 2002. This work was supported in part by the CICYT Spanish Project DPI2001-3219. This paper was recommended by Associate Editor W. Pedrycz.

The authors are with the Department of Computer Architecture and Technology, University of Granada, Granada 18071, Spain.

Publisher Item Identifier S 1094-6977(02)04681-3.

II. Application of ANOVA in the Design of a Genetic Algorithm

The ANalysis Of the VAriance (ANOVA) is one of the most widely used statistical techniques. The theory and methodology of ANOVA was developed mainly by R. A. Fisher during the 1920s [2]. ANOVA belies its name in that it is not concerned with analyzing variances but rather with analyzing the variation in means. ANOVA examines the effects of one, two, or more quantitative or qualitative variables (termed *factors*) on one quantitative response. ANOVA is useful in a range of disciplines when it is suspected that one or more factors affect a response. ANOVA is essentially a method of analyzing the variance to which a response is subject into its various components, corresponding to the sources of variation which can be identified [7]. Suppose the easy case that the number of factors affecting the outcome of the experiment is 2. We denote by $X_{i,\,j}$ $(i = 1, \ldots n1;\ j = 1, \ldots, n2)$ the value observed when the first factor is at the ith level and the second at the jth level. It is assumed that the two factors do not act independently and therefore that there exists an interaction between them. In this case, the observations fit the following equation:

$$X_{i,\,j,\,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,\,j} + \beta_j \tag{1}$$

where μ is the fixed effect that is common to all the populations; α_i is the effect associated with the ith level of the first factor; and $\varepsilon_{i,\,j,\,k^j}$ is the effect associated with the jth level of the second factor. The term $(\alpha\beta)_{i,\,j}$ denotes the joint effect of the presence of level i of the first factor and level j of the second one; this, therefore, is denominated the interaction term. The term $\varepsilon_{i,\,j,\,k}$ is the influence on the result of everything that could not be assigned or of random factors. The null hypothesis is proposed that each term of the above equation is independent of the levels involved; i.e., on the one hand we have the two equality hypotheses for the levels of each factor

$$\begin{aligned} &H_{01}\colon \alpha_1 = \cdots = \alpha_i = \cdots = \alpha_{n1} \\ &H_{02}\colon \beta_1 = \cdots = \beta_j = \cdots = \beta_{n2} \end{aligned} \tag{2}$$

and on the other hand, the hypothesis associated with interaction, which can be expressed in an abbreviated way as

$$H_{03}\colon (\alpha\beta)_{ij} = 0, \qquad \forall\, i,\, j. \tag{3}$$

The hypothesis of the equality of several means arises when a number of different treatments or levels of the main factors are to be compared. Frequently, one is interested in studying the effects of more than one factor, or the effects of one factor when certain other conditions of the experiment vary, which then play the role of additional factors.

With ANOVA, we test a null hypothesis that all of the population means are equal against an alternative hypothesis that there is at least one mean that is not equal to the others. We find the sample mean and variance for each level of the main factor. Using these values, we obtain two different estimates of the population variance. The first one is obtained by finding the sample variance of the n_k sample means from the overall mean. This variance is referred to as the *variance between the means*. The second estimate of the population variance is found

TABLE I
VARIABLES USED FOR THE STATISTICAL STUDY

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Population size	Decrement 40%	Decrement 20%	Nominal Value	Increment 20%	Increment 40%	
Number of Generations	Decrement 40%	Decrement 20%	Nominal Value	Increment 20%	Increment 40%	
Crossover probability: *pc*	0.4	0.5	0.6	0.7	0.8	
Mutation probability: *pm*	0.05	0.1	0.15	0.2	0.25	
Type of Crossover	One-point	Two-points				
Type of Mutation	Bit-flip	Inversion				
Type of Selection	Roulette Wheel	Elitist Roulette Wheel	Deterministic			
Type of Experiment	Knapsack problem	Prisoner's dilemma	Riolo function	Michalewicz function 1	Michalewicz function 2	Michalewicz function 3

by using a weighted average of the sample variances. This variance is called the *variance within the means*. Therefore, ANOVA allows us to determine whether a change in the measure of a given variable is caused by a change in the level of a factor or is just originated by a random effect. In this way, it allows us to distinguish between the components which cause the variations appearing in a set of statistical data and to determine whether the discrepancies *between* the means of the factors are greater than would reasonably be expected according to the variations *within* these factors.

The two estimates of the population variance are then compared using the F*-ratio* test statistic. Calculating the sum of the squares of the observations extended to the levels of all the factors (S_T) and the sum of squares within each level (S_R), and dividing S_T and S_R by the appropriate number of degrees of freedom ($D.F$), obtaining s_T and s_R, respectively, the F-ratio is computed as s_T/s_R. This calculated value of the F-ratio for each factor is then compared to a critical value of F of Snedecor with the appropriate degrees of freedom to determine whether we should reject the null hypothesis. When there is no treatment effect, the ratio should be close to 1. If a level of a main effect has a significant influence on the output variable (observed variable, in our case the error index), the observed value of F will be greater than the F-Snedecor distribution, with a sufficiently high confidence level (usually 95%). In this case, the null hypothesis is rejected and it is argued that at least one of the levels of the analyzed factor must affect the response of the system in a different way. The F-ratio test assumes normal populations with equal variance and independent samples. The analysis is sensitive to inequality of variance (heteroscedasticity) when the sample sizes are small and unequal and care should be taken in interpreting the results. The comparison between the F-ratio and the F-Snedecor distribution is expressed through the significance level (*Sig. Level*). If this significance level is lower than 0.05 then the corresponding levels of the factor are statistically significant with a confidence level of 95%.

The levels of a factor that are not statistically different form a homogeneous group and therefore the choice between the various levels belonging to a given homogeneous group has no significant repercussion on the response. Thus, once we discover that some of the factors involved in the design of an GA do not fulfill the null hypothesis, a study is carried out of the levels of this factor that may be considered statistically nonsignificant, using *multiple range test* tables for this purpose; these tables describe the homogeneous groups possible for each of the levels of the factor being analyzed.

III. Factors Considered in the Statistical Analysis

In the statistical study performed in Section V, the factors considered are the crossover and mutation probabilities, the population size, the number of generations, the type of selection, crossover, and mutation operators, and the type of experiment. Table I gives the different levels considered in each factor when carrying out multifactorial ANOVA (this is not a one-way ANOVA because we considered all the factors simultaneously). Each of these factors has different levels. For example, Roulette Wheel, Elitist Roulette Wheel, and Elitist Selection are the levels considered for the type of selection used for the reproduction process. The response variables used to perform the statistical analysis are maximum and average fitness in the last generation. The changes in the response variables are produced when a new combination of crossover probability, mutation probability, population size, etc. is considered, thus changing the design of the GA.

IV. Experimental Setup

Since a GA's performance (and its parameters setting) depends on the fitness function being optimized, problems of different types were used (one was combinatory, one was a strategy search, and others were numerical optimization, including multimodal and deceptive functions), in order to study the influence of the various factors on the solution found by the algorithm.

We have selected the following problems:

- 0/1 Knapsack problem;
- Riolo function;
- Prisoner's dilemma;
- Michalewicz's functions [8]

$$\begin{aligned} &F1\colon -x\sin(10\pi x)+1 && -2.0\le x\le 1.0\\ &F2\colon \text{integer}(8x)/8 && 0.0\le x\le 1.0\\ &F3\colon x\,\text{sgn}(x) && -1.0\le x\le 2.0. \end{aligned} \tag{4}$$

V. Results of the ANOVA Statistical Study

In this section, a statistical study is performed in order to determine the most relevant parameters in a GA design. The dynamics of GAs are analyzed from two viewpoints. The first is to study the best solution found by the system (maximum fitness). The second viewpoint is the diversity within the population of GAs; to examine this, the average fitness was calculated.

A. Analyzing the Best Solution

In this case, we are seeking the best of all the individuals within the last population examined. What really matters is to achieve a good individual during the execution of the algorithm. For this purpose, it is convenient to select the parameters for a broad-based exploration of the search space within the algorithm. Therefore, all the possible configurations of factors used (the crossover and mutation probability, the

TABLE II
ANOVA TABLE FOR THE ANALYSIS OF THE MAIN PARAMETER IN THE DESIGN OF A GA FROM THE VIEWPOINT OF THE BEST FITNESS

Source	Sum of Squares	D.F	Mean Square	F-Ratio	Sig.level
Main Factors					
Experiment	0.458	5	0.114	8.94	0.000
Population size	15.337	4	3.834	298.91	0.000
Number of Generations	0.407	4	0.101	7.95	0.000
Crossover probability: p_c	1.541	4	0.385	30.04	0.000
Mutation probability: p_m	0.345	4	0.086	6.73	0.000
Type of Crossover	0.081	1	0.081	6.34	0.012
Type of Mutation	1.721	1	1.721	134.18	0.000
Type of Selection	84.07	2	42.031	3277.14	0.000
Significant Interactions					
Cross. prob. & Generation	0.642	16	0.040	2.76	0.000
Population & Selection	4.879	8	0.610	41.96	0.000
Generation & Selection	3.782	8	0.473	32.53	0.000

population size, the number of generations, crossover, mutation, and selection operators) are evaluated for each of the different examples. Table II gives the four-way variance analysis for the whole set of problems studied (the function fitness is normalized in the range [0, 1], for comparison of all the examples). The analysis of variance table containing the sum of squares, degrees of freedom, mean square, test statistics, etc., represents the initial analysis in a compact form. This kind of tabular representation is customarily used to set out the results of ANOVA calculations.

As can be seen from Table II, all the different variables analyzed in the statistical process have a significant influence on the evolution of the GA, in terms of the best solution. Note that the selection operator type adopted, the population size and the type of mutation occurring present the greatest statistical relevance because the higher the F-ratio or the smaller the significance level, the greater the relevance of the corresponding factor. The crossover and mutation probabilities and the type of crossover are not so significant, even though several of the levels employed in this study produce statistically different behavior patterns concerning the output variable.

These conclusions are also confirmed by the multiple range tables for the different factors (see Table III). The multiple range table applies a multiple comparison procedure to determine which means are significantly different. By analyzing the different levels of each of these main factors, it is possible to determine their influence on the characteristics of the GA, enabling levels with the same response repercussion to be grouped homogeneously.

For one of the most significant factors, the selection operator, there exist three homogeneous groups that have no intersection. This means that the behavior of the three operators is different and therefore is statistically significant. Within each column, the levels containing Xs form a group of means within which there are no statistically significant differences. The method currently used to discriminate between the means is Fisher's least significant difference (LSD) procedure. With this method, there is a 5.0% risk of labeling each pair of means as significantly different when the actual difference is zero.

There is a considerable difference between the determinist method and the other two, based on the Roulette Wheel. This is due to the way in which each of the selection methods functions. The deterministic method is the most elitist of the three, due to the way in which it assigns a number of copies in the new population that are directly proportional to the fitness of each individual. As can be seen in Table III, individuals with a high level of fitness survive among the population using the deterministic method; therefore this is the selection operator with the highest value for the Best Fitness. However, as mentioned in [1], [9], and [12], the population of the GA must also possess diversity; otherwise there might occur a premature convergence of the algorithm. With respect to the two Roulette Wheel-based methods, the variation that ensures that at least one copy of the best solution is obtained in the new population produces results that are statistically different from those produced by pure Roulette Wheel; the latter gives better results in terms of the best fitness.

For the population size, five homogenous groups are identified using columns of Xs. Logically, as the number of individuals in the population increases, there is a greater probability that the fitness of the best individual will be higher. It should be noted that several researchers have investigated population size for GAS from different perspectives. Some have provided a theoretical analysis of the optimal population size [3], [12]. Usually, however, most effort was dedicated to the empirical finding of the "optimal" population size [11]. Also, the relationship between replacement operator and population size was analyzed in [10]. The experimental results presented in Table III corroborate the importance of population size in terms of Best Fitness; nevertheless, as described in the following subsection, for diversity within the population of GAs (average fitness and its standard deviation), there exist more influential factors than population size.

Concerning the number of generations, there are four homogeneous groups that are not disjointed, and thus there exist levels of this factor which can be classified within two groups simultaneously. The first of these comprises level 1 (the number of generations is 40% lower than the nominal value) and the last comprises level 5 (an increase of 40%). As can be inferred from these tables, the greater the number of generations or population size, the more possibilities there are of achieving a good individual from the current population, as there exists a greater variety of individuals and these have evolved through several generations.

Table III describes the results for the mutation probability, showing three nondisjoint homogeneous groups (some values correspond simultaneously to two homogeneous groups). The levels of this factor are ordered such that the lowest mean LS corresponds to the lowest mutation probability value used (PM1), while the Best Fitness is obtained with a high mutation probability (PM5). The utility of the mutation operator, together with the probability of applying it to the population elements, is that it provides diversity by introducing extra variability into the population. Due to its behavior, the mutation probability is not so significant from the point of view of the Best Fitness, but it is very important

TABLE III
MULTIPLE RANGE TESTS FOR THE VARIABLE ANALYZING THE BEST FITNESS

Levels of variable "Type of Selection"	LS Mean	Homogeneous Groups		
S1: Roulette Wheel	0.8119	X		
S2: Elitist Roulette Wheel	0.8582		X	
S3: Deterministic	0.9584			X

Limit to establish significant differences: ± 0.004

Levels of variable "Population"	LS Mean	Homogeneous Groups				
POP1: Decrement 40%	0.8325	X				
POP2: Decrement 20%	0.8619		X			
POP3: Nominal Value	0.8871			X		
POP4: Increment 20%	0.8943				X	
POP5: Increment 40%	0.9053					X

Limit to establish significant differences: ± 0.005

Levels of variable "Generation"	LS Mean	Homogeneous Groups			
GEN1: Decrement 40%	0.8591	X			
GEN3: Nominal Value	0.8752		X		
GEN2: Decrement 20%	0.8757		X	X	
GEN4: Increment 20%	0.8802			X	
GEN5: Increment 40%	0.8860				X

Limit to establish significant differences: ± 0.005

Levels of variable "Mutation Probability"	LS Mean	Homogeneous Groups		
PM1: 0.05	0.8701	X		
PM2: 0.1	0.8746	X	X	
PM3: 0.15	0.8771		X	
PM4: 0.2	0.8773		X	
PM5: 0.25	0.8821			X

Limit to establish significant differences: ± 0.005

Levels of variable "Crossover Probability"	LS Mean	Homogeneous Groups			
PC1: 0.4	0.8624	X			
PC2: 0.5	0.8716		X		
PC4: 0.7	0.8793			X	
PC3: 0.6	0.8831			X	X
PC5: 0.8	0.8847				X

Limit to establish significant differences: ± 0.005

Levels of variable "Type of Crossover"	LS Mean	Homogeneous Groups
C1: One-point	0.8743	X
C2: Two-points	0.8771	X

Limit to establish significant differences: ± 0.003

Levels of variable "Type of Mutation"	LS Mean	Homogeneous Groups	
M1: Bit-flip	0.8675	X	
M2: Inversion	0.8850		X

Limit to establish significant differences: ± 0.003

Levels of variable "Type of Experiment"	LS Mean	Homogeneous Groups			
EXP2: Prisoner's dilemma	0.8621	X			
EXP1: Knapsack problem	0.8898		X		
EXP3: Riolo function	0.8949		X	X	
EXP4: Michalewicz function 1	0.9011			X	X
EXP6: Michalewicz function 3	0.9061				X
EXP5: Michalewicz function 2	0.9071				X

Limit to establish significant differences: ± 0.007

when the diversity of the GA is analyzed. For the crossover probability, there are four overlapping homogeneous groups. In this case, the probability levels are not totally ordered, although the lowest values of p_c produce a lower mean LS value. With respect to the crossover operator, it is clear that there do not exist two nondisjoint groups of the operator type. Therefore, the crossover operators that have been designed present a very similar behavior pattern (one- and two-point crossover).

Regarding the mutation operator, from Table III it is clear that there are two homogeneous groups with no intersection, of which the bit-flip-type produces a lower mean LS. The reason for this is to be found in the functioning of the mutation operator. The inversion operator produces a higher average number of changes in the bits among the population elements. This could lead to a greater diversity among the population, although average fitness would be lower; on the other hand, new zones within the search space could be explored, where high fitness solutions might be found.

We also consider the type of experiment performed to be a factor that should be taken into account in statistical analysis. Table III shows there are differences between the various examples, although this factor is not the most relevant in the analysis. As mentioned above, the type of selection, type of mutation, and population size present a higher F-ratio. Nevertheless, it is interesting to analyze the four nondisjoint groups that make up the six levels of the "type of experiment" factor. In the first group, with no intersection with the others, is the Prisoner's dilemma. The second and third groups include the Knapsack, Riolo, and Michalewicz function 1 problems, which do

TABLE IV
ANOVA TABLE FOR THE ANALYSIS OF THE MAIN PARAMETER IN THE DESIGN OF A GA FROM THE VIEWPOINT OF THE AVERAGE FITNESS

Source	Sum of Squares	D.F	Mean Square	F-Ratio	Sig.level
Main Factors					
Experiment	0.255	5	0.0638	5.34	0.000
Population size	1.015	4	0.253	17.4	0.000
Number of Generations	6.033	4	1.508	103.7	0.000
Crossover probability: p_c	0.1312	4	0.0328	3.50	0.017
Mutation probability: p_m	30.25	4	7.562	631.7	0.000
Type of Crossover	0.0335	1	0.0335	3.58	0.058
Type of Mutation	1.85	1	1.855	127.6	0.000
Type of Selection	384.7	2	192.37	13232	0.000
Significant Interactions					
Population & Selection	4.879	8	0.609	41.96	0.000
Mutation prob. & Crossover prob.	0.754	16	0.047	3.94	0.000
Generation & Selection	3.782	8	0.472	32.53	0.000

intersect. The final group includes the three minimization functions presented in Michalewicz.

An important analysis that must be carried out is that of the interaction between the main effects considered Table II. Although all the interactions between the factors have been analyzed, the most statistically significant interactions are between the population size and the selection operator, between the number of generations and the selection operator and between the number of generations and the crossover probability.

B. *Analyzing the Diversity: The Average Fitness*

The methodology employed is the same as that described in the previous subsection. In this case, we are calculating the average fitness of all the individuals within the last population examined. Table IV gives variance analysis for the average fitness.

The type of selection operators, the mutation probability and type, and the number of generations present the greatest statistical relevance. On analyzing the selection operator in the multiple range test table for the average fitness (see Table V), we find that the determinist selection operator presents the highest mean, while the Roulette Wheel produces the lowest. The reason for this is that the Roulette Wheel operator produces the greatest diversification in the GA solutions. Despite the fact that the likelihood of the number of copies of each solution is directly proportional to their fitness, due to the randomness of the process, the number of copies obtained may vary considerably, thus increasing the diversity of the population and reducing average fitness. Nevertheless, the determinist algorithm presents the highest Average Fitness values.

When we analyze the mutation probability factor, it is important to note that the different levels of this parameter produce five disjoint sets; thus it is highly relevant to a study of Average Fitness. The lowest mean value is found with the highest mutation probability (PM5); obviously, this is due to the fact that the latter is associated with diversity among the population, with low fitness individuals lowering the Average Fitness value. Table V shows that the levels of this factor are ordered from highest probability (lowest mean LS) to lowest probability (highest mean LS). It should be noted that the mutation probability is a determining factor in the evolution of the GA in terms of diversity (average and standard deviation), because its function is to explore new areas within the search space by carrying out random changes within the chains of bits.

Regarding the mutation operator, inversion or specular reflection produces a greater number of changes in the bits of the individuals comprising the population, which could result in greater diversity and thus a lower average fitness value than the bit-flip mutation operator.

As expected, when the number of generations increased, so did the average fitness. In this case, the different levels of this factor are ordered from lowest to highest. Of the other factors analyzed, it is noteworthy that population size is not such a relevant factor in average fitness, as it was in best fitness, because the former is obtained for the total number of individuals in the final generation and, statistically, this mean is very similar for the different population sizes tested. Nevertheless, when the population size is very large (POP5), there exists the possibility of exploring a greater number of regions within the input space, although some of these will present a small fitness value. This would result in a decrease in Average Fitness (in Table V, for POP5 the increase of 40% presents the lowest mean LS, and the different levels are ordered from highest to lowest).

Finally, the factors with least influence on population diversity are the type of crossover used and the application probability. For the Crossover Probability factor, there are two groups with a high degree of intersection. This means that the choice between different crossover probabilities does not greatly alter the behavior of the GA with respect to diversity. It should be noted that, on the contrary to the other factors analyzed, the levels of p_c are not perfectly ordered, although there does exist a tendency for a lower probability of the latter to be associated with a probability of higher Average Fitness in the final population. With regard to the type of crossover, just as with Best Fitness, there do not exist two nondisjoint groups of the crossover type and the operators that have been designed present a very similar behavior pattern.

VI. CONCLUSION

A statistical study of the different parameters involved in the design of a GA has been carried out by using the analysis of variance (ANOVA), which consists of a set of statistical techniques that analyze and compare experiments by describing the interactions and interrelations between either the quantitative or qualitative variables (called *factors* in this context) of the system. The motivation of the present statistical study lies in the great variety of alternatives that a designer has to take into account when designing a GA. Thus, instead of relying on intuitive knowledge, it is necessary to gain a more precise understanding of the significance of the different alternatives and their interaction. For example, the selection operator, the number of generations, the mutation probability, and the size of the population within a GA are factors

TABLE V
MULTIPLE RANGE TESTS FOR THE VARIABLE ANALYZING THE AVERAGE FITNESS

Levels of variable "Type of Selection"	LS Mean	Homogeneous Groups		
S1: Roulette Wheel	0.3149	X		
S2: Elitist Roulette Wheel	0.4476		X	
S3: Deterministic	0.6337			X
Limit to establish significant differences: ± 0.004				

Levels of variable "Population"	LS Mean	Homogeneous Groups		
POP5: Increment 40%	0.4556	X		
POP4: Increment 20%	0.4605	X		
POP3: Nominal Value	0.4657		X	
POP2: Decrement 20%	0.4715			X
POP1: Decrement 40%	0.4737			X
Limit to establish significant differences: ± 0.005				

Levels of variable "Generation"	LS Mean	Homogeneous Groups				
GEN1: Decrement 40%	0.4379	X				
GEN2: Decrement 20%	0.4579		X			
GEN3: Nominal Value	0.4687			X		
GEN4: Increment 20%	0.4784				X	
GEN5: Increment 40%	0.4841					X
Limit to establish significant differences: ±0.005						

Levels of variable "Mutation Probability"	LS Mean	Homogeneous Groups				
PM5: 0.25	0.4184	X				
PM4: 0.2	0.4391		X			
PM3: 0.15	0.4611			X		
PM2: 0.1	0.4842				X	
PM1: 0.05	0.5242					X
Limit to establish significant differences: ± 0.005						

Levels of variable "Crossover Probability"	LS Mean	Homogeneous Groups	
PC5: 0.8	0.4572	X	
PC3: 0.6	0.4601	X	X
PC4: 0.7	0.4613	X	X
PC2: 0.5	0.4628		X
PC1: 0.4	0.4644		X
Limit to establish significant differences: ± 0.005			

Levels of variable "Type of Crossover"	LS Mean	Homogeneous Groups
C1: One-point	0.4622	X
C2: Two-points	0.4646	X
Limit to establish significant differences: ± 0.003		

Levels of variable "Type of Mutation"	LS Mean	Homogeneous Groups	
M2: Inversion	0.4563	X	
M1: Bit-flip	0.4745		X
Limit to establish significant differences: ± 0.003			

Levels of variable "Type of Experiment"	LS Mean	Homogeneous Groups			
EXP1: Knapsack problem	0.4319	X			
EXP6: Michalewicz function 3	0.4397		X		
EXP5: Michalewicz function 2	0.4405		X	X	
EXP4: Michalewicz function 1	0.4417		X	X	
EXP3: Riolo function	0.4448			X	
EXP2: Prisoner's dilemma	0.4529				X
Limit to establish significant differences: ± 0.007					

of great importance for the dynamics and quality of the convergence of a system. However, determining these parameters for a particular problem is still an open question, and it is also necessary to bear in mind the impact of the experimental setup on the conclusions derived. It is also important not to isolate or eliminate the different interactions of each of the above factors with the others. In summary, it would be of great interest to perform an analysis of the influence of modifying the main factor involved in the design of the GA, while simultaneously taking into account all the other parameters (application probabilities of the genetic operators, type of selection, mutation, and crossover operators, number of generations, and population size). The methodology based on ANOVA makes it possible to classify different configurations (here called *levels*) that can be used for given factors. Thus it is possible to obtain homogeneous groups of levels with similar characteristics.

One of the goals of this study was to analyze the dynamics of GAs when confronted with modifications to the principal parameters that define them, taking into account the two main characteristics of GAs; their exploration and exploitation capacity. Therefore, the dynamics of

GAs have been analyzed from two viewpoints. We studied the best solution found by the system, to observe its ability to obtain a local or global optimum. The second viewpoint is the diversity within the population of GAs; to examine this, the average fitness was calculated. For the first viewpoint, the most important factors were selection operator, type of mutation, the population size, and the number of generations. It is noteworthy that the type of crossover factor (one point/two points) produces practically identical results, although the application probability (p_c) does present statistically significant differences in the evolution of the GA from the perspective of Best Fitness. Regarding the diversity of the population in the final generations, analysis of the average fitness revealed that the most important factors are the selection and mutation operators and the mutation probability.

ACKNOWLEDGMENT

The authors appreciate the comments from the anonymous referees.

REFERENCES

[1] L. Davis, *The Handbook of Genetic Algorithms*. New York: Van Nostrand, 1991.

[2] R. A. Fisher, "Theory of statistical estimation," in *Proc. Cambridge Philos. Soc.*, vol. 22, 1925, pp. 700–725.

[3] D. E. Goldberg, K. Deb, and J. H. Clark, "Genetic algorithms, noise, and the sizing of populations," *Complex Syst.*, vol. 6, pp. 333–362, 1992.

[4] J. J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, pp. 122–128, Jan./Feb. 1986.

[5] F. Herrera and M. Lozano, "2-loop real-coded genetic algorithms with adaptive-control of mutation step sizes," *Appl. Intell.*, vol. 13, no. 3, pp. 187–204, 2000.

[6] I. Jagielska, C. Matthews, and T. Whitfort, "An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems," *Neurocomputing*, vol. 24, pp. 37–54, 1999.

[7] R. Mead, *The Design of Experiments. Statistical Principles for Practical Application*. Cambridge, U.K.: Cambridge Univ. Press, 1988.

[8] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd ed. New York: Springer-Verlag, 1999.

[9] H. Mühlenbein, "How genetic algorithms really work—Part I: Mutation and hillclimbing," in *Proc. 2nd Conf. Parallel Problem Solving form Nature*, R. Männer and B. Manderick, Eds. Amsterdam, The Netherlands, 1992, pp. 15–25.

[10] M. O. Odetayo, "Relationship between replacement strategy and population size," in *Proc. MENDEL*, P. Osmera, Ed., 1996, pp. 91–96.

[11] M. Ryynänen, "The optimal population size of genetic algorithm in magnetic field refinement," in *Proc. 2nd Nordic Workshop Genetic Algorithms and Their Applications*, J. T. Alander, Ed., 1996, pp. 281–282.

[12] R. Smith, "Population size," in *Handbook of Evolutionary Computation*, T. Bäck, D. Fogel, and Z. Michalewicz, Eds. New York: IOP, Bristol, U.K., and Oxford Univ. Press, 1997, pp. E1.1:1–5.

Pursuit Evasion: The Herding Noncooperative Dynamic Game—The Stochastic Model

Pushkin Kachroo, Samy A. Shedied, and Hugh Vanlandingham

***Abstract*—This correspondence proposes a solution to the herding problem, a class of pursuit evasion problem, in stochastic framework. The problem involves a "pursuer" agent trying to herd a stochastically moving "evader" agent to a pen. The problem is stated in terms of allowable sequential actions of the two agents. The solution is obtained by applying the principles of stochastic dynamic programming. Three algorithms for solution are presented with their accompanying results.**

***Index Terms*—Admissible policy search stochastic shortest path, policy iteration, value function, value iteration.**

I. INTRODUCTION

This correspondence presents the herding problem as a class of pursuit evasion problems. However, in pursuit evasion problems, the terminal state satisfies the spatial coordinates of the pursuer and the evader to be the same [1]–[3]. Meanwhile, the terminal state in the herding problem relates to the evader having reached and satisfied at the same time fixed spatial coordinate point. In another paper [4], we have studied the herding problem in a deterministic setting where the evader is passive. This correspondence studies the stochastic version of the problem where the evader dynamics involves randomness. A classic pursuit evasion game in a stochastic framework was studied [5], but with different terminal state than that of the problem studied here.

This problem can be viewed as a modified version of stochastic shortest path problems. Despite the fact that shortest path techniques, like label correcting algorithms [6] and auction algorithms [7], provide a solution to shortest path problems, these techniques fail to deal with situations like the one we study in this correspondence.

The correspondence is organized as follows. In Section II, we give a detailed description of the system dynamics since it represents the backbone of our proposed solution technique. Based on these dynamics, some characteristic properties of the system are derived in Section III. In Section IV, we introduce a mathematical statement for the system model. Finally, the proposed solution techniques with simulation results and graphs are given in Sections V and VI, respectively.

II. AN $N \times N$ STOCHASTIC PURSUER–EVADER PROBLEM

In this section, we introduce the pursuer–evader problem in an $N \times N$ grid and present the dynamics. The pursuer can occupy one of the $N \times N$ positions, as may the evader. However, they cannot both have the same initial position. The objective of the pursuer is to drive the evader to the pen, (0, 0) position, in minimum expected time.

The state vector at time k, $\boldsymbol{x}(k)$, is determined by the position of the evader and the pursuer, i.e.

$$\boldsymbol{x}(k) = [x_e(k) \quad y_e(k) \quad x_p(k) \quad y_p(k)]$$

Manuscript received February 12, 2001; revised December 27, 2001. This work was supported by the Office of Naval Research under Project N00014-98-1-0779. This paper was recommended by Associate Editor M. Shahidehpour.

The authors are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute of Technology and State University, Blacksburg, VA 24061 USA (e-mail: pushkin@vt.edu).

Publisher Item Identifier S 1094-6977(02)04679-5.